# T&DI Congress 2011

## Integrated Transportation and Development for a Better Tomorrow

ASCE

EDITED BY
Imad Al-Qadi, Ph.D., P.E.
Scott Murrell, P.E.

T&DI
TRANSPORTATION & DEVELOPMENT INSTITUTE

# T & DI CONGRESS 2011

## *INTEGRATED TRANSPORTATION AND DEVELOPMENT FOR A BETTER TOMORROW*

PROCEEDINGS OF THE FIRST TRANSPORTATION AND DEVELOPMENT INSTITUTE CONGRESS

March 13–16, 2011
Chicago, Illinois

SPONSORED BY
The Transportation & Development Institute (T&DI)
of the American Society of Civil Engineers

EDITED BY
Imad Al-Qadi, Ph.D., P.E.
Scott Murrell, P.E.

**Published by the American Society of Civil Engineers**

# Copyright and Disclaimer

ISBN: 978-0-7844-1167-4



American Society of Civil Engineers
ASCE International Headquarters
1801 Alexander Bell Drive
Reston, VA 20191-4400 USA

Call Toll-Free in the U.S.: 1-800-548-2723 (ASCE)
Call from anywhere in the world: 1-703-295-6300
Internet: http://www.pubs.asce.org

# Foreword

"Integrated Transportation and Development for a Better Tomorrow" is the theme of the First Congress of the 15,000 member Transportation and Development Institute of the American Society of Civil Engineers.

This First Congress brings together transportation and development researchers, engineers, planners, designers, project managers, construction managers, and contractors from around the world to discuss integrated strategies focusing on smart development and efficient multi-modal movement of people and goods to support the economic pulse of our nation and the world.

More than 308 abstracts were submitted, of which 166 were accepted for paper submission. Of the submitted papers, only 127 presentations were accepted in five technical tracks:

- Transportation Operations and Safety
- Pavements and Transportation Materials
- Advanced Technologies and Infrastructure Systems
- Airport Planning and Design
- Smart Development and Sustainability

The Congress also includes five pre-conference professional development workshops:

- Retrofitting Transportation Design for Smart Development
- Lessons Learned: Case Studies in Airport GIS Submissions and Electronic Airport Layout Plans (eALPs)
- 2010 Highway Capacity Manual
- FAA PAVEAIR Pavement Evaluation and Maintenance Computer Program
- Young Attendees' Professional Development

Sitting at the transportation crossroads of the nation, where a significant percentage of the nation's goods and people pass through each day, Chicago is an appropriate location for this inaugural event, and three local sites are highlighted sites for guided technical tours:

- CenterPoint Intermodal Center Tour in Joliet
- Chicago's O'Hare International Airport with a focus on the ongoing O'Hare Modernization Program
- NAVTEQ Traffic Studio and R&D "Garage" Tour

Due to this diverse program and exciting location, the inaugural Congress provides a unique opportunity to enhance our knowledge of the planning, design, and construction of transportation projects and therefore continue our contributions to developing integrated strategies for tomorrow's transportation.

We would like to acknowledge our fellow steering committee members, Katie Chou, Ph.D., P.E., M.ASCE and William A. Fife, P.E., M.ASCE, the many individuals who helped by reviewing the submitted papers, the scientific committees, which are listed below for each of the five themes, and the local organizing committee and T&DI/ASCE staff also listed below, for their continuing input and suggestions throughout the planning process.

*Theme Scientific Committees*

**Transportation Operations and Safety:** Robert Bryson **(Co-Chair),** Hani Mahmassani **(Co-Chair),** Ghassan Abu-Lebdeh, Ahmed Al-Akhras, Ray Benekohal, Ronny Chowdhury, John Corbin, Sunanda Dissyanake, Reza Jafari, Mohammad Qureshi, Mitsu Saito, Edward Smaglik, Kevin St. Jacques, Xiaodun Sum, and Richard Weeks.

**Pavements and Transportation Materials:** Shiraz Tayabji **(Co-Chair),** Bill Buttlar **(Co-Chair),** Ross Anderson, Karim Chatti, Katie Chou, Andy Collop, Dave Dratnol, Jeff Gagnon, Katherine Keegan, Wayne Lee, James Mania, Louay Mohammed, Steve Muench, John Naughton, Dave Newcomb, Ray Rawe, Amy Schutzbach, Charles Schwartz, Terry Sherman, Mike Smejkal, Sam Tyson, and Leif Wathne.

**Advanced Technologies in Transportation Systems:** Kelvin C. Wang **(Co-Chair),** Gary Ritter **(Co-Chair),** James Bryant, Louis Cohn, Ronnie Chowdhury, James Delton, Robert Dorer, Jack Finn, Gerardo Flintsch, Christopher Garlick, Peter Gertler, Nasir Gharaibeh, Greg Kruger, Kenneth Maser, Justin McNew, Dave McKay, Stephen Richie, Roger Roberts, Mitsu Saito, Brian Smith, William Vavrik, Karen Van Dyke, Harry Voccola, Yinhai Wang, Ping Yi, Stan Young, and Zhanmin Zhang.

**Airport Planning and Design:** Brian McKeehan **(Co-Chair),** Amiy Varma **(Co-Chair),** Geoffrey Ax, Geoffrey Baskir, Alexandre deBaros, William Fife, Edward Gervais, Geoffrey Gosling, Jim Hall, Michael Hewitt, Michael McNerney, Elisha Novak, Christopher Oswald, Saty Satyamurti, Paul Schonfeld, Jim Smith, William Sproule, David Tomber, Vojin Tosic, and Konstantinos Zografos.

**Smart Development and Sustainability:** Marsha Anderson Bomar **(Co-Chair),** Matthew M. McDole **(Co-Chair),** Adjo Amekudzi, Timothy Jackson, Marsha Kaiser, Alan Kercher, Tim Lomax, Barbara McCann, Roger Millar, Larry Mugler, Sandra Otto, Leanne Redden, and Robert Stevens.

**Local Organizing Committee:** Ross Anderson, Jamil Bou-Saab, Bill Buttlar, Steve Gillen, Bret Johnson, Paul kovacs, David Lipper, Hani Mahmassani, Jeff Roesler, and Yanfeng Ouyang,

**T&DI and ASCE Staff:** Jonathan C. Esslinger, P.E., F.ASCE, CAE, (Director, T&DI), Andrea Baker (Manager, T&DI) and Elaine Watson (Conference Manager, ASCE)

Conference Co-Chairs
Imad L. Al-Qadi, Ph.D., P.E., Dist.M.ASCE (University of Illinois)
Scott Murrell, P.E., M.ASCE (Port Authority of NY and NJ)

# Table of Contents

## Advanced Technologies in Transportation Systems

### Pavement Condition and Inventory



### Mobility, Rail, and Traveler Information



### Monitoring Pavement Systems



### Traffic Monitoring and Control and Travel Time

## Mechanistic-Empirical Design of Pavements



## Testing to Improve Pavement Performance—Part 1



## Pavement Management and Operation

## Improved Pavement Modeling Techniques



## Concrete Pavements



## Sustainable Pavement Technologies



## Testing to Improve Pavement Performance—Part 2

# Transportation Operations and Safety

## Multi-Modal Traffic Operations and Planning



## Strategies to Improve Highway Safety



## Topics in Advanced Traffic Management



## Incident Management and Emergency Traffic Operations

# Road Data Collection and Database Establishment in Developing Countries- Three Case Studies

D. Mihajlovic[1] and I. Jokanovic[2]

[1]Dragan Mihajlovic, Ph.D., Professor, Faculty of Architecture and Civil Engineering, University of Banja Luka, 77 Vojvode Stepe Stepanovica Boulevard, 78000 Banja Luka, Bosnia and Herzegovina, e-mail: dram@inecco.net
[2]Igor Jokanovic, Ph.D., Assistant Professor, Faculty of Architecture and Civil Engineering, University of Banja Luka, 77 Vojvode Stepe Stepanovica Boulevard, 78000 Banja Luka, Bosnia and Herzegovina, e-mail: jokanovic@blic.net

## ABSTRACT

Several decades ago and not long after developed countries, the developing world also recognized the need for the availability of road asset data. However, systematic data collection, establishment of databases and creation of management systems has been failing because of poorly developed networks requiring expansion and modernization, as well as a lack of funding, equipment and staff. The former Yugoslavia, once a leader in the Balkans, began to develop a data collection methodology and to create a road database in the late 1980's; however, the subsequent conflicts brought most of these activities to a halt. There was a marked discrepancy between the levels of achievement in the different parts of the former federal state, although the activities commenced at the same time, with equal resources allocated for the purpose. In this paper, the authors present the latest developments and their experiences gained in three former Yugoslav republics (Bosnia and Herzegovina, Serbia and Macedonia), clearly demonstrating the problems and challenges faced by their local road management agencies, as well as lessons learned from the projects in which they participated.

## INTRODUCTION

Road management is a framework for cost-effective resource allocation, programming and management decisions. It combines engineering principles with sound business practices and economic theory, and provides tools to facilitate a better-organized, logical and comprehensive approach to decision-making. An inventory of road assets and tools to assess their condition and model their performance enables road agencies to identify investment requirements for both short- and long-term improvements as well as decision-making.

Useful and reliable data are crucial instrumental for efficient road management. Road management is a data-intensive process that involves gathering, retrieval, storage, analysis and communication of enormous quantities of data. The information drawn

from these data is essential for the process of decision-making underlying road management. Information is required to evaluate and monitor the condition and performance of the asset inventory, develop performance objectives and measures, identify cost-effective investment strategies and conduct asset value assessments. Information is also required to monitor the economic feasibility of road management. Although it is not necessary to store all transportation system data in a single repository, it is crucial that the data be readily accessible and comparable. Data integration and data sharing, therefore, are vital components of road management.

Data, i.e. inventory, condition, traffic, environmental and cost data, are vital to the success of any management system. Without data, it is impossible to conduct proper analysis and monitoring of a road network. Problems with data are some of the most frequently cited causes of failure of the Road Management System (RMS) or Pavement Management System (PMS).

Road agencies must inevitably collect huge amounts of data. Each data item requires time, effort and money to collect, store, retrieve and use. The task of the staff members of an agency, working on data collection and analysis, is to support its management in strengthening the decision-making process and increasing productivity. At the same time, they should create a balance between those data items that may be considered desirable, interesting or, possibly, useful in the future, and those that are essential.

## ROAD DATA COLLECTION IN THE FORMER YUGOSLAVIA

The history of road database implementation in the former Yugoslavia goes back to the early 1980's, when a group of experts, called the Expert Committee, were appointed at the federal level. In particular, their tasks were:

- developing a common federal basis for the establishment of a road database;
- creating a road database dictionary;
- developing and assisting with the implementation of the road data collection methodology.

Different interests within the road sector in the former Yugoslav republics governed the Committee's success. However, the Committee managed to lay a common foundation for developing the database and creating a related dictionary. The conclusion was to recommend further studies and implementation at the level of individual road agencies in the former republics.

The Highway Institute in Belgrade produced the first version of the computerized road database (RDB) for Serbia, Montenegro, Bosnia and Herzegovina and Macedonia. The then existent documentation and field surveys results (from 1981 to 1984) were used as the source of information on the road inventory (geometry, pavement width, structures and traffic equipment and signalization). This data is still available, mostly as a hard copy of the system in graphic form (usually referred to as

the road cadastre). Similar surveys were conducted in Croatia and Slovenia. This database contained only the most basic data on roads, without any information regarding the condition of pavements and equipment, traffic volume, works and costs, etc.

In 1988, one of the main targets of the World Bank 3rd Sector Loan was the establishment of a road database for all republics of the former Yugoslavia and the introduction of a PMS based on the HDM-III. However, due to the well-known unfavorable historical circumstances, nothing extensive was achieved at that time.

Later on, in the early 1990's, the engineers of the Belgrade Institute for Materials Testing succeeded in designing a road survey vehicle. It was developed based on a number of survey technologies available at that time, but it only contained devices for measuring road geometry and distances and video logging equipment (front and rear cameras). However, other survey devices, like those for the measurement of longitudinal roughness, lateral profile, texture or similar, were not included. The same group of experts generated a database (Informix platform) and populated it with the data gathered in the original 1992-1993 survey. The database was updated in 1996 and 1997 for approximately one-third of the network. In addition, a survey program was added to distinguish between main/trunk roads and regional roads.

In the meantime, the situation has changed dramatically in only two of the former Yugoslav republics. Slovenia and Croatia, utilizing the knowledge of local consulting companies, have developed their own systems and, by integrating the RDB into their management systems (pavement management, bridge management and asset management), they have created conditions for the efficient management of their road networks.

## LATEST DEVELOPMENTS

In the last 10 to 15 years, enormous financial resources have been allocated for the improvement of the condition of roads as well as new developments in the Balkan states. In order to support these improvements and build the capacity of the local road agencies, international financing institutions have landed numerous road management support loans/credits for projects which, almost as a rule, included data collection and road database development.

Of the states created after the breakup of the old federation, Bosnia and Herzegovina (B&H), Republic of Serbia (RS) and Republic of Macedonia (RM) were the ones that did not manage to make improvements to the formerly developed databases and data collection methodologies. This is justifiable on the grounds of insufficient national funding during this period, not only for those services dealing with data collection, but also for the overall maintenance of their state road networks. However, the reason for this can be found in the failure of the international community to fully support changes to the structure and management practices in the road sectors in these countries, but also in their political instability.

Once the required changes became inevitable and the support evident, all the three countries started road management support projects within 5 years after the first project of this kind was launched in B&H (in 2004). Shortly afterwards (2007) a project was started in the RS, followed by the RM as the last (September 2009). All these projects recognize the need to provide relevant information for solid decision-making. Terms of Reference were prepared based on good practices and recommendations from road sectors worldwide, emphasizing the facts presented further in the text.

An information system usually serves several goals of a road agency, i.e. road planning, design and investment, as well as needs connected with the organization of routine maintenance and protection. This complexity involving multiple goals also imposes different, complex requirements on the information system.

Information is needed on items such as network details, traffic and axle loads, costs, road conditions, etc., which in turn require the existence of relevant data on which the information can be based. The need to assess the physical condition, safety, level of service and efficiency of operation of a road system is widely recognized. In addition to knowing the characteristics of the existing system, it is becoming increasingly important to be able to predict the effects that proposed policies are likely to have in future. Such predictive capabilities enable the decision-maker to test alternative courses of action to determine which policies and strategies will be the most effective in accomplishing the desired goals with the resources available.

An RDB should include the following data about roads and structures:

- technical data about roads and structures (road network data);
- information about expenses on roads (works and costs data);
- information about the use of roads (traffic and vehicle data).

An RDB comprises descriptive, numeric, graphic, pictorial and other data about roads and structures thereon. Its purpose is to provide:

- an overview of the condition of roads and structures thereon;
- the data needed for the administration, construction, maintenance and protection of roads and traffic thereon;
- the data needed for the official statistics to be used by administrative bodies, other institutions and individuals.

More often than not, a simple, flexible, scalable and modularized software package is recommended for a usable and sustainable RMS/PMS which can be easily and economically extended over time, as the database structure implemented might be required to incorporate different data needs related to the planning process, different levels of data complexity might be required to be included in the database, and more reports and management functionality might be required with time. The database

structure should be an open one, based on personal computers, and it should serve as a support to any type of the RMS/PMS, with the ability to evolve and add data as required and desired.

**Bosnia and Herzegovina.** The Roads, Bridges and Tunnels Database Project (BHRBTD, 2004-2005) was a part of the World Bank sponsored Road Management and Safety Project carried out between 2002 and 2007. This part of the project comprised the definition and establishment of a database and the survey of the complete B&H network (approximately 8,500 km). It has to be noted that the B&H road network is managed by two different road agencies at the entity level. During the project, these two agencies demonstrated significantly different levels of organization and performance.

At the time of the project commencement, some activities had already been completed by one of the entities (Republic of Srpska). The two most important ones were the establishment of a road reference system (node/section-based) and the complete transformation of the road maintenance system (towards a market-oriented one), which helped a lot during the road survey. As for the other entity (Federation of B&H), its road agency had only recently been established as a separate body under the auspices of the Ministry of Transport and Communications.

As a prerequisite for all activities related to the survey and database, a complete reference system had to be established and a lot of time was spent on reaching an understanding and finding common ground for defining the nodes and sections. On the other hand, the road survey went surprisingly well, with the utilization of a sophisticated ARAN vehicle (Czech Republic), a semi-manual Viziroad system (France) mounted in a passenger vehicle, and some additional pavement survey equipment (Kern device, Dynamic Cone Penetrometer (DCP), Falling Weight Deflectometer (FWD)). Along with the collection of road inventory data, it was also possible to collect data on pavement condition and other road elements.

During the road survey, the consultant developed the database structure and interface with the HDM-4 (Highway Development and Management). Actually, the database is not a real database, but rather a series of MS Excel tables (csv format) connected by means of the French software Routen (commercial-off-the-shelf (COTS) product), a kind of a simplified database management system (Figure 1).

The database contains all the necessary road inventory data (geometry, cross section, drainage, traffic signalization and equipment, structures, settlements, etc.), road condition data (distress type and intensity), as well as data needed for specific analyses (traffic volume and forecast, climate zones, etc.), grouped in 38 tables.

Beside the Routen package, another four (COTS) computer programs were delivered, namely: Sillage (strip diagrams), Carten (simplified GIS), Couplen (reference system management) and Routen-HDM (HDM-4 interface).

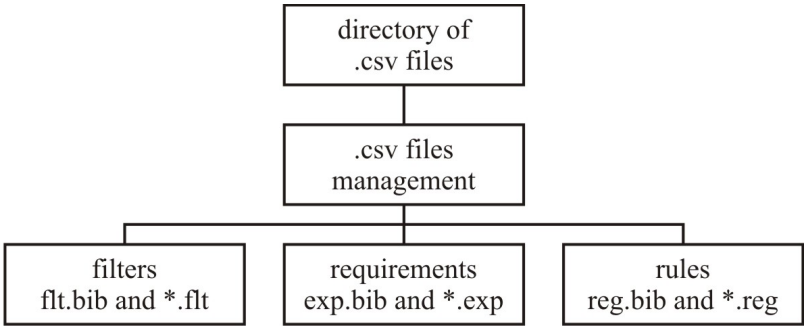



**Figure 1. Database management organization**

The establishment of the database and the road survey were completed relatively fast and smoothly, and relevant training sessions were organized for the staff of the two road agencies. In addition, a data update plan was developed. Immediately following the project completion, the Republic of Srpska Road Agency procured two road survey devices (FWD and longitudinal profiler) for measuring deflections and IRI parameters.

However, some problems arose after the completion of the project:

- a thorough crosscheck of the data revealed deficiencies, mainly associated with road geometry and pavement widths, as well as the absence of data for parts or even complete sections where the survey speed dropped under 30 km/h;
- both road agencies failed to adopt or develop their own survey and data update plans;
- road survey devices are only used on sections planned to undergo major maintenance or rehabilitation work, which does not exceed 200-250 km/year;
- the database has not been updated since the project completion in 2005;
- the data have only been used once by both road agencies to conduct five-year studies on road rehabilitation, and occasionally to present road accident data;
- the staff initially trained to use the database and additional programs have either left the agencies or changed their duties.

**Republic of Serbia.** Similar to the project in B&H, the Transport Rehabilitation Project in Serbia was also financed by the World Bank (RSRDB, 2007-2009) and included the road database component. The principal goals to be accomplished were the review and improvement of the existing reference system, road network survey, database implementation and installation of the HDM-4.

The main problem occurred early in the project. Even though the reference system was created some 15 years ago, not even the road agency was fully familiar with the network, in particular with regional roads. This was ascertained after road sections were identified to be non-existent or still under construction. Nonetheless, the problem was solved by entering corrections into the existing system and introducing data attributes needed to address these problems properly. In addition, most of the

roads were traveled on before the main survey to identify the itinerary and organize and conduct the survey accordingly.

The data on road inventory and pavement condition as well as other road elements were collected using an ARAN vehicle (Czech Republic), a ROMDAS vehicle (RS), and additional pavement survey equipment (Kern device, DCP, FWD, Skid Resistance Tester (SRT), Ground Penetrating Radar (GPR)). The lack of familiarity with the network and very bad pavement condition on regional roads (the ARAN vehicle was forced to travel along unpaved earth roads) resulted in the survey taking longer than estimated at the beginning (9 months versus 5, as planned initially).

The database was built using the MS SQL Server platform as the back-end database, and the HIMS (HDM-4 Information Management System) software as the front-end of the RDB. The HIMS has a very robust and user-friendly interface, and is divided into two components. One of the components deals with the data (Road Information System) and the other, called PMS, carries out data analyses, automatic road sectioning (with manual refinement enabled) and preparation of data for HDM-4 studies. A total of 72 data tables are organized in eight modules, namely: reference system, sections, assets, traffic, pavements, costs, visual documentation and organization.

The structure of the system is shown in Figure 2. The objective of the GIS module included in the HIMS is to provide end users with a facility or platform to view objects such as the road network, province/administrative boundary in a map viewer, etc. It enables the user to generate thematic maps showing the current road condition, inventory, traffic data and assigned PMS programs. The GIS-enabled component does not allow the creation or modification of any GIS objects (layers) such as nodes, roads, sections, etc. However, GIS objects and maps can be imported into the system after being created outside it in an industry standard map objects format such as tab, shape, mif, etc.

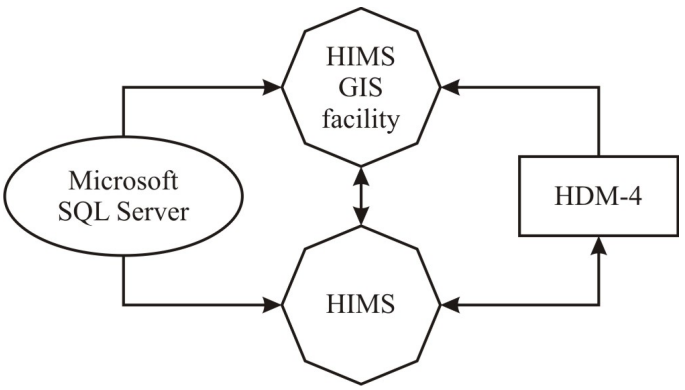



**Figure 2. Organization of the system**

During the project, simultaneously with the road survey and database development, several additional documents were drawn up, such as a data update plan, an HDM-4

calibration study program, a review of the road survey equipment with recommendations for future procurement and road survey contractual arrangements, etc. Also, comprehensive training was provided.

The system was handed over at the end of 2009; some problems remain, though, mainly associated with the staff not being sufficiently familiar with the data (type and quantity) and system architecture.

**Republic of Macedonia.** The project in the RM (RMTA, 2009-2010) is similar to the two presented above, but only to some extent. It is financed by the European Bank for Reconstruction and Development and commenced in the autumn of 2009. The main objective of technical assistance is to help the state road agency with the implementation of the new Law on Public Roads and with the transfer of duties from a maintenance contractor, previously commissioned for both data collection of any type and the database, to the agency. However, these earlier activities were seriously hampered by a lack of funding and trained staff.

The reference system, as it could be expected, was not implemented. It was generated in the early 1990's, but nothing was done in the meantime to update it or put it into effect in the field, or use it to track up the activities on the network. The proposal of the project consultant, which is strongly supported by the client, is to establish a reference system similar to the two discussed previously, and develop a data collection methodology to be executed on a contractual basis (through outsourcing).

The initial survey of the local capabilities and available equipment showed that the local engineers and technicians are not familiar with modern data collection technologies and that road survey equipment is generally not available. The road agency owns an FWD and a roughness device, but they are not in working condition.

The task of the project consultant is to analyze the needs and propose the system organization, carry out the initial basic survey of the road inventory (limited to section lengths, pavement types and widths), develop the database and propose options for data collection and system management. So far, the first activity has been completed and the system proposed is the same as the one implemented in the RS. In addition, a database structure was proposed, with several improvements made upon the experience in B&H and RS.

A comprehensive survey of the network assets, pavement condition and other road elements remains to be conducted by the client. Based on the B&H and RS experience and after the initial RM capacity survey, this can be pointed as the main obstacle for full database population and further activities of data update and system management.

## LESSONS LEARNED

The experience gained from participation in these three projects can be summarized as follows:

- road networks in developing countries are usually not completely known, which is especially the case with secondary and tertiary roads;
- the staff specialized in and trained to collect and process data and manage the database are scarce and often unwilling to cooperate;
- road survey equipment is either old and broken or is not used for regular surveys;
- there are no methodologies developed and/or adopted for data collection, processing, verification and update;
- local consulting companies are not interested in providing these types of services or, if they are, they are usually understaffed or underequipped.

## SUSTAINABILITY

The experience of various countries and road agencies shows that many existing data collection technologies and analysis tools are not used to their full potential to influence investment decision-making. This under-utilization concerns the capabilities of the tools themselves, the credibility of the input data and models and organizational factors. The most successful application occurs when an organization has made a sustained, multi-year commitment to integrating the use of data and tools in its decision-making processes and supporting an internal “champion” to improve those tools over time based on feedback from end-users and decision-makers.

All systems rely on three fundamental components: processes, people and technology. Appropriate funding is an additional prerequisite, although it is not the only one and does not suffice as such. If any of these components are lacking, the system will not be successful. The best technology in the world will ultimately fail if implemented in an environment where there are no people to run it, or where the processes are not in place to utilize it.

## CONCLUSION

The establishment and population of an RDB, in its initial phase, can be a very discouraging task, if there is no previous quality information on all the data required. Probably the most difficult task is selecting what data has to be acquired in order to meet management goals. In any case, it is important to avoid collecting data that is totally irrelevant for the purpose. Experience suggests that it is easy to collect and store data and to develop software for data management, but it is very hard to estimate what data will be needed in upcoming periods, both short- and long-term, and to ensure regular information update.

Additionally, there is a wide range of technologies available to road agencies for measuring road network attributes. There is the challenge of selecting the right

equipment to match the local conditions and opting for the right way in which the data will be used. The technology of road condition survey and monitoring is continuing to develop worldwide. Equipment manufacturers are continually upgrading pavement condition data collection and processing equipment to incorporate the latest technologies. Much of this effort is inspired by the desire for more real-time data analysis, as more agencies collect more data.

Based on the experience from the three projects presented above and a thorough review of current practices worldwide (considering data availability, data collection methodologies, institutional, equipment and staff capacity and capability), the following is recommended to road agencies in developing countries:

- determine the type and amount of data necessary for proper network and program level studies, in line with the database structure and requirements;
- establish a simple but detailed data collection policy;
- establish a sound and constantly growing budget for data collection and management;
- outsource data collection and processing requiring specialist equipment (including seasonal and main pavement inspection and inspection of structures) on the basis of detailed terms of reference for the services and specification of the equipment to be used;
- outsource data management, including verification/validation, database upload, database management, data update, etc., on the basis of detailed terms of reference for the services and specification of the equipment to be used (these two contracts can be joined for a period of at least three years);
- require the establishment of a database and data analysis procedure based on COTS products to allow for rapid prototyping and implementation, ensuring that the road agency has the maximum amount of time during the project to become familiar with the system and understand its use in the organization;
- establish a precise and detailed data quality assurance policy, and require the data collection and management contractor to produce its own quality assurance plan in line with the policy;
- require the data collection and management contractor to train assigned agency staff (lectures and on-the-job training);
- establish and implement a system of technical (internal and/or external) auditing of data and systems, and ensure its recommendations are acted on.

## REFERENCES

Consulting Services for Road Data Base (RSRDB). (2007-2009). *Project Documentation*, Belgrade, Republic of Serbia

Roads, Bridges and Tunnels Database (BHRBTD). (2004-2005). *Project Documentation*, Banja Luka/Sarajevo, Bosnia and Herzegovina

Technical Assistance to the State Road Agency in Implementing the New Law on Roads (RMTA). (2009-2010). *Project Documentation*, Skopje, Republic of Macedonia

# The Effect of Pavement Condition Data Sampling on Project Boundary Selection

C. M. Dean, G. Y. Baladi, T. A. Dawson, and S. W. Haider[1]

[1]Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824, deanchr2@msu.edu

## ABSTRACT

Pavement condition data are either continuously collected or sampled. In sampling it is assumed that the time and costs of data collection can be reduced and the pavement condition of shorter segments represent the conditions of larger sections. In a study sponsored by the Federal Highway Administration, pavement distress data along several miles of roads were obtained from four State Highway Agencies. Each department collects and stores distress data on a continuous basis for each 0.1 mile section along the network. The continuous data were sampled and the impacts on the accuracy of pavement decisions were analyzed. Results of the analyses are discussed herein. It is shown that, for variable pavement conditions, ten percent sampling leads to inaccurate decisions. Pavement sections in need of repair are ignored whereas healthy sections are selected for repair. The costs incurred due to inaccurate decisions could be much higher than the saving incurred by sampling.

Keywords: Pavement distress data, sampling, pavement management, cost-effective decisions

## INTRODUCTION

Some State Highway Agencies (SHAs) collect pavement condition data on a continuous basis while others use sampling techniques to reduce the time and costs of data collection. It is typically assumed that the pavement condition of a pavement surveying length (e.g., one mile, two miles, or longer) can be represented by the pavement conditions along a shorter segment (such as 100, 200, or 500 ft) of the survey section. Two sampling techniques are typically utilized; randomly selected short segment or a fixed short segment (the first 0.1 mile segment of each mile).

Many SHAs videotape the pavement condition along the entire network. Some digitize the images along the entire network (100 percent sampling) and store the data for each one mile or 0.1 mile of pavement. Others digitize the data along a portion of one mile (such as 100, 200, 500, or 1000 ft) and assign the same data to the entire mile. Still, others walk the first few hundred feet of each mile and observe and record the pavement condition for the entire mile.

In a study sponsored by the Federal Highway Administration (FHWA) titled "Optimization of and Maximizing the Benefits from Pavement Management Data Collection", rut depth, International Roughness Index (IRI), and cracking data were obtained from four SHAs; Michigan Department of Transportation (MDOT), Louisiana Department of Transportation and Development (LADOTD), Colorado Department of Transportation (CDOT), and Washington State Department of

Transportation (WSDOT). The four agencies collect sensor (IRI and rut depth) data and digitize the videotapes of the pavement surface conditions along the entire pavement network. They store the data for each 0.1 mile of pavement. The data from the four SHAs were used in this study to simulate sampling. Although continuous data collection yields the most accurate data, it may not be the most cost-effective procedure. Sampling reduces the immediate cost of distress data collection at the expense of data accuracy. Data accuracy affects the quality of the decisions regarding the selection of the optimum project boundaries, timing of the fix, and the fix type to be applied. The costs of lower quality decisions due to sampling could be much higher than the saving incurred by sampling. Hence, the decision to sample or not to sample must be based on the ratio of the cost of lower quality decisions and the saving incurred by sampling. The lower is this ratio the more cost-effective is the data collection technique. A previous study suggested that 70% sampling (data collected from 70% of the pavement surface) will nearly eliminate statistically significant error. Another study developed a methodology for analyzing sample sizes based on error and costs to agencies and users (Ong 2008 and Mishalani 2007). In this study, two sampling techniques were used, random and fixed length. In the fixed length technique, the sample size for each one mile of pavement was increased from 10 to 60 percent in 10 percent increments. The sampled data were then used to study the effects of sampling on the accuracy of the pavement management decisions. Due to space limitations, this paper addresses only 10 percent fixed sampling.

## BACKGROUND

Evaluation of the time series pavement condition data is essential for the application of cost-effective pavement preservation programs (Cafisco 2002). Any pavement data collection practice is limited by the technical, practical, and economical constraints within the SHAs. The National Cooperative Highway Research Program (NCHRP) Synthesis of Highway Practice 222 states that some of the practical constraints can be attributed to the size and variability of the condition of the pavement network. The accuracy of the data is limited by the data collection frequency, by the selected representative samples, and by sample size (Zimmerman 1995). Sampling of pavement condition data requires less time and money than continuous data collection. Sampling is meant to support accurate pavement condition predictions without excessive amounts of data (Robertson 2004). In the NCHRP 2004 Synthesis of Highway Practice 334, 42 states, the District of Columbia, 2 FHWA offices, 10 Canadian provinces and territories, and Transport Canada (airfields) were surveyed regarding pavement condition data collection (McGhee 2004 and Tremblay 2004). It was found that:

1. Most agencies use a semi-automated means for data collection along the entire outer traffic lane every other year.

    a) For pavement cracking; nine agencies survey 100% of the lane to be evaluated, three agencies collect cracking data on a varying sample length basis, five agencies sample 10% to 30% of the roadway using a random

sampling technique, and others, videotape 100% of the survey lane but, for each one-mile, they digitize the data along 50 to 1,000 feet segments.

b) For pavement roughness; many agencies collect data along the entire surveying lane and they report the data for each 0.1-mile interval. The Canadian provinces report the roughness data at 50-meter to 100-meter intervals and the State of Arizona uses a reporting interval of 1 mile.

## EFFECTS OF SAMPLING ON PAVEMENT MANAGEMENT DECISIONS

As stated above, IRI, rut depth, and cracking data were received for segments of nine roads from each of four SHAs. The nine roads from WSDOT are listed in Table 1. In this paper and because of space limitations, analyses of the data from WSDOT for the highlighted three roads in Table 1 (SRID 161, SRID 082, and SRID 005) are presented and discussed. It is important to note that the reason the IRI and rut depth data are used in the analysis is that they are collected by sensors, they are less variables than the cracking data and hence, the effects of sampling are minimal. The effects of sampling on the cracking data are much higher than those on the IRI and rut depth data.

Table 1 Washington State DOT PMS data for segments of nine roads

| SRID | ARM | | SRMP | | Pavement type | Direction | Length (mile) |
|---|---|---|---|---|---|---|---|
| | Begin | End | Begin | End | | | |
| 161 | 8.93 | 17.56 | 8.95 | 17.58 | Composite | B | 8.63 |
| 099 | 2.5 | 13.29 | 6.15 | 16.94 | Composite | D | 10.79 |
| 005 | 59.79 | 79.26 | 59.72 | 79.19 | Composite | I | 19.47 |
| 082 | 100.8 | 122.14 | 100.83 | 122.17 | Rigid | D | 21.34 |
| 005 | 124.71 | 135.6 | 124.65 | 135.54 | Rigid | I | 10.89 |
| 195 | 42.54 | 60.34 | 44.4 | 62.2 | Rigid | B | 17.8 |
| 005 | 263.55 | 273.92 | 263.49 | 273.86 | Flexible | D | 10.37 |
| 020 | 32.61 | 41.17 | 33.01 | 41.58 | Flexible | B | 8.56 |
| 090 | 226.96 | 234.96 | 229.24 | 237.24 | Flexible | D | 8 |

Once again, WSDOT reports and stores the measured IRI, rut depth, and the digitized pavement distress data for each one tenth of a mile along the road network. Such data are labeled herein continuous data. Once again due to space limitations, this paper addresses the impact of sampling on the IRI data along SRID 161, SRID 082, and SRID 005 and the rut depth data along SRID 161 only. In this analysis the IRI and rut depth data of the first 0.1 mile of each one mile of road were assumed to represent the pavement conditions of the entire one mile section. The continuous and sampled data for ten data collection years were analyzed and the results are discussed below.

In order to present the data in a meaningful manner, both the continuous and the sampled data for each distress survey year were normalized relative to the sampled data of that year. Hence, for each distress survey year, the sampled data are represented by the 100 percent line along the entire road as shown in Figures 1

through 3. The three figures depict the normalized sampled and continuous data for SRID 161, SRID 082, and SRID 005, respectively. The data in the figures indicate:

- For SRID 161, the continuous IRI data could be as little as about 25% of the sampled data and as high as about 240%.
- For SRID 082, the continuous IRI data could be as little as about 50% of the sampled data and as high as about 250%.
- For SRID 005, the continuous IRI data could be as little as about 20% of the sampled data and as high as about 300%.

These general trends and results were found to be similar to those of the other six pavement sections in WSDOT and the pavement sections of other SHAs.

Similar analyses were conducted on the rut depth data of SRID 161. The recorded rut depth of the first 0.1 mile of each mile was assumed to represent the rut depth of the entire mile. In addition, the average of the continuous data along the entire mile was calculated. Based on the sampled, the continuous, and the average rut depth data, SRID 161 was divided into 7 uniform sections based on the following rut depth brackets; 0.0 to 0.1, 0.11 to 0.2, 0.21 to 0.3, 0.31 to 0.4, 0.41 to 0.5, 0.51 to 0.6, and greater than 0.6 inch. The percentages of SRID 161 within each rut depth bracket for each of the 10 distress survey years are listed in Table 2 based on the continuous, the sampled, and the average rut depth along each one mile. The data in the table indicate that, for each data collection year, the sampled, continuous, and average rut depth data show significant variation in the percentage of road in each bracket. For instance, the sampled data indicate that, in 1999, 65.5% of SRID 161 was in rut depth bracket 0.21 to 0.3 inch, the continuous data indicates that only 29.9% of the road was in this bracket (more than 50% error) and the average continuous data show only 31%. Similar results can be seen in other years.

The differences between the sampled, continuous, and average per mile data shown in Figures 1 through 3 and listed in Table 2, would likely affect the pavement management decisions. Specifically, the variability may lead to the selection of inaccurate project boundaries. To illustrate this point, the sampled, continuous, and average per mile rut depth data were used to divide SRID 161 into two uniform sections, one with a rut depth higher than 0.4-inch, and the other is lower. These uniform sections mark the boundaries of projects. The reason for using more than 0.4-inch rut is that most SHAs would take action at such rut depth. Nevertheless, Figure 4 shows the uniform section boundaries along SRID 161 for years 1999 through 2002. It can be seen from the figure that:

- The uniform section boundaries based on the average rut depth data are different from those of the continuous data. However,
- The average rut depth data do not completely miss the uniform sections based on continuous data. Indeed, for the year 2002, the boundaries of the uniform sections based on the average and continuous rut data are nearly identical.
- The average rut depth, though, may not identify hot spots as seen in the years 2000 and 2002.

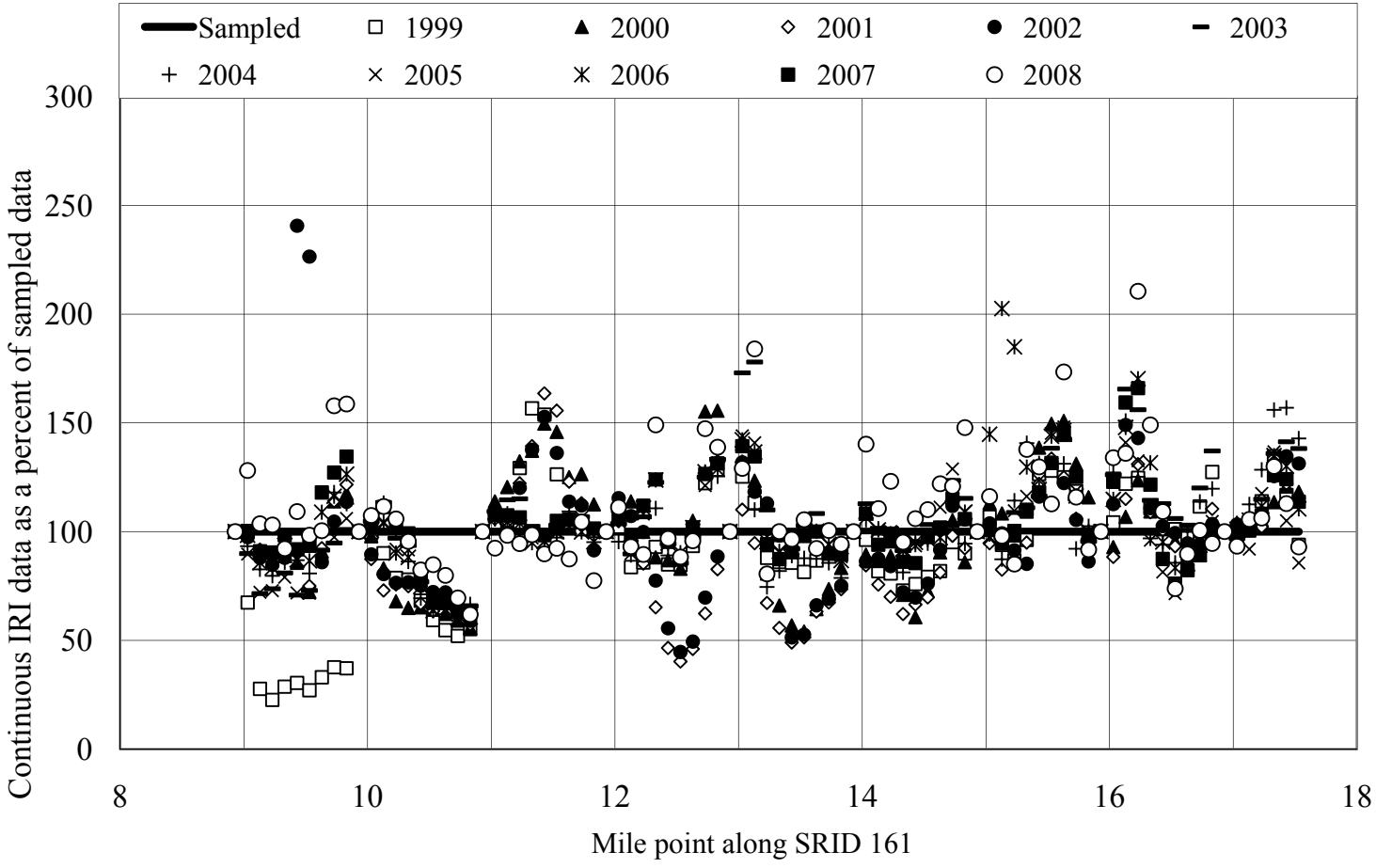



Figure 1 Continuous time series IRI data as percent of the sampled data along SRID 161, Washington

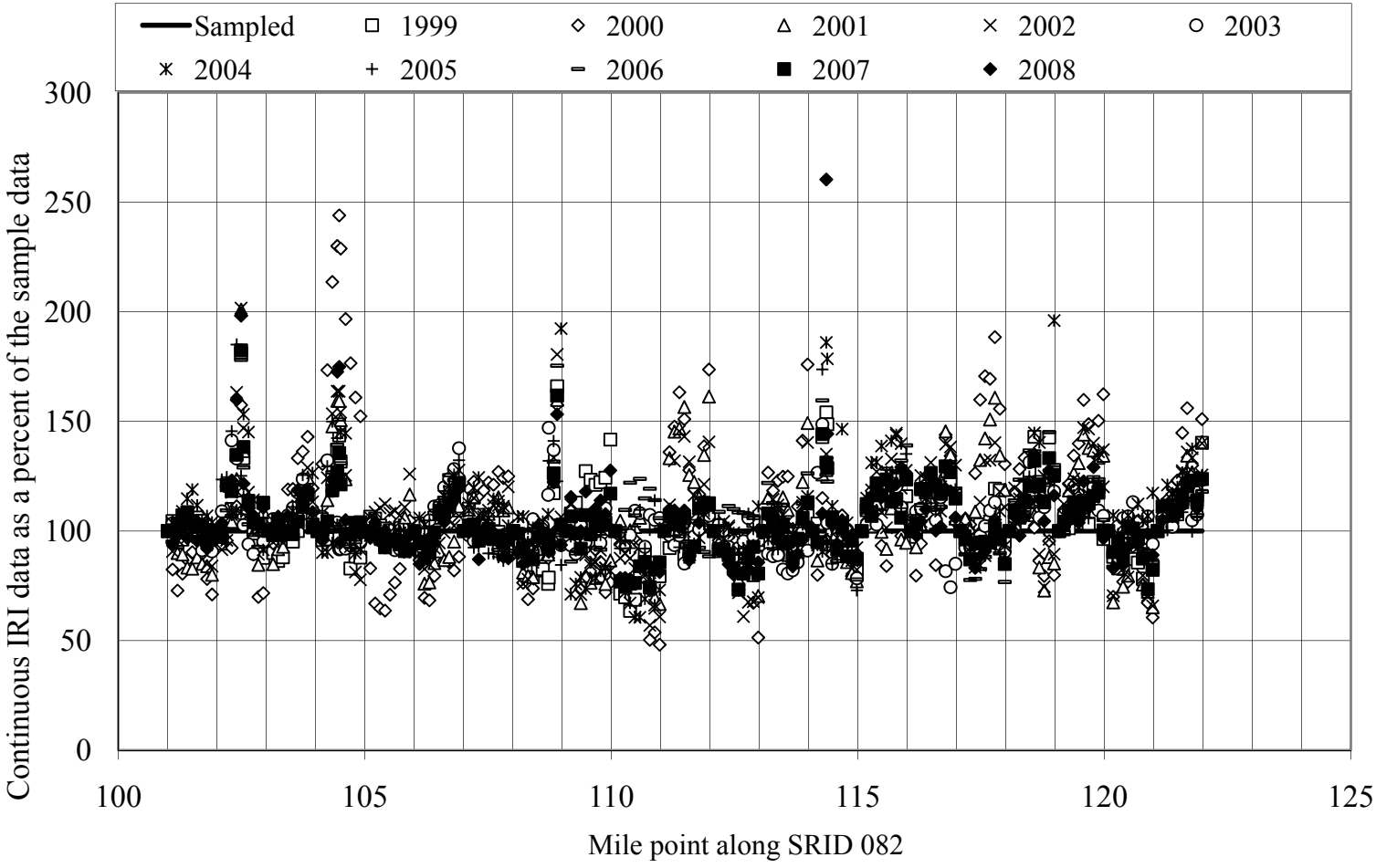



Figure 2 Continuous time series IRI data as percent of the sampled data along SRID 082, Washington

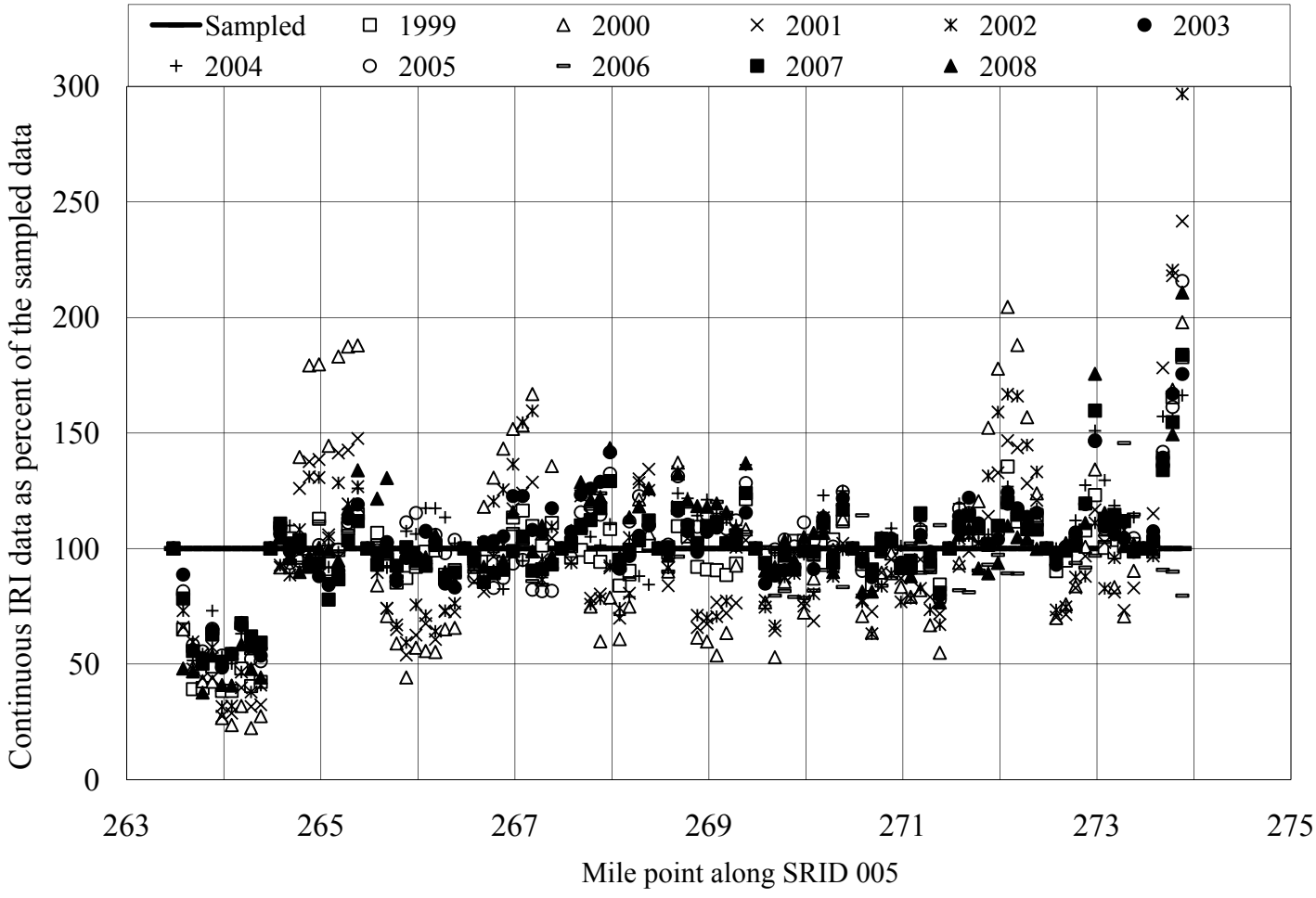



Figure 3 Continuous time series IRI data as percent of the sampled data along SRID 005, Washington

Table 2 Time series rut depth data SRID 161, Washington

| Data type | Rut depth brackets (inch) | Percent of road in various rut brackets in ten years | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
| Sampled | 0.0 to 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 92.0 | 57.5 | 31.0 | 88.5 | 0.0 | 11.5 |
| | 0.11 to 0.2 | 0.0 | 11.5 | 0.0 | 0.0 | 0.0 | 34.5 | 69.0 | 11.5 | 88.5 | 65.5 |
| | 0.21 to 0.3 | 65.5 | 34.5 | 11.5 | 23.0 | 0.0 | 8.0 | 0.0 | 0.0 | 11.5 | 23.0 |
| | 0.31 to 0.4 | 0.0 | 31.0 | 54.0 | 19.5 | 8.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 0.41 to 0.5 | 23.0 | 0.0 | 34.5 | 46.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 0.51 to 0.6 | 0.0 | 11.5 | 0.0 | 11.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | > 0.6 | 11.5 | 11.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | Total | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | | | | | | | | | | | |
| Continuous | 0.0 to 0.1 | 0.0 | 1.1 | 3.4 | 2.3 | 79.3 | 64.4 | 58.6 | 93.1 | 8.0 | 11.5 |
| | 0.11 to 0.2 | 8.0 | 12.6 | 13.8 | 6.9 | 4.6 | 19.5 | 41.4 | 6.9 | 87.4 | 72.4 |
| | 0.21 to 0.3 | 29.9 | 36.8 | 20.7 | 21.8 | 9.2 | 13.8 | 0.0 | 0.0 | 4.6 | 16.1 |
| | 0.31 to 0.4 | 29.9 | 25.3 | 35.6 | 37.9 | 5.7 | 2.3 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 0.41 to 0.5 | 21.8 | 9.2 | 24.1 | 20.7 | 1.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 0.51 to 0.6 | 6.9 | 10.3 | 2.3 | 10.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | > 0.6 | 3.4 | 4.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | Total | | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | | | | | | | | | | | |
| Average per mile | 0.0 to 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 80.5 | 69.0 | 65.5 | 100 | 0.0 | 0.0 |
| | 0.11 to 0.2 | 0.0 | 11.5 | 0.0 | 0.0 | 0.0 | 11.5 | 34.5 | 0.0 | 100 | 100 |
| | 0.21 to 0.3 | 31.0 | 42.5 | 46.0 | 34.5 | 11.5 | 19.5 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 0.31 to 0.4 | 34.5 | 23.0 | 31.0 | 42.5 | 8.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 0.41 to 0.5 | 34.5 | 11.5 | 23.0 | 23.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 0.51 to 0.6 | 0.0 | 11.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | > 0.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | Total | | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

It is important to remember that the boundaries of uniform sections are the boundaries of candidate projects. As shown in Figure 4, the sampled data often identify different project boundaries. Given that continuous data are the most accurate and complete data, SHAs using sampled data would not accurately identify candidate project boundaries. Hence, they either spend money where is not needed, or neglect pavement sections that requires action. To illustrate this point and its impact on the cost of data collection, consider a SHA that controls 25,000 lane-mile of pavement, and spends annually $300,000,000 on pavement actions (preservation, maintenance, rehabilitation, etc.). The cost of collecting and digitizing continuous distress data (excluding IRI and rut depth) every year is about $1,000,000 (based on $80.00 per lane-mile). Suppose that the agency decided to sample the distress data (videotaping the entire system and digitizing the data based on a sampling technique), the maximum saving would be about $500,000. As shown above, the resulting sampled data could be different from the continuous data by as much as 250 percent. If such

difference causes only 2 percent error in the project boundaries, the true cost of sampling will be 2 percent of the budget or $6,000,000.00. It is absolutely true that, in the above illustration, data sampling precipitated $500,000 in saving. The minimum hidden cost of sampling is $6,000,000.00.

## CONCLUSION

Based on the results of the data analyses of three pavement sections in Washington that are presented in this paper and on the results of the analyses of other pavement sections that are not included in the paper because of space limitation, the following conclusions were drawn:

1. Data sampling accurately assesses uniform pavement condition; however, it does not represent variable pavement conditions.
2. The use of sampling to indentify uniform pavement sections results in improper identification of project boundaries. Hence, the sampling data cannot be used to arrive at accurate and cost-effective decisions.
3. It is evident that sampling reduces the direct costs and time for data digitization. Yet, the hidden costs of sampling could be several folds higher than the saving.

## ACKNOWLEDGEMENT

The authors wish to thank the Federal Highway Administration for their financial support and for providing the necessary materials for the study. To the Pavement Management Engineers and staff of the MDOT, WSDOT, LADOTD, and CDOT many thanks for your cooperation in providing the pavement distress data for the analyses. Finally, many thanks to Ryan Muscott, Nick Tecca, Matthew McCloskey, and Corbin St. Aubin for their efforts in the data mining phase of this study.

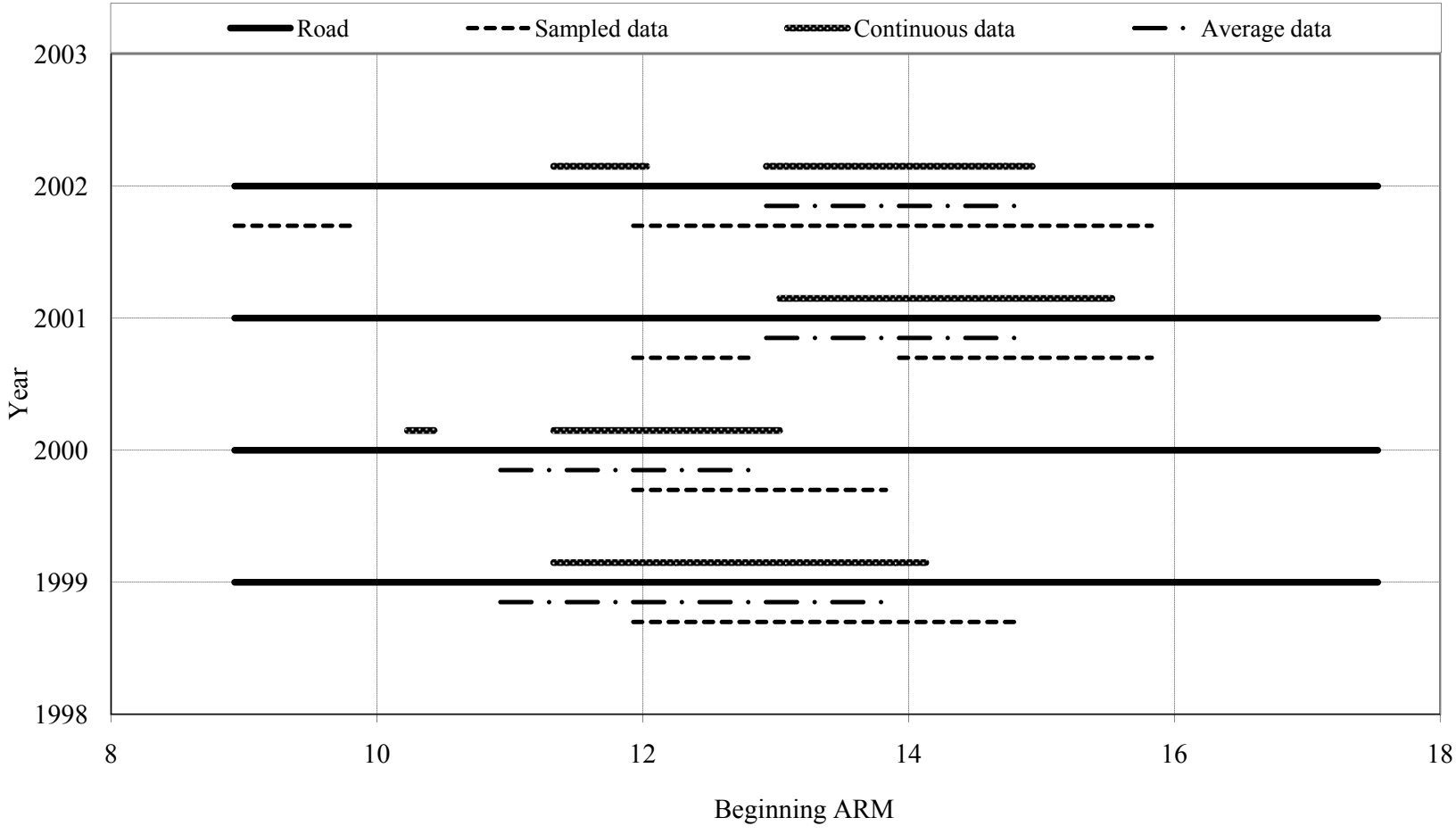



Figure 4 Differences of project boundaries based on continuous, sampled, and average rut depth along SRID 161, Washington

## REFERENCES

Cafisco, S., A. Di Graziano, H.J. Kerali, & J.B. Odoki. (2002). Multicriteria Analysis for Pavement Maintenance Management. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1816,* Transportation Research Board of the National Academies, Washington, D.C.

McGhee, Kenneth H. (2004). NCHRP *Synthesis of Highway Practice 334: Automated Pavement Distress Collection Techniques,* Transportation Research Board, National Research Council, Washington, D.C.

Mishalani, R. G. and L. Gong. (2007). "Evaluating the Impact of Pavement Condition Sampling Advances on Life-cycle Management", Transportation Research Board (TRB), Washington, D.C.

Ong, G. P., S. Noureldin, and K. C. Sinha. (2008). "A Methodology to Evaluate Quality of Pavement Surface Distress Data Collected from Automated Techniques", Transportation Research Board (TRB), Washington, D.C.

Robertson, N.F., Latimer, D.P. and Nata-atmadja, A. (2004). *A Study of the Impact of Road Segmentation Schemas on Predicted Maintenance Investment Outcomes.* Sixth International Conference on Managing Pavements, Brisbane, Australia.

Tremblay, Guy, Mathieu Grondin, Diane Leroux, & Jean Carrier. (2004). *Efficient Way to Measure Pavement Distress Manually: Ministère des Transports du Quebec Experience.* Sixth International Conference on Managing Pavements, Brisbane, Australia.

Zimmerman, Katherine A. (1995). NCHRP *Synthesis of Highway Practice 222: Pavement Management Methodologies to Select Projects and Recommend Preservation Treatments,* Transportation Research Board, National Research Council, Washington, D.C.

# Rapid 1-mm Mapping of Runways in an Integrated Environment

Kelvin C.P. Wang[1], Vu T.D. Nguyen[2], and Quintin B. Watkins[3]

***Abstract***

A critical component of runway pavement survey is visual inspection of surface conditions. The dominant practice today is still based on manual survey, which causes delays, poses safety hazard, labor intensive and error-prone. In recent years, Hartsfield-Jackson Atlanta International Airport (HJAIA) has been using digital imaging technology to acquire pavement images for condition survey purposes. Initially, 3-mm resolution digital technology was used. In the most recent iteration, the developer used 1-mm technology and laser imaging to cover entire runway pavements. In this paper, the authors introduce data collection techniques on lasers and imaging. The laser based imaging system works day or night without shadowing problem or the impact of sun light. A vehicle is used to collect 14-ft wide continuous images of a runway at highway speed. A methodology is used to stitch images of multiple data collection runs for the same runway, so that a single uniform virtual imagery is formed for one entire runway. GPS coordinates are also established for multiple points on the images. A procedure to conduct Pavement Condition Index (PCI) evaluation on runway pavement is presented. Tests are conducted to demonstrate that this methodology is superior to manual approach. Condition survey of runways based on the 1-mm images can be substantially more accurate and effective than any manual approach.

## INTRODUCTION

Traditional runway pavement inspection is conducted through manual observation in the field. It is hazardous, error-prone, lack of consistence, and time and cost consuming. Many existing systems can automatically collect the data (Wang, 2000). However, they share common characteristics of limited database automation level. Even though pavement agencies of roadway and airports share technologies in data collection and interpretation, there exist substantial differences in using and analyzing the data sets. For example, roadway is normally analyzed based on its transverse width of 12-feet for one lane, while in airports, runways as the critical pavement asset has a width of over 100-feet. If the surface is made up of concrete, 25-feet wide slabs may be used to form the runway surface.

Therefore, it is important to present the entire runway as a seamless image to facilitate the condition survey of runway pavements. This paper describes the application of laser imaging technology to acquire 1-mm resolution images covering 13-feet wide passes of runways. Technical approaches and algorithms for merging images of multiple passes of a runway are presented to form a virtual runway. Positioning information from the linear referencing based Distance Measurement Instrument (DMI) and Differential Global Positioning System receiver are used as baseline references for merging the images. Imaging processing techniques are employed to merge longitudinally and transversely the collected images to form an integrated database environment for the evaluation of runway pavement surface. Distress survey based Pavement Condition Index (PCI) evaluation with stitched virtual runway images is also introduced in this paper. This new technology (MHIS-Airport) was tested on concrete runways of Hartsfield-Jackson Atlanta International Airport (HJAIA).

[1] Department of Civil Engineering, University of Arkansas, Fayetteville, U.S.A. kcw@uark.edu

[2] Department of Civil Engineering, University of Arkansas, Fayetteville, U.S.A

[3] Prime Engineering, Inc., Atlanta, USA, formerly Hartsfield-Jackson Atlanta International Airport (HJAIA)

## DESCRIPTION OF HARDWARE SYSTEM

The Digital Highway Data Vehicle (DHDV) has evolved into a new platform with laser based imaging technology for pavement surface data collection. The DHDV is multi-functional and includes a sub-system for pavement surface imaging, a sub-system for Right-Of-Way imaging, and a sub-system of laser road profiling. Figure 1 illustrates the basic components and data flow of DHDV.

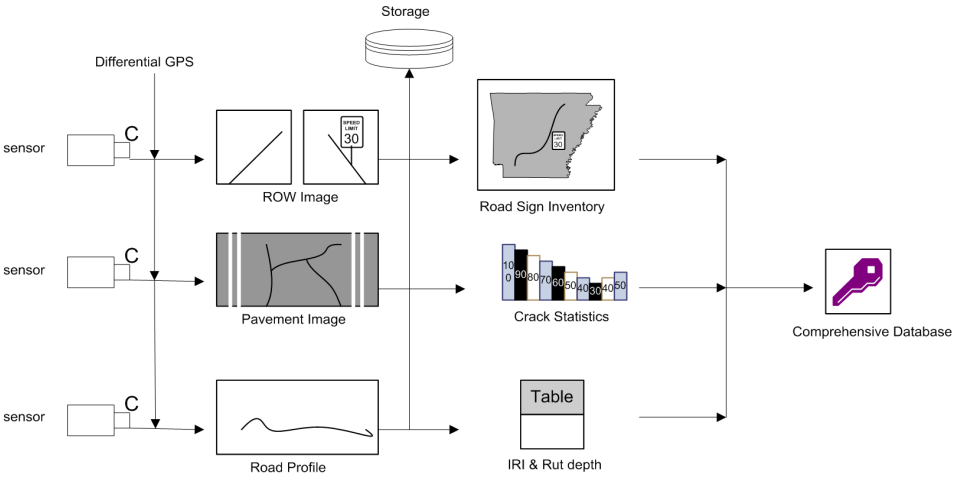



Figure 1. Dataflow of the DHDV System

For several decades, the data collection industry of roadway pavements struggled to provide high-quality and shadow free images. The implementation of digital image collection in recent years did not help ease the problem. Since early 2006, laser was successfully implemented as the illuminating device to capture pavement images. The underlying principle is to illuminate the pavement surface with a laser line light within a narrow spectrum that sun light has limited energy influence. Camera lens is outfitted with a filter that only allows energy in the narrow spectrum to pass through. The end result is that the camera can only receive information reflected from the pavement surface in the narrow spectrum. The two major advantages of this technique are (1) shadow free images are obtained all the time, day or night, (2) energy consumption is limited to 200 watts or less. Systems based on traditional lighting techniques require thousands of watts of energy, still not able to rid of the shadowing problem. In Figure 2, the laser system uses both high speed/high resolution line-scan cameras in conjunction with high power laser line projectors that are aligned in the same plane in a symmetrically crossed optical configuration.

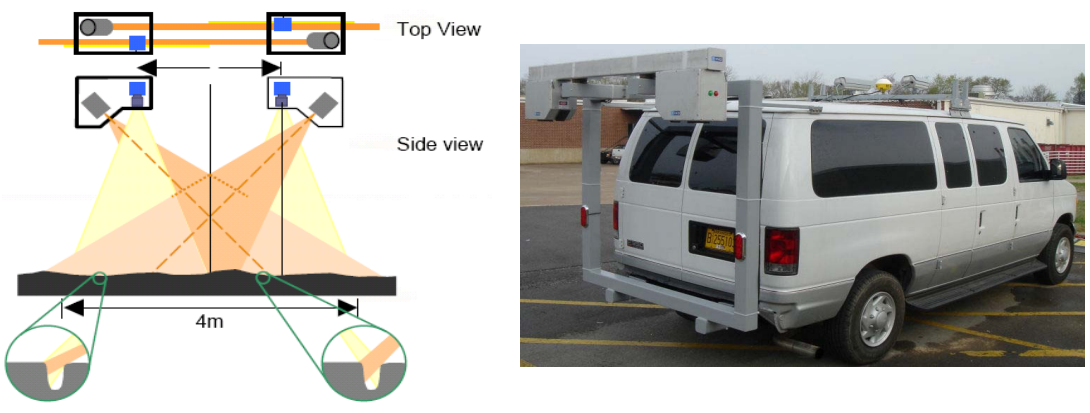



Figure 2. DHDV with LRIS in Operation

DHDV was used in multi-pass operation to collect runway images for HJAIA (Wang 2007). Typically, a 25-feet wide slab was used as the runway-lane. DHDV ran over the same slab

from one end of the runway to the other end with 3 multiple runs: left, right, and the middle of the slab. There are a total of six slabs for one runway, resulting in a total of 18 runs of the DHDV over the entire length of a runway in HJAIA. Figures 3 and 4 show before-stitching and after-stitching images from six passes for two slabs.

## TECHNICAL APPROACHES TO MERGING RUNWAY IMAGES

The difficulties of stitching the images for one runway are:

1) The size of the data: Each image has the size of 4-meter by 2-meter, resulting in 8 MB*18=144 MB raw image size just for a single image. To cover one entire runway, the total storage for all raw images can run up to several 100 gigabytes. The current compression for images is about 8:1. Custom software shall be made to read and stitch the images, as conventional packages of imaging software are not designed to handle this type of work.

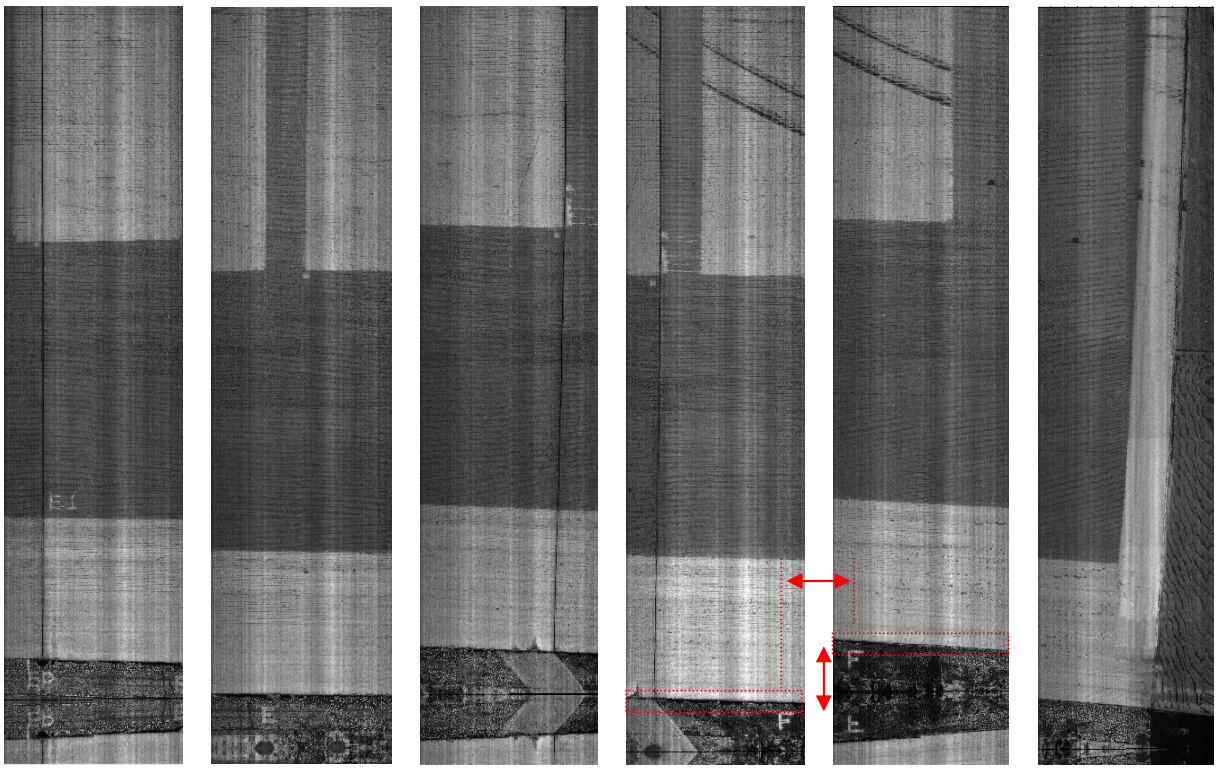
2) The paces of different data collection passes are not exactly equivalent: for a single pass data collection, an accurate alignment for the start location does not mean that the images collected follow a perfect alignment for the rest of the pass in both longitudinal and transverse directions. For example, the routes of the passes or runs are not perfectly parallel to each other. Due to variation of vehicle wandering on the runway, the total numbers of scanned lines by the laser-illuminated cameras with multiple passes are not always consistent. Consequently, for the same longitudinal slab path, three different collection passes will give different longitudinal lengths.

3) Overlap or gaps among adjacent data collection passes: due to the need to ensure complete coverage of runway surface, three passes are used to cover each 25-ft slab path. In most cases, there exist overlaps of pavement surfaces among adjacent images in the transverse direction. However, due to driver error or simply vehicle wandering, gaps among images can also be observed. An essential challenge in stitching the images is to identify the overlaps and gaps.

4) Infeasible manual stitching: theoretically, issues of misalignment, overlapping, and gapping of the images can be adjusted manually. Due to the sheer size of the data, and large quantity of image numbers, it would be impossible to conduct this stitching work manually in a computer with any imaging software. First, both longitudinal alignment and transverse alignment are required. And the alignment of the images from neighboring runs needs to be adjusted periodically, such as every several tens of feet.

5) Automation by imaging techniques: features on the runway images are very similar. A critical power of computer based imaging is to identify common features and conduct relevant processing. There are jut not enough unique features available from runway pavement images.

The runway images in individual passes are captured seamlessly along longitudinal direction during data collection. To stitch or merge all the images from the 18 passes for an entire runway, the longitudinal and transverse offsets for individual passes need to be determined. This is achieved by using techniques in image processing and the differential GPS positioning data.



| Runway 8L WE E1 | Runway 8L WE E2 | Runway 8L WE E3 | Runway 8L WE F1 | Runway 8L WE F2 | Runway 8L WE F3 |
|---|---|---|---|---|---|

Figure 3. Airport Runway Images of Six Passes for Two Slabs
(Arrows show the longitudinal offset and transverse offset among the individual collections)

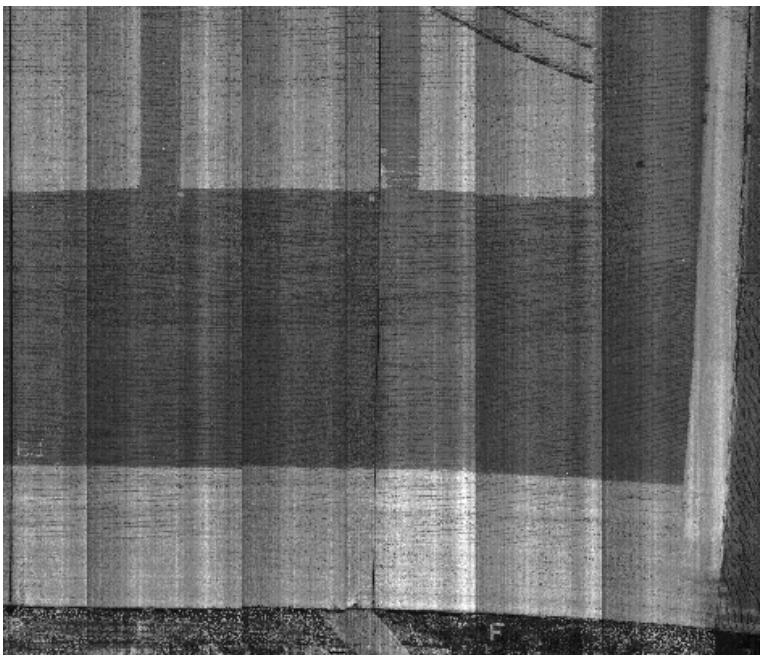


Figure 4. Runway Image after Stitching for Two Slabs (50-ft Wide)

Longitudinal offset of images from two adjacent passes is determined by locating the same features on the images from the two passes, such as slab joints. GPS coordinates related to each joint was then used to identify the unity of the joints. Transverse offset is more difficult to determine. In the first iteration of the software application, it depended on manual adjustments for alignments at visible feature points. The relative distance calculated from the GPS coordinates in the two parallel collections is used to adjust the alignment offset in the final implementation.

The longitudinal stitching procedure is as follows:

1) Segmentation

An adaptive thresholding method is used to conduct the segmentation. An initial threshold (T) is chosen from the histogram. Then the image is segmented into object and background pixels, creating two sets:

$G_1 = \{f(m,n) : f(m,n) > T\}$ (object pixels)

$G_2 = \{f(m,n) : f(m,n) \le T\}$ (background pixels)

$f(m,n)$ is the value of the pixel located in the $m_{th}$ column, $n_{th}$ row.

The average of each set is computed:

$m_1$ = average value of $G_1$

$m_2$ = average value of $G_2$

A new threshold is created that is the average of $m_1$ and $m_2$

$$T' = (m_1 + m_2)/2$$

Using the new threshold $T'$ to replace the initial threshold $T$ is repeated based on the above procedure until convergence is achieved when $T'$ matches $T$. Figure 6 shows the original grayscale runway image and the binary image after the segmentation. Normally after this step the slab joint or other distinctive objects will be segmented from the background.

2) Projection

From the binary image in Figure 5, projections are conducted horizontally and vertically, with tolerance within a few degrees. Slab joint, as a simple feature, will easily be recognized considering the projection, the direction of the object, and the uniformity along a specific direction. However, it is difficult to identify the unity of one joint by image processing only.

3) GPS Coordinates

Differential GPS coordinates of each detected joint can be obtained by interpolation with known GPS data gathered during each data collection pass. Comparing the GPS coordinates of two candidate joints will help identify whether they are the same joint on the runway. Once the joints are identified as one, the longitudinal offset is determined for that specific location. Differential GPS coordinates obtained in the open field on runways are shown to have very good precision.

4) Frequent Adjustment

Based on experimentation, proper offsets determined at one location for stitching may not be applicable to another location a few hundred feet away, primarily due to differences in pixel counts for the same length of runway because of vehicle wandering. Therefore, adjustments to obtain proper offsets need to be made for the images at many locations along the runway.

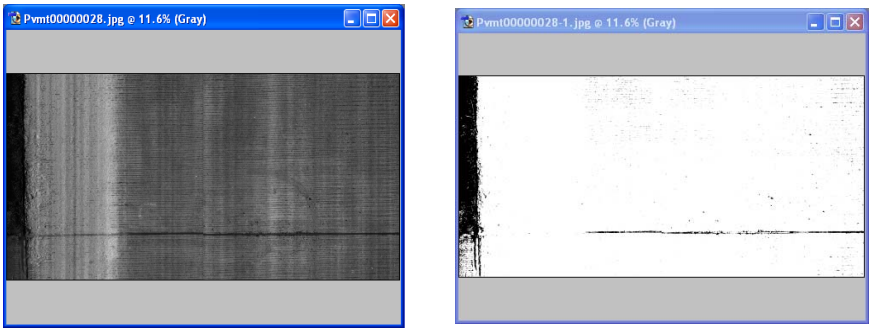



Figure 5. Segmentation

The transverse stitching procedure is as follows:

1) Initial Alignment in Featured Points
   Initial alignment relies on a manual effort. The user has to drag the mouse and align the images of adjacent passes at an easily recognizable location. If data collection passes are sufficiently parallel, the transverse offset obtained from two adjacent passes is applicable for the entire two passes.

2) For some passes with transverse joints present, the same technique used for the longitudinal offset can be used to determine the transverse offset.

3) If data collection passes are not entirely straight, such as vehicle wandering on the runway which is common, GPS coordinates are used to adjust the alignment. Figure 6 shows the drawing of the GPS coordinates of data collection passes. The transverse offset of two parallel routes can be calculated to adjust the alignment.

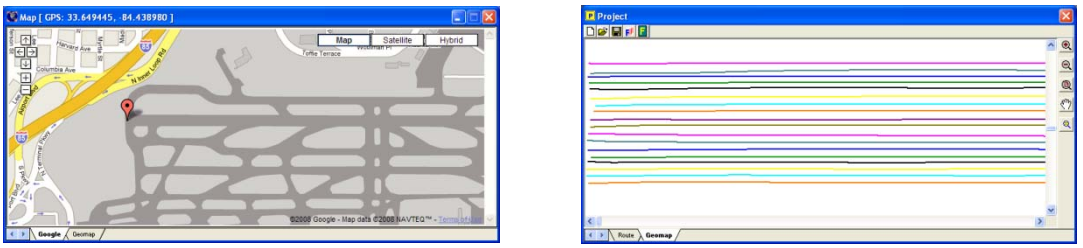



Figure 6. GPS Coordinates from Multiple Data Collection Passes

**AIRFIELD PAVEMENT DISTRESS SURVEY**

Various types of pavement distress for bituminous and concrete pavements generally fall into one of the following broad categories: cracking, distortion, disintegration and loss of skid resistance (FAA, 2007). Based on stitched virtual runway image, distresses will be collected visually or manually for PCI surveys. The following procedure is to collect distresses for concrete pavement based on data collected at HJAIA. A grid standing for a joint grid is first generated as in Figure 7.

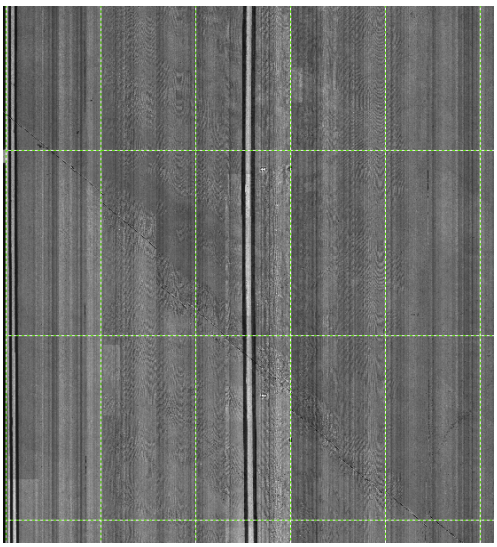


Figure 7. Joint Grid

Then, each considered slab is evaluated by distress type and its severity with distress evaluating table as in Figure 8. Distress information will be saved into database and can be displayed for further reference as in Figure 9.

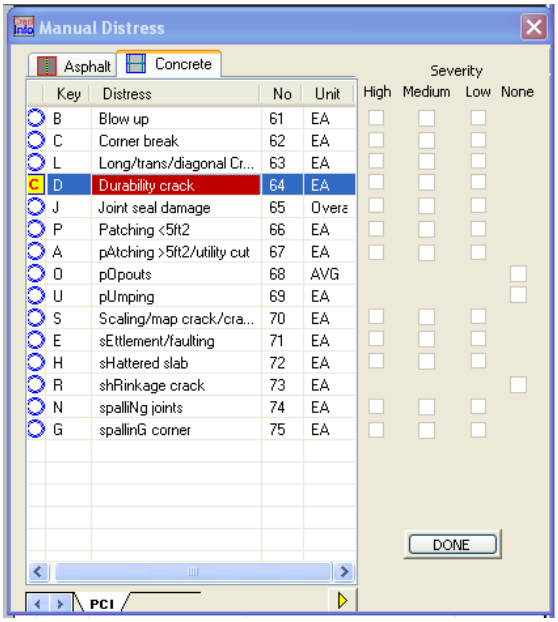



Figure 8. Distress Evaluating Table

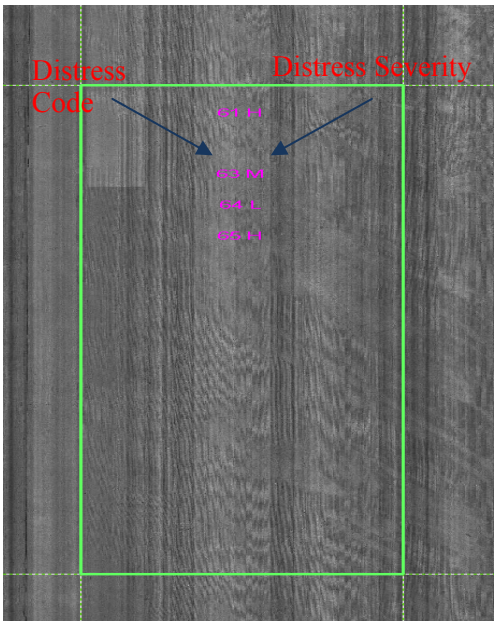



Figure 9. Display Distress on Stitched Image

**AUTOMATED AIRFIELD PAVEMENT CONDITION INDEX (PCI) EVALUATION**

After 1-mm laser images are collected and stitched, PCI frames are reviewed manually, using a computer monitor to visually determine distress type, severity, and quantity. The following steps show how to implement automated PCI evaluation with the software application MHIS-Airport.

Step 1: Assign sample unit on slab grid as in Figure 10.

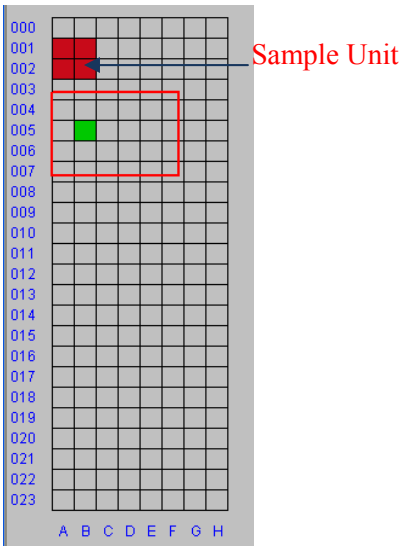



Figure 10. Assign Sample Unit on Stitched Image

Step 2: Evaluate slab distress of sample unit as in Figure 11.

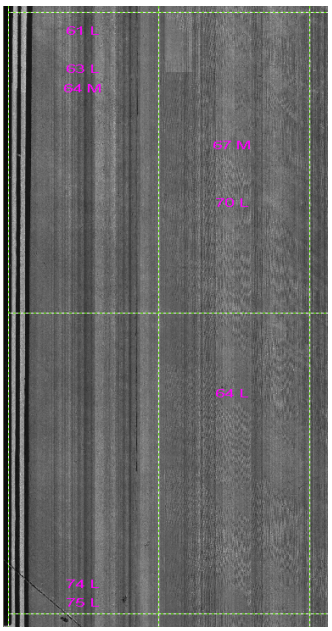



Figure 11. Assign Distress of Sample Unit on Stitched Image

Step 3: Calculate PCI.
Based on distress information, PCI value of an examined section is automatically calculated by using American Society for Testing and Materials (ASTM) standard D 5340-04. The deduct values, corrected deduct values and PCI value are shown in Figures 12 and 13.

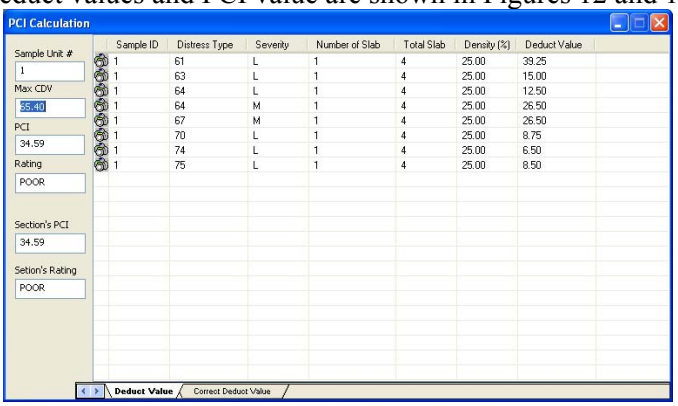



Figure 12. Deduct Values of Examined Sample Unit

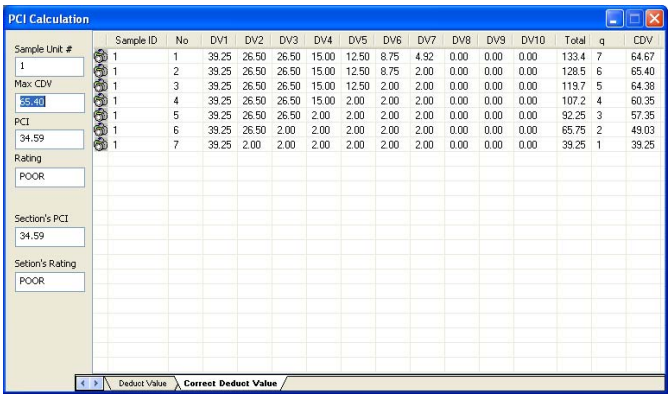



Figure 13. Correct Deduct Values of Examined Sample Unit

## CONCLUSIONS

A single virtual image of an entire airport runway can be reproduced in the MHIS-Airport software with 1-mm/pixel resolution. This application was successfully used for surveying 150-ft wide runways at HJAIA. This new technology is designed to assist airport pavement engineers to rapidly obtain high-resolution runway surface images, collect distresses and conduct PCI condition survey. Based on past surveys at HJAIA, it took about four hours to collect images of two full-length runways and a few more hours to produce a fully stitched single virtual image of a runway. The research team has recently developed sensor-level 1-mm resolution 3D imaging technology for pavement survey. It is anticipated that such technology will be used in the near future on airport pavement survey and evaluation.

## REFERENCES

Wang, K.C.P. "Design and Implementation of Automated Systems for Pavement Surface Distress Survey", ASCE Journal of Infrastructure Systems, Vol.6, No1, March, pp. 24-32, 2000

Wang, K.C.P, Hou, Z.Q, and et al, Automated Imaging Technique for Runway Condition Survey, Proceedings of 2007 FAA Worldwide Airport Technology Transfer Conference and Exposition, April 15 – 18, 2007 • Atlantic City, New Jersey

FAA (2007), Guidelines and Procedures for Maintenance of Airport Pavement, Federal Aviation Administration, U.S. Department of Transportation, Washington, D.C.

# Strategies for Flexible Pavement Rehabilitation based on Case-based Reasoning

Lin Li[1] and Kelvin C. P. Wang[2]

**Abstract:** Maintenance and rehabilitation strategies are critical components of a pavement management system. The on-time and effective rehabilitation not only improves the pavement performance, but also prolongs the service life of an existing pavement. The process to select feasible rehabilitation alternatives from a pool of available alternatives for pavement rehabilitation can range from using simple judgment to relying on a decision tree based expert system. However, the process of knowledge retrieval in an expert system is tedious and rule revision is not convenient. Therefore, a new technique called Case-Based Reasoning (CBR) is used in this paper to reduce the complexity of the problem solving process by relying on solution parameters that were successfully applied in the previous similar problems. The solution parameters are contained in an active case base that is continually updated with new solution parameters when new problems are solved. In this paper, a CBR system is proposed for determining pavement rehabilitation strategies by comparing the similarity between available solutions and new solutions.

**Keywords:** Flexible pavement rehabilitation; Case-Based Reasoning (CBR); Case representation; Case retrieval

## Introduction

It is well-known that types and frequency of rehabilitation on pavements can significantly influence pavement performance and cost (Haas, et al., 1978). Indeed, the need for rehabilitation of our highways, streets and airports has never been greater. There is an increasing demand to make better use of shrinking resources for pavement rehabilitation and maintenance with a variety of strategies or alternatives. Selecting the most cost-effective strategy for pavement rehabilitation continues to be a significant challenge to the transportation professionals (Haas, et al., 1994). There are many methodologies for determining strategies of pavement rehabilitation. The process used to select feasible rehabilitation alternatives from a set of available alternatives for pavement rehabilitation can range from simple judgment to a decision tree based expert systems. Although using an expert system based on rule reasoning can also acquire proper determining strategies, the process of knowledge retrieval is difficult and the associated rule revising is not convenient.

A new method for determining rehabilitation strategies is presented in the paper based on Case-Based Reasoning (CBR), a unique and useful method used in solving many decision-making problems (Leake, et al., 1997). The technique of CBR is to find solution(s) to a problem from a case base where previous similar problems were successfully solved and their solutions are retained as cases in the case base.

[1] Lin Li, PhD Student, Department of Civil Engineering, Univ. of Arkansas, Fayetteville, AR 72701. PH (479)575-7468; E-mail: lxl025@uark.edu

[2] Kelvin C. P. Wang, Professor, 4190 Bell Engineering, Civil Engineering, Univ. of Arkansas, Fayetteville, AR 72701. PH (479)575-8425; E-mail: kcw@uark.edu

**Rehabilitation Strategies based on Distress Types for Flexible Pavements**

Even though input parameters for pavement rehabilitation can include various functional, structural, and performance information of a pavement, this paper uses distresses as the only parameters to illustrate the application of CBR for selecting rehabilitation strategies for flexible pavements. Surface distress types can be broadly grouped into the following five categories (LTPP, 2003), some of which are included in the first column in Table 1 (Smeaton, et al., 1985).

1) Cracking: including fatigue cracking, block cracking, edge cracking, wheel path longitudinal cracking, non-wheel path longitudinal cracking, reflection cracking at joints, transverse reflection cracking, longitudinal reflection cracking and transverse cracking;
2) Patching and Potholes: including patch, path deterioration and potholes;
3) Surface Deformation: including rutting and shoving;
4) Surface defects: including bleeding, polished aggregate and raveling;
5) Miscellaneous Distresses: including lane-to-shoulder drop-off and water bleeding and pumping.

Strategy options for flexible pavement rehabilitation depend upon local conditions and pavement distress types. An example of rehabilitation strategies and their corresponding selections in a city is shown in Table 1 (Smeaton, et al., 1985). In this example there are a total of 12 strategies for pavement rehabilitation; the maximum number of feasible strategies for given combination of conditions is five, which occurs in two cases; while the minimum number of feasible strategies is two, which occurs in one case. In the table the selection of rehabilitation strategies or feasible alternatives is determined by pavement distress types.

Table1 Feasible alternatives for Pavement Distresses (Based on Smeaton, et al., 1985)

| Distress Presence | COMBINATIONS OF DISTRESS | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cracking Major | N | N | N | Y | Y | Y | Y | N | N | N |
| Rutting > 30% | Y | N | N | N | N | N | N | N | N | N |
| Raveling > 30% | N | Y | N | N | N | N | N | N | N | N |
| Bleeding > 30% | N | N | Y | N | N | N | N | N | N | N |
| Alligator Crack > 30% | N | N | N | N | N | N | Y | N | N | N |
| Edge Crack > 30% | N | N | N | N | N | Y | N | N | N | N |
| Longitudinal Crack >30% | N | N | N | N | Y | N | N | N | N | N |
| Excess Crown Major | N | N | N | N | N | N | N | Y | N | N |
| Alligator Crack Major | N | N | N | N | N | N | N | N | Y | N |
| Rutting Major | N | N | N | N | N | N | N | N | N | Y |
| Feasible Rehabilitation Actions | 3<br>4<br>6<br>11 | 1<br>5<br>7<br>12 | 1<br>8<br>12 | 2<br>5<br>7 | 3<br>4<br>6<br>9<br>10 | 2<br>6<br>9<br>11 | 3<br>6<br>11 | 4<br>10 | 2<br>4<br>6<br>10 | 3<br>4<br>5<br>8<br>10 |

REHABILITATION CODES:

1. 1 in. Overlay
2. 2 in. Overlay
3. 3 in. Overlay
4. Mill 1 in. + Chipseal
5. Recycle 1 in. + 1 in. Overlay
6. Recycle 1 in. + 2 in. Overlay
7. Heater Plane 1 in. + 1 in. Overlay
8. Heater Plane 1 in. + 2 in. Overlay
9. Heater Plane 1 in. + 3 in. Overlay
10. Reconstruct 2 in. AC/4 in.ABC
11. Reconstruct 2 in. AC/6 in. ABC
12. Chipseal

## Case- Based Reasoning (CBR)

CBR is the process of solving new problems based on the solutions of similar past problems. CBR has been formalized for purposes of computer reasoning as a four-step process (Aamodt, et al. 1994):

1) Retrieve: Given a target problem, retrieve cases from memory that is relevant to solving it. A case consists of a problem, its solution, and, typically, annotations about how the solution was derived.
2) Reuse: Map the solution from the previous case to the target problem. This may involve adapting the solution as needed to fit the new situation.
3) Revise: Having mapped the previous solution to the target situation, test the new solution in the real world and, if necessary, revise.
4) Retain: After the solution has been successfully adapted to the target problem, store the resulting experience as a new case in memory.

The processes involved in CBR can be represented by a schematic cycle (Figure 1). A new case is matched with the old cases in the case base, and then one or more similar cases are retrieved from the case base. A solution suggested by the case base is then reused and tested. Unless the retrieved case is a close match with the new case, a new solution will have to be made based on case revisions, which will generate a new case that can be retained in the case base.

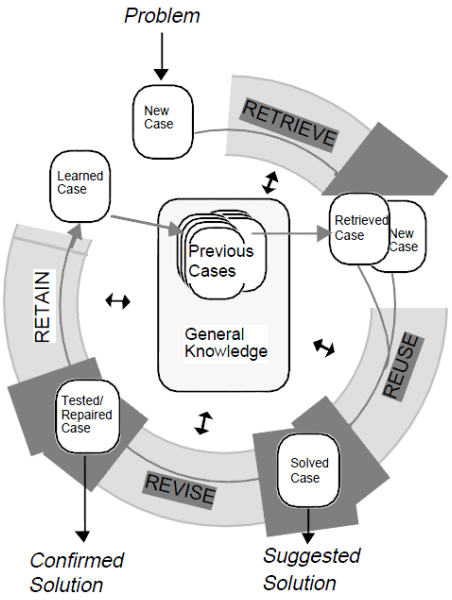



Figure1 the CBR Cycle (Aamodt,et al ,1994)

## The Reasoning Process in the Case Base

Because CBR is a method of analogy reasoning, and utilizes past experience and past cases to solve current problems, the reasoning process with the case base in Figure 2 is established for determining rehabilitation strategies of flexible pavements based on the CBR steps by Stottler (Stottler, 1989) for general decision-making.

1) Case Base. Case base is used to store the past cases. Case storage is an important aspect in designing efficient CBR systems in that, it should reflect the conceptual view of what is represented in the case and take into account

the indices that characterize the case. The case-base should be organized into a manageable structure that supports efficient search and retrieval methods.

2) Case Retrieve. Given a description of a problem, a retrieval algorithm, using the indices or attributes in the memory, should retrieve the most similar cases to the current problem or situation.
3) Case Adaptation. Once a matching case is retrieved, a CBR system should adapt the solution stored in the retrieved case to the needs of the current case. Adaptations look for prominent differences between the retrieved case and the current case and then apply rules that take those differences into account when suggesting a solution.
4) Case Repair. Case repair is referred to restore these cases that have been retrieved from case base according to cases' professional knowledge; parameter features and corresponsive rules.
5) Case study. Case study is used for extending and updating cases in the light of the problem to be solved, and makes sure that CBR system under development provides a long effective and reliable application platform.

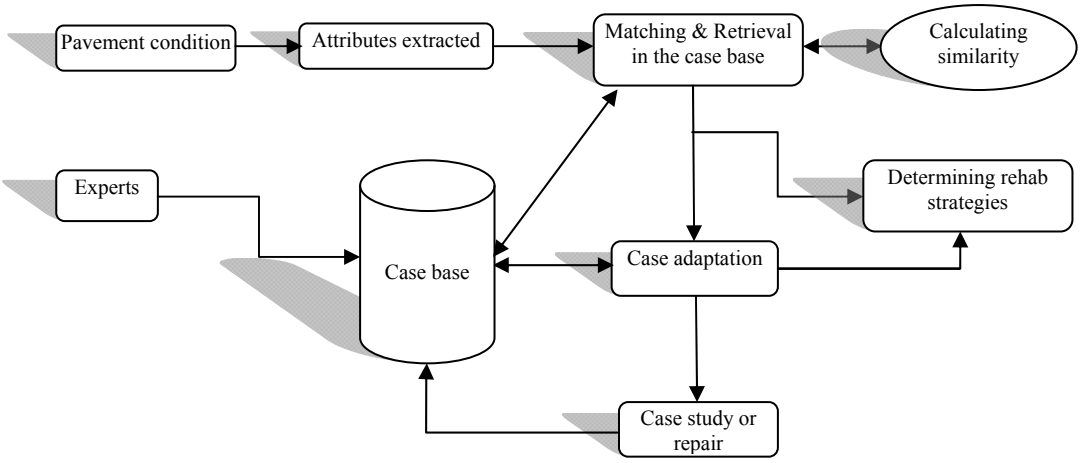



Figure2: the Reasoning Process of the Case Base

**Case Presentation for CBR**

Developing methods for representing cases and problems is required for CBR (Weber et al., 2006). A case representation must be expressive enough for users to accurately describe a problem or case. In addition, CBR systems must reason with cases in a computationally tractable fashion (Plaza, 1995). Cases can be represented in a variety of forms using the full range of artificial intelligence (AI) representational formalisms including frame-based (Plaza, 1995), object-oriented (Bergmann, 2002), semantic nets (Gray, 1984). Frame-based presentation is used here as a type of data structure for describing the cases' attributes which are the basic parameters for retrieving a previous case. A frame, a basic unit of knowledge presentation, is made of many structures called slots, and each slot can be grouped into several sides, and each side has several detailed values. A framework for the general frame-based representation is illustrated in Figure 3; one slot is used to describe the some aspect of the orientated case base; one side is used to describe the some attribute of the corresponding aspect.

In Figure 4, the framework of the case base of the pavement strategy problem contains three slots: the case number C1, the problem description C2 and problem solving methods C3. Slot C2 has 10 sides or distress attributes shown in Figure 4. The inter-linked framework of the cases' slots in the figure constitutes the entire case presentation for determining pavement rehabilitation strategies. In this framed-based presentation, the attributes of each case can be acquired by the problem description, and then the problem solving methods can be selected from the case base by calculating the similarities between the new case and old cases.

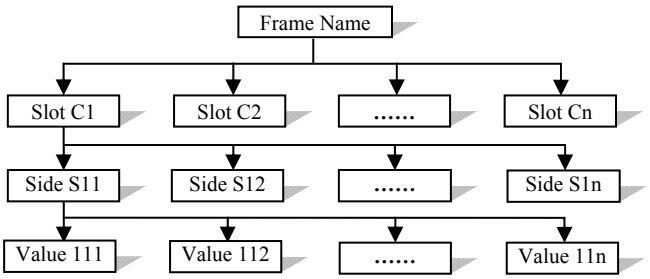



Figure 3 the General Structure of Frame-based Presentation

Case Name < the Case Base > -- Frame Name
Case Slot C1 < Case Number>
Case Slot C2 < Distress Types>- Problem Description
Slot Side S21 < Cracking Major >
Slot Side S22 < Rutting > 30%>
Slot Side S23 < Raveling > 30%>
Slot Side S24 < Bleeding > 30%>
Slot Side S25 < Alligator Crack > 30%>
Slot Side S26 < Edge Crack > 30%>
Slot Side S27 < Longitudinal Crack >30%>
Slot Side S28 < Excess Crown Major >
Slot Side S29 < Alligator Crack Major >
Slot Side S210< Rutting Major >
Case Slot C3 < Determining Strategies>-Description of problem-solving methods

Figure 4 Case Presentation based on Frame-based

Table 2 shows the contents of the case base. The 3rd column is the case attributes. The 4th column contains binary attribute values of cases (Y: the attribute included, and N: not included). J1 to J6 in the table present the pavement rehabilitation strategies for cases T1 to T6 which are the six previous cases in the case base.

**Case Retrieval for the CBR**

Successful application of CBR depends on efficient retrieval algorithms. Case retrieval must be equipped with heuristics that perform partial matches, since in reality there may not be an existing case that exactly matches a new case. Therefore, the methodologies for case retrieval are critical to the use of CBR. There are two well-known methods for case retrieval: nearest-neighbor approach (Lee, 2008), and knowledge guided induction (Michalski, 1978).

Table 2 Case Contents of the Case Base

| | Field | Case attributes | The attribute values of cases | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Case number | Case No | Case Number | T1 | T2 | T3 | T4 | T5 | T6 | |
| Problem description | CRACK_MAJOR | Cracking Major | Y | N | N | N | Y | N | … |
| | RUTTING | Rutting > 30% | N | N | Y | N | N | N | … |
| | RAVEL | Raveling > 30% | N | Y | N | N | N | N | … |
| | BLEEDING | Bleeding > 30% | N | N | Y | N | N | Y | … |
| | ALLI_CRACK | Alligator Crack > 30% | N | N | N | N | N | N | … |
| | EDGE_CRACK | Edge Crack > 30% | N | N | N | N | N | N | … |
| | LONG_CRACK | Longitudinal Crack >30% | Y | N | N | N | N | N | … |
| | CROWN_MAJOR | Excess Crown Major | N | N | N | Y | N | N | … |
| | ALLI_MAJOR | Alligator Crack Major | N | N | N | N | N | N | … |
| | RUT_MAJOR | Rutting Major | N | N | N | N | N | N | … |
| Problem solving | DETER_STRA | Determining Strategies | **J1** | **J2** | **J3** | **J4** | **J5** | **J6** | … |

Nearest neighbor approach is the most widely used technology in CBR since it is provided by the majority of CBR tools (Watson, 1997). First, previous cases in the case base are retrieved for similarity comparison if they have similar case attributes as in column three in Table 2. Then, the case retrieval is based on calculating the similarities of the new case and the old cases. The similarity of the problem case (new case) to a case in the case base is determined according to case attributes. The similarity of all attributes is calculated based on a weighted summary algorithm to provide a measure of the similarity of a pervious case in the case base to the problem case. This is a critical step for retrieving a proper case.

Assume object set of case base $C = \{a_1, a_2, \cdots, a_n\}$, and the attribute set of cases $a_i = \{a_{i1}, a_{i2}, \cdots, a_{ik}\}$, the description of similarity calculation of cases is shown by the equation (1) (Stottler, 1989).

$$Sim(p, a_i) = 1 - Dif(p, a_i) = 1 - \frac{\sum_{j=1}^{k} w_j Dif(P_j, a_{ij})}{\sum_{j=1}^{k} w_j} \quad (1)$$

$P_j$- The $j^{th}$ attribute of the new cases

$a_{ij}$- The $j^{th}$ attribute of the old case $a_i$, data in the 4th column in Table 2

$p$ - The new case

$a_i$ - The old case

$w_j$ - The weight, the importance extent of relevant attributes

$Dif(.)$ - A function or rule used for calculating the difference

This calculation is repeated for every case in the case base to rank cases. Similarities are normalized within a range of zero to one: zero means totally dissimilar and one means an exact match. Based on the values of similarity, the case that has the closest match with the new case may be selected from the case database as the solution.

Table 3 illustrates a calculation example: a new case T is presented with the distress attributes of “cracking major” and “longitudinal crack over 30%”, and three past cases T1, T2, and T3 have the distress attributes of “cracking major” and “longitudinal crack over 30%”, “raveling over 30%”, and “rutting over 30%”. The weight values are given based on engineering experience and the distresses’ impacts on pavement performance. J1, J2 and J3 represent specific strategies as follows:

*J*1– The rehabilitation strategy with Recycle 1 in. + 2 in. Overlay;
*J*2– The rehabilitation strategy with Recycle 1 in + 1 in. Overlay;
*J*3– The rehabilitation strategy with 3 in. Overlay.

Table 3 Attributes and Weights of Old Cases and Status of New Case

| Field | Case attributes | Weights of attributes | Attribute values | | | New case (T) |
|---|---|---|---|---|---|---|
| Case No | Case Number | | T1 | T2 | T3 | |
| CRACK_MAJOR | Cracking Major | 3 | Y | N | N | **Y** |
| RUTTING | Rutting > 30% | 2 | N | N | Y | **N** |
| RAVEL | Raveling > 30% | 1 | N | Y | N | **N** |
| BLEEDING | Bleeding > 30% | 1 | N | N | N | **N** |
| ALLI_CRACK | Alligator Crack > 30% | 2 | N | N | N | **N** |
| EDGE_CRACK | Edge Crack > 30% | 1 | N | N | N | **N** |
| LONG_CRACK | Longitudinal Crack >30% | 1 | Y | N | N | **Y** |
| CROWN_MAJOR | Excess Crown Major | 1 | N | N | N | **N** |
| ALLI_MAJOR | Alligator Crack Major | 2 | N | N | N | **N** |
| RUT_MAJOR | Rutting Major | 3 | N | N | N | **N** |
| DETER_STRA | Determining Strategy | | J1 | J2 | J3 | |

The similarity between the old case T1 and the new case T is calculated with the Equation 1 where subtractions with the same letters being zero, and subtractions with the different letters being one. Equation 2 shows the result of calculating Dif in Equation 1 with corresponding weight values of attributes in Table 3.

$$\text{Dif}(T, T1) = (3 * |Y - Y| + 2 * |N - N| + 1 * |N - N| + 1 * |N - N| + 2 * |N - N| + 1 * |N - N| + 1 * |Y - Y| + 1 * |N - N| + 2 * |N - N| + 3 * |N - N|) / (3 + 2 + 1 + 1 + 2 + 1 + 1 + 1 + 2 + 3) = (2*0 + 3*0 + 1*0 + 1*0 + 2*0 + 1*0 + 1*0 + 1*0 + 2*0 + 3*0) / 17 = 0 \quad (2)$$

Based on Equation 1, the similarity is calculated as Sim (T, T1) = 1 - 0 =1, meaning that the new case is the same as T1. The corresponding rehabilitation strategy is therefore J1: Recycle 1 in. + 2 in. Overlay.

Similarly, the similarity between the old case T2 and the new case T is calculated as Sim (T, T2) = 1 - Dif (T, T2) = 1 - 0.235 = 0.765. Likewise, Sim (T, T3) = 1 - Dif (T, T3) = 1 - 0.353 = 0.647.

$$\text{Dif}(T, T2) = (3 * |N - Y| + 2 * |N - N| + 1 * |Y - N| + 1 * |N - N| + 2 * |N - N| + 1 * |N - N| + 1 * |N - Y| + 1 * |N - N| + 2 * |N - N| + 3 * |N - N|) / (3 + 2 + 1 + 1 + 2 + 1 + 1 + 1 + 2 + 3) = 5 / 17 = 0.294 \quad (3)$$

The similarities between the new case and the three old cases (T1, T2 and T3) are 1, 0.706, and 0.647, respectively, indicating that the case T1 is the most similar to the new case. Therefore, J1 is selected from the case base and recommended as the rehabilitation strategy for the new case T with the distress attributes of "cracking major" and "longitudinal crack over 30%".

## Conclusion

The paper presents a new method to determine rehabilitation strategies with the Case-Based Reasoning (CBR) technique. An example is given based on pavement

distresses and a set of rehabilitation strategies for flexible pavements. The CBR technique is applied to establish the relationship between the attributes of pavement distresses and rehabilitation strategies. A case base is created and used to identify and retrieve the pavement rehabilitation strategies according to the attributes of pavement distresses. In order to execute this new technique efficiently and accurately, the frame-based presentation method and nearest-neighbor approach are used for knowledge or experience presentation and case retrieval. Lastly, an example is given in the paper to illustrate the computation of similarities between a new case and old cases using weight values of cases' attributes. A rehabilitation strategy is recommended in the example based on the values of similarities.

**Reference**

Aamodt, Agnar and Plaza, Enric. (1994). "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches." *Artificial Intelligence Communications. IOS* Press, Vol.7: 1, 39-59.

Bergmann, R. (2002). "Experience Management: Foundations, Development Methodology, and Internet-Based Applications." *Berlin: Springer*.

Gray, P. M. D. (1984). "Logic, Algebra and Databases." *Ellis Horwood Limited*, 200.

Watson, I. (1997). "Applying Case-Based Reasoning: techniques for enterprise systems." *Morgan Kaufmann*, Calif., US.

Leake, David, and Enric Plaza, eds. (1997) "Case-Based Reasoning Research and Development." *Proceedings of the Second International Conference on Case-Based Reasoning*. Berlin: Springer Verlag.

Lee, John. (2008). "A Nearest-Neighbor Approach to the Automatic Analysis of Ancient Greek Morphology."*Proceedings of the 12th Conference on Computational Natural Language Learning*, 127–134.

Long-Term Pavement Performance Program (LTPP). (2003). "Distress Identification Manual for the Long-term pavement performance program." *Publication No*. FHWA-RD-03-031, 1.

Michalski, Ryszard S. (1978). "Pattern Recognition as knowledge-guided computer induction." *Report No.927*, Department of Computer Science, University of Illinois, Urbana, Illinois.

Plaza, E. (1995). "Cases as terms: A feature term approach to the structured representation of cases," *in Veloso, M & Aamodt, A (eds.) Proceedings of the 1st International Conference on Case-Based Reasoning*, Berlin: Springer, 265–276.

Haas, Ralph and Hudson, W. Ronald. (1978). "Pavement Management System." *Hemisphere Publishing Corporation*, 335.

Haas, Ralph, Hudson,W. Ronald and Zaniewski, John.(1994). "Modern Pavement management." *Krieger Publishing Company*. 207-209.

Stottler, Richard H. (1989). "Rapid Retrieval Algorithms for Cased-Based Reasoning." *Proceedings of the Eleventh I JCAI*-89, vol 1:233-237.

Smeaton, W. K., Karan, M. A. and Haas, R. (1985). "Determining the Most Cost-Effective Combination of Pavement Maintenance and Rehabilitation for Road and Street Networks," *Proc., First North American Pavement Management Conf.,* Toronto.

Weber, Rosina O., Ashley, Kevin D. and Brunighaus, Stefanie. (2006)."Textural case-based reasoning." *The Knowledge Engineering Review*, Vol. 20:3, 255–260.

# Gateway Traveler Information System: Regional Traveler Information Services

Charles Sikaras, ITS Program Specialist
Illinois Department of Transportation (IDOT)
201 West Center Court, Schaumburg, IL 60196
(847) 705-4800, Charles.Sikaras@illinois.gov

John F. Dillenburg, Ph.D., Gateway Project Manager and Constance A. Kelly, Ph.D.,
Gateway Project Coordinator -University of Illinois at Chicago
Jerry Hron, Gateway Project Manager - IDOT
Scott Lee, P.E., Systems Engineer – Delcan Corporation

## Abstract

The Gateway Traveler Information System (Gateway) and www.travelmidwest.com web site (formerly www.gcmtravel.com) provide timely, accurate, and useful traffic information to a variety of public and private constituents. Recurring and non-recurring congestion directly impacts the economy, the environment, and the safety of the transportation network. The services and data provided by the Gateway have become a mission-critical service of the Illinois Department of Transportation (IDOT) and the geographically-connected roadway operators that are members of the Lake Michigan Interstate Gateway Alliance (LMIGA).

This paper details the background and history of LMIGA and the Gateway, specifically www.travelmidwest.com. It highlights the goals of the Gateway and identifies four challenges and the associated resolutions that IDOT and the LMIGA partners use to ensure the Gateway is able to maximize the dissemination of quality traveler information to the staffs of operating agencies, the traveling public, and third party data subscribers. This paper also describes the elements that will impact both the system and the regional partnership as technologies advance, traveler information demands increase, and policies evolve.

## Traveler Information Benefits and Congestion Costs

### *Benefits of Traveler Information*

Traveler information studies from across the country are available via the U.S. Department of Transportation Research and Innovative Technology Administration (RITA), through the Intelligent Transportation System (ITS) Benefits Database. The Federal Highway Administration (FHWA) study "Managing Demand Through Traveler Information Services" revealed that the state of Washington experienced benefits including reduced trip times, more predictable travel, less stressful conditions, and safer travel conditions as a result of their traveler information efforts.

### *Cost of Congestion*

The Texas Transportation Institute's (TTI) "Urban Mobility Report" (Shrank and Lomax 2009) placed the Chicago region second in rush hour travel among major metropolitan regions and third worst in the nation for wasted time, fuel, and total congestion costs ($4.2 billion annually).

The Metropolitan Planning Council's Report "Moving at the Speed of Congestion" (MPC 2009) painted an even more severe picture of regional congestion.

- The $7.3 billion total regional cost includes the cost of lost time ($6.98 billion), fuel ($354 million), and environmental damages ($33 million).
- Gridlock also increases labor costs, impeding the creation of 87,000 jobs.
- Lost time costs the Chicago-area economy and its drivers nearly 20 times more than the cost of wasted fuel.

Regionally, congestion adds 22 percent to peak period travel times. Within Chicago itself, congestion increases peak period travel times by about 40 percent.

The region has neither the space nor the financial resources required to build its way out of congestion. Traffic operations and traveler information play a vital role in helping support roadway users in the region by providing the best information possible to support time, route, and mode of travel choices.

### Regional Background

The Gateway system has a long history of serving as a central traveler information repository for the three-state Gary-Chicago-Milwaukee (GCM) ITS Priority Corridor established in 1993. Since 1998, the Gateway has provided traveler information to a growing geography. The mission originally centered on the Chicago area expressway system with www.travelinfo.org, which disseminated travel time and congestion information collected from the IDOT Traffic Systems Center (TSC). In 2001, as additional regional traffic data collection systems came on-line in Wisconsin and Indiana, and at the Illinois Tollway, www.gcmtravel.com became the new web site.

(Billerbeck 2009)The LMIGA is a multi-state, multi-disciplinary organization that includes the following members: Illinois Department of Transportation (IDOT), Illinois State Toll Highway Authority (ISTHA), Wisconsin Department of Transportation (WisDOT), Indiana Department of Transportation (InDOT), Michigan Department of Transportation (MDOT), Chicago Skyway, and Indiana Toll Road. Recently, LMIGA supported a web site rebranding effort which includes an updated look and feel and an opportunity to begin initiating new services.

The Gateway and www.travelmidwest.com web site provide tools that promote the value of ITS technologies and services, and represent the face of traveler information in the region. Figure 1 depicts the history of the Gateway and web site interface.

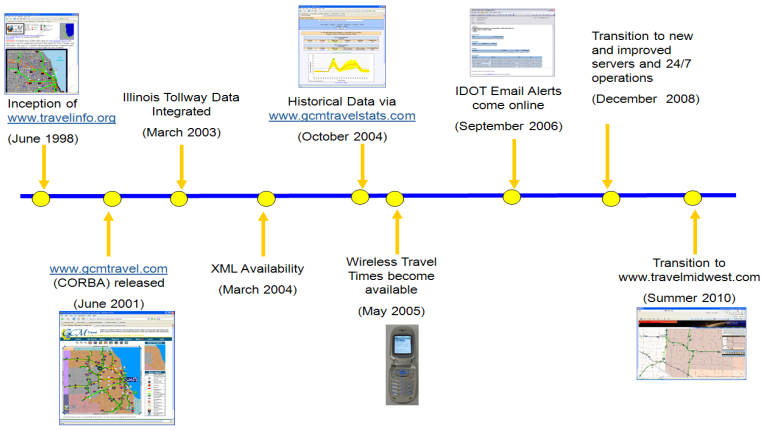



**Figure 1:** Gateway Development Timeline and Milestones

## Summary of Services

The Gateway provides an integrated electronic clearinghouse of regional real-time traffic data for consumption by public and private users. Information on a variety of roadway classes from interstate highways to locally-maintained roads in the LMIGA region is collected, concatenated, and then provided to the user community through a broad collection of reports, web pages, and standardized data feeds. The raw data from roadside detection equipment is used to compute travel times and congestion information, and provide visual interpretations via map displays. Additional information obtained from roadway signage and video cameras provides context for absolute travel times and computed congestion.

Numerous reports provide detailed information in tabular form for construction, incidents, congestion, travel times, special events, dynamic message signs, and vehicle detectors. While much data is automated; incident, special event, and construction data is manually entered from emails, faxes, radio frequency monitoring, and telephone notifications.

The Gateway provides data to approximately eighty external entities including research institutions, private sector traveler information service providers, and the media. The monthly web site demand includes 375,000 site visits and over 4,500,000 total page views.

The Gateway was selected as the Intelligent Transportation Society of the Midwest 2009 Project of the Year. It also garnered two national projects of the year and traveler information web site awards from FHWA and ITS America in the mid-2000s.

## Mission

The LMIGA mission is to ensure that traffic moves safely and efficiently. This is realized through interagency communication and coordination, improvement projects, training efforts, and region-wide planning.

IDOT and the LMIGA partners are committed to providing timely, useful, and accurate traffic information from a variety of agency sources via the latest technologies. This mission is accomplished with a focus on reliability and usability for the end users.

## Challenges

Four challenging areas influence operations and policy/investment decisions for IDOT and the LMIGA membership.

1) The Gateway's long-running use of standards in message and data formatting has had a national influence on Center-to-Center (C2C) systems. While the Gateway architecture and interfaces have remained stable for nearly ten years, there are both programmatic and human factors that remain a critical part of the Gateway's success.
2) The Gateway collects data from six unique systems and an array of technologies. Conveying a consistent definition of timely, accurate, and useful information requires commitment and routine evaluation of data within the Gateway and the source system processing.
3) The Gateway needs to maximize the utility of the system for everyday commuters, discretionary travelers, the trucking community and other commercial users, third party data subscribers, and regional transportation management system operators.
4) As situations change, the Gateway must balance being at the forefront of technology with the reality that not all contributing systems will necessarily be on the same timeline.

### *Challenge #1: Technology Standards*

The Gateway has long been a leader in the formation and application of national Intelligent Transportation System (ITS) standards programs which promote commonality of traffic data exchanges and provide all stakeholders ease and efficiency in utilizing the data. IDOT and the Gateway staff served on FHWA and American Association of State Highway and Transportation Officials (AASHTO) standards committees as early as the mid-1990s and played a major role as an early implementer of C2C standards using Common Object Request Broker Architecture (CORBA) and in the migration to Extensible Markup Language (XML) based data exchange.

The Gateway is designed to be the central hub to facilitate corridor-wide system integration and inter-operation of various ITS systems. Additionally, the Gateway is also the primary information dissemination channel to the general public and to private sector traveler Information Service Providers (ISP).

When the Gateway system architecture was defined in the late 1990s, two C2C communications approaches were being addressed by the National Transportation Communications for ITS Protocol (NTCIP) C2C Work Group: DATEX (Data Exchange) and CORBA. These standards attempted to harmonize use of the Traffic Management Data Dictionary (TMDD) and Message Sets for External Traffic Management Center Communication (MS/ETMC2). It is noteworthy that the bulk of the original Gateway development occurred while these standards were under development.

(Lee 2005) In the early 1990's, the region chose to adopt CORBA as the primary

system interface platform for the Gateway primarily because it is an open and object oriented system integration platform that is independent of hardware platform, operating system (OS), and programming language. The adoption of CORBA provided a reliable and open platform for corridor-wide system integration and inter-operation.

Even though XML was not on the radar screen of the NTCIP C2C work group at that time, the great potential of XML technology and planning blossomed to support a transition to XML-based system interfaces and data services.

XML was originally used in the Gateway to manage internal system configuration data. Later, a mechanism was developed to map the Gateway's CORBA data structures to XML. Data logged in XML formats allowed system developers to efficiently perform various system debugging and monitoring activities.

Though the Gateway's CORBA-based interface was successful, it became clear that the cost and effort associated with developing and maintaining a CORBA application presented an obstacle to a number of existing and potential Gateway user systems, from both a private sector ISP perspective, and the perspective of existing and future agencies wishing to exchange data. Some of the ISP's elected to 'scrape' the web pages for data, employing a user program to read and extract information from html-encoded data. One of the major issues inherent in web scraping is that the web scraping process will fail if changes, such as a table layout change, are made to the web pages.

To encourage broader usage of information provided by the Gateway, XML was added as a distribution method. XML drastically reduced the overhead of providing data to ISPs on several levels. The XML subscribers are presently comprised of a variety of ISPs. No sensitive information is part of the data sets making the internet a suitable method of transport to the ISPs, so no additional routers nor networking equipment are required. Similarly, the use of XML meant that third party subscribers had easier access to interface expertise. Thus, IDOT no longer had to provide additional software support.

Notably, the Gateway has also seen a migration of all C2C exchanges to XML. For example, the Illinois Tollway TIMS has migrated to XML exchanges as they have made internal system upgrades and found the migration is more cost effective to implement due to licensing arrangements for CORBA. Additionally, the XML interface is easier to maintain. Simply stated, as a production system serving real-time users, LMIGA agencies, and researchers alike, the technologies used by the Gateway, specifically XML, have been selected to support and attract traffic management and traveler information data sharing and maximize the Gateway's overall utility. The Gateway fuses the data from individual agencies and their collection technologies (inductive loops, radar, transponders, Bluetooth) into a best-fit common view of traffic data for the entire region.

***Challenge #2: Consistency in Quality and Timeliness***

The Gateway uses standard interfaces and a few custom interfaces, where needed, to bring the region's data to a single access point and a common format that is

appropriate for all users. Since going to a 24/7 operation, IDOT has been able to support a more aggressive approach to evaluate data accuracy and consistency. A few specific examples stand out.

*Dynamic Message Sign (DMS) Performance Monitoring:* The Gateway staff monitoring of DMS reports has resulted in the following changes:

- The Gateway has updated its approaches to status monitoring, time stamping of source information, and handling of multi-phase messaging.
- The source systems have been able to use the Gateway review to support identification of operational issues and DMS maintenance activities.
- Regional message consistency has gained a renewed focus through this effort and the efforts of the LMIGA's Traffic Center Communications Work Group.
- DMS sign information is validated against congestion data and travel time information to ensure that all contributing systems are in synch.

*Travel Time and Congestion Monitoring:* Travel times and congestion information are served to the public and private data consumers through a variety of reporting mechanisms including several technology modes (web map, html report varieties, Really Simple Syndication (RSS), and feeds).

The partner agencies deploy multiple technology solutions to measure congestion and travel times. Travel times distributed through any Advanced Traveler Information System (ATIS) provide either a series of estimates from a wide variety of technologies with specific behaviors or direct collection of travel times that immediately are past tense once collected. As the Gateway has monitored the data and feedback from the public, the partner agencies have made additional commitments to ensure accuracy.

*Incident Data Collection:* Traffic incidents and their severity have at least as much impact on the transportation system users as recurring congestion. A combination of distribution mechanisms, including emails and forwarded electronics reports, monitoring of responder radios via scanner, monitoring of broadcast radio, and electronic interfaces (such as Lake County and the Illinois Tollway connections to dispatching), support Gateway incident data collection. The Gateway operations staff monitors congestion patterns to identify potential incidents before they become major traffic events. This also ensures the location and impact of incidents are accurately conveyed to the public.

***Challenge #3: Maximizing Data Utility and Usability***

The Gateway has had a long history of providing the following datasets to both web site users and data subscribers:

- Incidents
- Construction
- Special Events
- Vehicle Detectors
- Congestion Links
- Travel Times
- Dynamic Message Signs

- Closed Circuit Television (CCTV) Images
- Links to other traveler information services

The web site data is available through RSS feeds, XML data subscriptions, mobile device-friendly web pages, customized travel times, and standard data reports. Additional links are also available throughout the web site to more generic DOT pages, major construction project links, and transportation-related services. The current www.travelmidweststats.com web site provides historical travel times and is directly integrated with the Gateway. This information serves as a tool to support trip planning and to compare specific current travel time conditions with historical data. The data provides a relative comparison of congestion - Is this route experiencing worse than normal congestion levels and, if so, how much worse?

The effort to establish www.travelmidwest.com was undertaken with an understanding that www.gcmtravel.com was already a very successful endeavor. The GCM acronym was consistent with the GCM Corridor through 2001. The LMIGA partnership reflects the geographic addition of Michigan along with expansion of Illinois coverage to the Quad Cities (Iowa-Illinois) area. The www.travelmidwest.com web site provides coverage for the entire LMIGA region. The web site contains an easier to identify brand with a more recognizable URL, improved presentation methods, usability options, web site navigation, expanded graphics use, and a new Truckers Report.

The Truckers Report has been introduced in order to better meet the needs of the freight community. IDOT participated in the Midwest Truck Show in both 2009 and 2010 with an explicit goal of generating more awareness for the web site and soliciting feedback to improve its utility to freight users, a major stakeholder and consumer of the data. The rebranding effort uncovered the fact that the commercial vehicle operators were underserved by the available traffic information providers even though they were a significant source of traffic. The trucking community expressed an interest in streamlining the available information to filter the data down to reflect the most critical conditions in the region – highest travel times, highest impact construction, and highest impact incidents.

While the Gateway itself is a production system serving transportation system users (the public) and transportation system operators as the primary audiences, the availability of a wide variety of data has supported numerous and varied transportation research efforts and ITS standards development since its inception in 1998. Both CORBA and XML feeds have been provided to numerous public research institutions and private researchers since 2001.

Third parties interested in becoming a data subscriber must submit an application to IDOT and LMIGA partners defining their proposed use of the data and desired data elements. Once approved, the users are assigned permissions and are provided the Gateway External Interface User Guide (Parsons 2005) that documents detailed available traffic data elements and the fundamentals of how the users should configure their data subscription.

Feedback from web site users, third party data subscribers, and LMIGA traffic operation center officials is encouraged through webmaster@travelmidwest.com, and

drives the evaluation and deployment of system changes. User feedback was the driving force in the rebranding effort. Realistically, IDOT and the LMIGA partnership recognize that feedback is more likely when users perceive problems in presentation or quality of information provided. Fortunately, over the eight year history of www.gcmtravel.com, the majority of feedback has centered on requests for more services or enhanced presentation of traffic information.

***Challenge #4: Keeping the Gateway System at the Leading Edge***

IDOT and the LMIGA partners have monitored regional changes in operations and national/international technology developments as they apply to traveler information. This includes monitoring the requirements contained in the FHWA proposed rule to establish a Real-Time System Management Information Program (FHWA 2009).

The rule being considered recommends the following traveler information service requirements:

| | Timeliness for Delivery | | | |
|---|---|---|---|---|
| Category of Information | Metropolitan Areas (min) | Non-metro Areas (min) | Availability (%) | Accuracy (%) |
| Construction Activiites: Implementing or Removing Lane Closures | 10 | 20 | 90 | 85 |
| Roadway or Lane Blocking Traffic Incident Information | 10 | 20 | 90 | 85 |
| Roadway Weather Observation Udpates | 20 | 20 | 90 | 85 |
| Travel Times Along Highway Segments | 10 | NA | 90 | 85 |

**Figure 3:** FHWA's proposed rule for Real Time System Management Information Program Minimum Requirements

Fortunately for the LMIGA region, the Gateway and the supporting systems already in place provide a significant head start in fulfilling this rule, particularly in the metropolitan areas connected through the Gateway. This rulemaking will impact the design and implementation of all future roadway improvement projects. IDOT has already implemented changes in construction projects on the Dan Ryan and Eisenhower Expressways where temporary detection technologies replaced out-of-service permanent sensors to measure speed and congestion.

Another major impact of the rule is the focus on non-metropolitan areas and arterials across the region. Presently, there is limited availability of mature systems to provide some of the proposed data.

**Arterial Data**

The Gateway has expanded its commitment to providing users construction data for arterials. Collection of real-time congestion data on the arterial network has been historically difficult.

The Lake County, Illinois Passage System is providing arterial congestion data to the Gateway as of this paper's conference date. Overlaying many data types and sources for expanded services on arterial streets provides a very complete picture of the surface transportation network. Keeping the site usable, while accommodating a growing geographic area and inventory of data types, will be an on-going challenge.

The aggregation of transit data from sources such as the Regional Transportation Authority (RTA) and its service boards (Pace, Metra, and the Chicago Transit Authority) will also add to the challenge.

Expanding the collection of arterial data on the Gateway will follow the development timelines of arterial systems. Traffic condition data will be collected in a manner similar to that used by Lake County and the City of Chicago Midway Airport Advanced Traffic Management System (ATMS).

**Trending Technologies**

While IDOT and substantial portions of the LMIGA region are relying on the traditional ITS technology choices to collect traffic data, IDOT has remained an engaged observer of new technologies such as probe data projects and Vehicle Infrastructure Initiatives (VII). IDOT expects to implement economically feasible technologies that have matured from the "bleeding edge" to the "leading edge". During the 2010 highway construction season, IDOT deployed Bluetooth technologies to support real-time travel time collection during resurfacing of Interstate 290, where pavement milling eliminated the existing detection infrastructure.

It can be anticipated that probe data and some of the national initiatives being considered to collect a national data set of speed and congestion conditions will help fill the data gaps in arterial and non-metropolitan regions.

The age of social media, such as Twitter, provides one example of change in accepted end-user technology that impacts the expectations for systems such as the Gateway. Fundamentally, the Gateway is committed to ensuring that the users who need the information have the best possible access. Providing information in formats that the users find most comfortable is as important as providing correct and timely information.

The impact of incidents on travel times and congestion is recognized by traffic operations staff and motorists. The Gateway and its partners are committed to promoting electronic integration among dispatch services, traffic management centers, and traveler information hubs. The safety of responders and motorists approaching an incident scene is not limited to the physical location of vehicles, debris, and responders associated with the incident. Informed motorists can choose alternative routes or modes of travel and help reduce congestion near an incident scene. Just as important, those motorists remaining on their selected route will arrive at an incident scene aware and prepared to encounter congestion and response teams, thus resulting in a safer environment.

Similarly, the standards-based implementation for basic traffic data and traveler information can be expected to undergo a detailed review by IDOT and LMIGA partners as data sources mature and evolve over the coming three to five years.

**Summary**

The Gateway compiles and provides a wealth of real-time traffic data and traveler information to the regional traffic management systems, planners, researchers, third party information service providers, and the public. The system's success has hinged

on its ability to provide timely, accurate, and useful information to all consumers. The Gateway agency partners, through LMIGA and the Gateway's development and enhancement team managed by IDOT, are committed to the use and maintenance of advanced technologies for data processing and distribution. These technologies are implemented to meet the needs of the data consumers, encourage system feedback, and focus a keen eye on the future of traffic information collection and dissemination.

**References**

Research and Innovative Technology Administration (RITA), U.S. Department of Transportation (US DOT), "ITS Benefits Database", http://www.itsbenefits.its.dot.gov/its/benecost.nsf/SingleTax?OpenForm&Query=Traveler+Information

Federal Highway Administration (FHWA), "Managing Demand Through Travel Information Services", EDL Document No.: 14072 Publication No.: FHWA-HOP-05-005, http://ops.fhwa.dot.gov/publications/manag_demand_tis/travelinfo.htm

Schrank, David and Lomax, Tim, Texas Transportation Institute, "2009 Urban Mobility Report", highlights via http://mobility.tamu.edu/ums/

Metropolitan Planning Council, 2009, "Moving at the Speed of Congestion", http://www.metroplanning.org/uploads/cms/documents/mpcreport_movingatthespeedofscongestion.pdf

Billerbeck, Lisa, Transmart Technologies, "2009 Lake Michigan Interstate Gateway Alliance Annual Report", http://www.gcmcommunicator.com/public-library/pic-reports/pic-corridor-program-plan/09_LMIGA_AnnRept_FINAL.pdf

Lee, Scott and Zavattero, David and Wu, Wei, "Application of XML by the GCM Corridor Facilitating Traveler Information and Integration", Illinois Department of Transportation - ITS Program Office, U.S.A., 120 W. Center Ct., Schaumburg, IL 60195, U.S.A., ITS World Congress 2005

Parsons Corporation, "Gateway External Interface User Guide", May 25th, 2005 http://www.gcmcommunicator.com/public-library/pic-technical-docs/pic-trav-info-system/Gateway%20External%20Interface%20User%20Guide.doc/view

Federal Highway Administration, 2009, "Real-Time System Management Information Program: Proposed Federal Rule for Real Time Traveler Information Program", January 14th, 2009 http://edocket.access.gpo.gov/2009/E9-392.htm

# HIGH SPEED RAIL DESIGN THROUGH A URBAN/ SUBURBAN CORE

Genaro Mejia, PE
Arup
12777 Jefferson Blvd, Los Angeles, CA 90066;
Tel: (310) 578-4462; e-mail: genaro.mejia@arup.com

## ABSTRACT

The California High Speed Rail project is an 800 mile long high speed rail network that will ultimately connect the cities of Los Angeles, San Francisco, Sacramento and San Diego. Phase I of the project will connect Los Angles to San Francisco through the Central Valley; Phase II will add connections to Sacramento and San Diego completing the overall system. Phase I is divided into eight sections including Palmdale to Los Angeles, Arup is a part of a joint venture team with Hatch Mott Macdonald (HMM) and URS to develop 30% design drawings to certify the Environment Impact Report/ Environmental Impact Statement (EIR/EIS) and to provide a basis for contractors to submit bids for construction. Palmdale to Los Angeles has two distinct sections including Palmdale to Sylmar which passes through rural and suburban cities with terrain ranging from flat lands to mountainous and Sylmar to Los Angeles Union Station (LAUS) which passes through a densely populated mixture of urban and suburban cities in the San Fernando Valley adjacent to the Los Angeles River. Particular challenges are faced when designing a high speed rail corridor with design speeds reaching 220 miles per hour (mph) through a dense urban suburban core such as in the Sylmar to LAUS section of the project. The paper will discuss the criteria used to design the rail alignment, the challenges and constraints encountered along the route and the construction methods employed to mitigate the constraints.

## INTRODUCTION

In April 2009, the Federal Railroad Administration (FRA) released a long term plan for high speed rail in the United States [2]. The plan identified high speed rail corridors throughout the United States that would be a part of a national high speed rail network. To help make this a reality the US Department of Transportation (USDOT) distributed $8 billion through the American Reinvestment and Recovery Act (ARRA) in January 2010, with an additional distribution of $2.5 billion in October 2010 to help jump start various high speed rail programs throughout the country. The total value of applications submitted was $55 billion in January and $8 billion in October. The ration for the value of submitted applications to funds available was six to one, showing a high demand for a new form of transportation. In California, the California High Speed Rail Authority (CHSRA) was able to pass Proposition 1A, a high speed rail bond that allocated $9.95 billion for high speed rail development [1]. Furthermore, California capitalized on its advanced development of the California

High Speed Train (CHST) project and received $2.25 billion in federal ARRA grants in January 2010 and $750 million in October 2010 from the $10.5 billion available. High Speed rail already exists around the world, with developed systems in France, Spain, Japan and China to name a few, for the United States this is a new form of transportation with California being the most developed program in the country.

The CHST project is planned as a fully electrified system that will provide an alternative form of transportation within the state. The corridor will be fully grade separated and will be a steel wheel on steel rail system [4]. Trains will be powered by an overhead cantenary system (OCS) and will have the ability to travel to speeds up to 220 mph. As established by Proposition 1A the HST will have a travel time between Los Angeles and San Francisco of 2 hours 40 minutes. Currently the CHSRA has commissioned the completion of segment specific Environmental Impact Statements (EIS) and Environmental Impact Reports (EIR) to satisfy the federal and state environmental laws NEPA and CEQA [4]. As a part of this process the CHSRA has developed alignment design criteria to establish rail corridors for environmental evaluation. For the Palmdale to Los Angeles (PL) segment the aim is to design the corridor that will attract the most ridership, with the fast travel times, with the least amount of impacts on the environment and the communities, while minimizing costs and maximizing benefit. Within this segment stations will be provided at the north end in Palmdale, the south end in Los Angeles, and a third station in the center of the segment. Two main sections within the PL segment can be characterized, Palmdale to Sylmar is comprised of rural and suburban cities with mountainous terrain. Sylmar to Los Angeles, is generally a suburban to urban area running through the heart of the San Fernando Valley. It is the goal of this paper to discuss the alignment design through this dense urban/ suburban core and how the physical and environmental constraints are balanced with the need to satisfy minimum design criteria by implementing the appropriate construction method for the particular constraint.

## SYLMAR TO LOS ANGELES SECTION DESCRIPTION

The Sylmar to Los Angeles portion of the PL section is approximately 21.5 miles long and passes through the cities of Los Angeles, San Fernando, Burbank and Glendale. The high speed rail alignment runs along an existing Los Angeles Metropolitan Transportation Authority (LAMTA) right of way parallel to San Fernando Road. The existing corridor is owned by LAMTA and is utilized by a commuter rail service Metrolink and freight service Union Pacific Railroad (UPRR), the existing rail corridor is currently designed for top speeds of 80 mph [4]. The right of way is adjacent to a regional airport and an international airport, in addition the corridor is crossed by freeway structures and waterways. Towards the southern end of the segment the alignment passes through state and regional parks as well as

primary and secondary schools. The terminus of the segments is at the existing Los Angeles Union Station.

## ALIGNMENT DESIGN CRITERIA

For the track alignment design the California High Speed Train project Technical Memoranda (TM) has been used [3]. The criteria was developed from many sources including Asian and European standards and adapting them to AREMA standards. It is noted that the criteria as discussed below is not finalized and still a work in progress.

The criteria for HST is different from conventional rail as design speeds are much higher, ranging from 35 mph to 220 mph. As a part of developing an alignment for the EIR/ EIS process it is important to be as conservative as possible when applying the criteria as the design can be more refined at later stages of design. Generally criteria was set for desirable, minimum and exceptional conditions, where using exceptional criteria would require specific approval on a case by case basis from the program management team (PMT). The driving goal behind the criteria is passenger comfort, since trains will be travelling at high speeds it is imperative that passengers feel comfortable and safe while riding the train.

### Horizontal Alignment

For specific horizontal alignment criteria see TM 2.1.2 from the CHST Technical Memoranda, below is a summary of key points for the horizontal alignment design [3].

- Minimize use of reverse curves
- Overlapping horizontal and vertical curves not desirable
- Vertical curve within spiral curve is an exceptional condition
- Tangent lengths need to meet minimum element length criteria
- Applied super elevation is equal to approximately two-thirds of balanced super elevation
- Clothoid Spirals defined for speeds up to 80 mph
- Half-sine Sprials defined for speeds greater than 80 mph

The table below is a summary using the required minimum radius, minimum element length, super-elevation, and minimum spiral length for a given design speed. Note that desirable values are higher, while exceptional values are lower. For preliminary design purpose the intent is to satisfy the minimum design criteria wherever possible.

**Table 1: Design speed and minimum horizontal criteria**

| Speed (mph) | Radius (ft) | Min element length (ft) | Balanced Super elevation (in) | Applied Super elevation (in) | Unbalance (in) | Min spiral length (ft) |
|---|---|---|---|---|---|---|
| **220** | 22000 | 775 | 8.8 | 6 | 2.8 | 1720 |
| **180** | 15000 | 475 | 8.6 | 6 | 2.6 | 1410 |
| **140** | 9000 | 370 | 8.7 | 6 | 2.7 | 1100 |
| **100** | 4500 | 264 | 8.9 | 6 | 2.9 | 780 |
| **65** | 1900 | 172 | 8.9 | 6 | 2.9 | 710 |
| **35** | 1000 | 92 | 5 | 3 | 1.9 | 360 |

**Vertical Alignment**

For specific vertical alignment criteria see TM 2.1.2 from the CHST Technical Memorandum, below is a summary of key points for the vertical alignment design [3].

- Desirable vertical grades at less than 1.5% for HST
- Minimum vertical grades at less than 2.5% for HST
- Exceptional vertical grades shall not exceed 3.5% for HST
- Maximum vertical grades at less than 2.0% for passenger/ freight rail
- Minimum vertical grades in trench or tunnel shall not be less than 0.25%
- Tangents between vertical curves to meet minimum tangent length criteria
- Vertical curves within spiral curves are an exceptional condition

The table below is a summary of minimum vertical curve lengths grouped by grade difference. As speed and grade difference increase length of curve increases. Note that desirable values are higher, while exceptional values are lower. For preliminary design purpose the intent is to satisfy the minimum design criteria wherever possible.

**Table 2: Design Speed and minimum vertical criteria**

| | Grade Difference (A) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 5 | 4.5 | 4 | 3.5 | 3 | 2.5 | 2 | 1.5 | 1 | 0.5 |
| **Speed (mph)** | **Min. Curve Length (ft)** | | | | | | | | | |
| **220** | 6510 | 5860 | 5210 | 4560 | 3910 | 3260 | 2610 | 1960 | 1310 | 780 |
| **180** | 4360 | 3920 | 3490 | 3050 | 2620 | 2180 | 1750 | 1310 | 880 | 480 |
| **140** | 2640 | 2380 | 2110 | 1850 | **1590** | 1320 | 1060 | 800 | 530 | 370 |
| **100** | 1350 | 1210 | 1080 | 950 | 810 | 680 | 540 | 410 | 270 | 270 |
| **65** | 1000 | 900 | 800 | 700 | 600 | 500 | 400 | 300 | 200 | 180 |
| **35** | 1000 | 900 | 800 | 700 | 600 | 500 | 400 | 300 | 200 | 100 |

**CONSTRUCTION METHOD**

At-grade – HST runs at grade along the existing ground surface. Where grade separations exist with a rail over road condition the existing rail bridges will most likely need to be replaced. If the grade separation is above grade, adequate vertical clearance needs to be provided to protect the overcrossing. In cases where grade separations do not exist, a road over or road under grade separation may be considered. At grade HST is generally the least expensive of all construction methods.

Viaduct – HST runs above grade on a pre-cast box structure approximately 30 feet from the existing ground surface to top of rail. This provides for 16.5 feet vertical clearance existing ground to bottom of structure. A viaduct structure may be considered when crossing waterways or crossing consecutive at grade streets. In cases where multiple at grade streets are crossed it may be more cost effective than constructing road grade separations over an at grade HST. Viaduct is more expensive than at-grade but less expensive than trench or tunnel.

Trench – HST runs below ground, generally 35 feet below with retaining structures anchored into existing ground. If trenches are greater than 35 feet below ground horizontal struts may be required. Trenches require 27 feet vertical clearance from top of rail to bottom of any road structure crossing the trench. Trench is recommended in situations where the HST may conflict with airport clearance envelope, where HST may conflict with existing parks or on approaches to cut and cover or bored tunnels. Trench is more expensive than viaduct or at-grade but less expensive than tunnel.

Tunnel – HST runs underground either in a cut and cover tunnel 35-45 feet below ground or a bored tunnel at least 60 feet below ground. Bored tunnels are generally 30 feet in diameter with 68.5 feet spacing between centers. Tunnels should only be used in extreme cases where constraints override all other construction methods. Tunnels are the most expensive construction method.

Road Grade Separation – Road grade separations can be used when HST runs at grade and a major street needs to cross the corridor. Since HST requires that it runs fully grade separated there are many grade separations that need to be built within this corridor. Grade separations are disruptive to local communities as they require long approaches with grades not to exceed 8%. For road over HST grade separations 27 feet vertical clearance is required from top of rail to bottom of road structure. For road under HST grade separations 16.5 feet vertical clearance is required from top of road to bottom of rail bridge. Generally, road over HST grade separations are easier to construct but are visually obtrusive and require longer approaches, road under HST grade separations are more difficult to construct due to the intensive utility relocations and pump design for drainage.

## PHYSICAL CONSTRAINTS AND CHALLENGES

### Freeway Structures

Challenge – The HST alignment will cross major freeways including I-5, SR-110, SR-134 and SR-118. The goal is to cross the freeways at grade through the existing bridge spans so as to not modify the existing structures. Adequate horizontal clearance is necessary from centerline of rail to bridge pier or abutment to provide for crash barriers and vehicle travel envelope.

Solution – To avoid reconstruction of freeway structures the HST should pass through them at grade, viaduct structures over freeway structures should be avoided as they will be high structures measuring 60-80 feet in height. Trench structures should also be avoided as they may undermine freeway structure footings.

### Airports

Challenge – The HST alignment will be passing by two airports, Whiteman Airport a regional airport located in City of San Fernando, and Burbank Airport located in the City of Burbank. Airport clearance envelopes were obtained from both airport authorities adhering to federal regulations to determine the areas where the HST could not pass through without affecting the clearance.

Solution – In order to provide adequate clearance for any OCS systems vertical clearance measuring 27 feet from top of rail to flight clearance envelope is required. To provide adequate clearance a trench will be needed for the HST to pass through and avoid the airport clearance envelope.

### Existing Tracks

Challenge – The HST corridor is planned to be placed in an existing right of way measuring 105 feet in width, the right of way is owned by the Los Angeles Metropolitan Transportation Authority (LAMTA). Currently the tracks are operated by Metrolink, Amtrak and a freight provider Union Pacific Railroad (UPRR). The existing tracks were designed for speeds of 80 mph, while proposed speeds for HST will be 220 mph.

Solution – In order for new HST tracks and to be placed within the same corridor, the existing tracks will need to be relocated with an additional track provided for a portion of the corridor. Ultimately the rail corridor will consist of two passenger/ freight rail tracks and two HST tracks.

### Existing Stations

Challenge – There are five existing Metrolink/ Amtrak stations within the corridor, the largest being Los Angeles Union Station (LAUS). For LAUS, this will be the major point of connection for this corridor. There are also smaller Metrolink stations along the existing corridor including Burbank, Glendale, Sun Valley and San Fernando Stations.

Solution – The smaller Metrolink stations are generally situated on the eastern side of the right of way and would most likely need to be relocated or reconfigured as a part of the CHST project. Within the Sylmar to SR-2 section there will be one CHST station that will co-locate with a Metrolink station providing for connectivity between systems.

**Existing Road Network**

Challenge – The existing LAMTA rail corridor is generally bordered by San Fernando Road on the west and a frontage road on the east for a distance of 18 miles (referred to as Sylmar to SR-2). From Sylmar to SR-2 there are currently 30 grade crossings going across the existing LAMTA corridor with 23 of these being at grade. The remaining are existing grade separations that were built either as undercrossing or overcrossing grade separations.

Solution – In order to provide a functional HST corridor all grade separations will need to be eliminated.

**Built environment**

Challenge – The HST corridor will be passing through four cities including Los Angles, Burbank, Glendale, and San Fernando. The corridor passes through many different built environments including industrial, commercial, residential, and open space.

Solution – Graphic Information Systems (GIS) were use to help characterize these areas into different zones existing land use data was used from each of the cities. This helped group different areas by the planned use. It is recommended to pass through these areas at grade or on viaduct unless; one of the other constraints forces the HST to be in a trench or in a tunnel.

**Seismic zones**

Challenge – Two faults exist within this section, the San Fernando Fault and the Verdugo Fault. Both faults are deemed as potentially active and need to be studied further to understand the potential for fault rupture.

Solution – To minimize damage during a seismic event it is the intent to cross the fault zones at grade. This will also ensure that HST service could be restored quickly if a seismic event caused substantial damage to the infrastructure, by requiring HST to be at grade within seismic zones large structures either below or above ground would not have to be rebuilt.

**ENVIRONMENTAL CONSTRAINTS AND CHALLENGES**

**Historical and Archaeological sites**

Challenge – There are numerous historical sites throughout the corridor, particularly in the SR-2 to LAUS portion. Historical sites include historic bridges crossing the Los Angeles River, a state historic park with archaeological significance,

religious institutions, and other sites listed on the California Historical Resources Information System (CHRIS) [4].

Solution – All the historical sites were mapped within the vicinity of the corridor to identify potential conflicts. To avoid the historical sites it is recommended to go around them so as to not affect them or below them in a bored tunnel deep enough to avoid disruption.

**Waterways**

Challenge – The HST alignment will cross Tujunga Wash, Pacoima Wash, Verdugo Wash, Arroyo Seco and the Los Angeles River.

Solution – In order to not affect the waterways the 100 year flood zone was determined for each waterway taking data from the Los Angeles County Flood Control District. The vertical alignment will be set to provide adequate freeboard over the flood zone and by taking into account the structural depth and rail formation. Due to difficulty of construction and disruption of waterways it is recommended to cross above with a viaduct or below with a bored tunnel.

**Parks and Schools**

Challenge – The HST alignment will pass by many parks throughout the corridor with the major parks being Elysian Park, Los Angeles State Historic Park, and Rio de Los Angeles State Park. Additionally the HST corridor may pass by existing elementary schools, middle schools and high schools as well as a newly constructed high school.

Solution – To minimize impacts on these facilities tunnels and trenches are being considered.

## CONCLUSION

The design of the Sylmar to LAUS portion of the PL segment has many challenges. Beginning with the relocation of existing LAMTA tracks, relocation of Metrolink stations, grade separations throughout the San Fernando Valley, crossing two airports, parks, waterways and ending with the historical and archaeological sites in the SR-2 to LAUS portion of the project. This coupled with the stringent criteria meant to provide a high speed alignment where efficiency and passenger comfort are the paramount objective. By utilizing the various construction methods and recommending the appropriate solution for the particular challenge a solution can be obtained to provide the maximum benefit with the least cost and limited impacts.

**REFERENCES**

[1] California High-Speed Rail Authority, December 2009, “Report to the Legislature: Business Plan”.

[2] U.S. Department of Transportation, Federal Railroad Adminstration, April 2009, “Vision for High Speed Rail in America”.

[3] California High-Speed Rail Authority, March 2009, “Technical Memorandum: Alignment Design Standards for High-Speed Train Operation”.

[4] California High-Speed Rail Authority, July, 2010, “Preliminary Palmdale to Los Angeles Section: Alternatives Analysis Report”.

# TRAVEL TIME RELIABILITY ESTIMATION: USE OF MEDIAN TRAVEL TIME AS MEASURE OF CENTRAL TENDENCY

M. Arezoumandi[1] and G. H. Bham, Ph.D.[2]

[1] 302 Engineering Research Lab, 500 W. 16th Street, Rolla, MO 65409-0440, Phone: 573-341-6372, Email: ma526@mst.edu

[2] Missouri University of Science and Technology, Civil, Architectural and Environmental Engineering, 135 Butler Carlton Hall, 1401 N. Pine Street, Rolla, MO 65409-0030, Phone: 573-341-6286, Email: ghbham@mst.edu

## ABSTRACT

This paper proposes use of median travel time compared to mean travel time as the measure of central tendency to calculate Buffer Time Index (BTI), one of the most important measures of travel time reliability. This study examined gamma, exponential, largest extreme value, smallest extreme value, logistic, Weibull, lognormal and log-logistic distributions to determine the best fit to travel time data for two cases: 24 hours and peak hours for a highway with Variable Speed Limit (VSL) conditions. The goodness-of-fit tests indicated that the log-normal distribution best represented the travel time data for peak periods and log-logistic distribution better fitted the travel time data for 24 hours. The log-logistic distribution and log-normal distributions were positively skewed and for skewed distributions, median is the best estimate of central tendency. Hence, median travel time is proposed for calculating BTI. The difference in values of BTIs were evaluated when mean and median travel time were used. Travelers must add extra time to their average travel time to ensure on-time arrival. It was found that this extra time increases from 11% to 30% when median travel time is used as the average travel time compared to mean travel time for calculation of BTI.

## INTRODUCTION

Travel time reliability is one of the most important transportation performance measures as it allows travelers to make better use of their time and provides shippers reliable travel time information that allows them to remain competitive. Therefore, travel time reliability can be considered a key performance measure for planning purposes.

Previous studies have defined several methods to measure travel time reliability. Polus (1) defined the measure of travel time reliability on arterial routes as the inverse of the standard deviation of the travel time distribution. The Florida Department of Transportation (2) used the percentage of travel time less than the median travel time plus a certain acceptable additional time such as 5%, 10%, 15%, and 20% above the expected travel time to estimate travel time reliability. Chen et al. (3) defined the measure of travel time reliability by the variability between the average travel time and the actual travel time in terms of the standard deviation of travel time distribution. Lomax et al. (4) presented three measures of travel time reliability: 1) the percent variation which was represented with respect to the average travel time as a percentage measure; 2) the misery index, represented by the average number of minutes that the worst trips exceed the average; and 3) the buffer time index which represented the amount of extra time needed to be on time for 95% of the trips.

The Federal Highway Administration (FHWA) (5) used the following measures to quantify travel time reliability including: 1) the $90^{th}$ or the $95^{th}$ percentile travel time which indicates the amount of delay on the heaviest travel days; however, this measure is not easily compared across trips with different lengths, 2) the travel time index represents the ratio of the average travel time to the free-flow travel time, 3) the buffer time index is defined as the difference between the 95th percentile travel time and average travel time, and then divided by average travel time, which represents the percentage of extra travel time that most travelers add on to their trip in order to ensure on-time arrival. 4) the planning time index is defined as the ratio of the 95th percentile travel time to the free-flow travel time, which represents total time needed for planning a trip to be on-time arrival 95% of the time. 5) The congestion frequency which represents congestion which exceeds some expected threshold such as the percent of days or time that travel speed falls below a desired speed. Bogers et al. (6) stated that different travel time reliability measures such as the variance of travel time, buffer time index, and the misery index may draw different conclusions even when applying to the same data. They proposed the degree of skew of the travel time distribution. Bertini et al. (7, 8) used travel time reliability measures to prioritize freeway segments according to their travel time reliability which can be beneficial in regional transportation planning, operations, and bottleneck prioritization.

Research has used different distributions to model travel time. Emam and Al-Deek (9) compared the log-normal, gamma, Weibull, and exponential distributions for modeling of travel time data and concluded that a log-normal distribution provided the best fit. Hsing-Chung Chu (10) compared the gamma, largest extreme value, log-logistic, log-normal, and Weibull distributions for modeling of travel time data for freight trips during mid-day and concluded that a log-logistic distribution provides the best fit.

The main scope of this study is to identify the distributions that provide the best fit for both 24 hours and peak periods. According to this finding we can state

median or mean, which one is better estimate as average travel time and also correcting the indexes depend on average travel time such as BTI, one of the most significant factors in travel time reliability.

## STUDY SITE AND TRAFFIC DATA

This study selected three congested segments of I-270 in St. Louis County, Missouri, USA for studying travel time. The segments were 1.4, 3.7, and 4.7 miles long. Travel time was estimated from 30-second time mean speed data measured by remote traffic microwave sensors (RTMS). The Variable Speed Limit (VSL) system was initiated on May, 2008 on this highway. This analysis aggregated archived data over 5-minute intervals for clear weather conditions from October, 2008 to May, 2009 as post-VSL conditions.

Some segments of I-270 have five lanes, others have four. The lanes next to the right most lanes, lanes 2 and 3 had the most stable traffic flow, and changes to traffic flow patterns in these two lanes are found to be more than random fluctuations in traffic. Lanes 2 and 3, therefore, were the focus of this research. Week days were selected for the analysis. Figure 1 shows their locations.

.

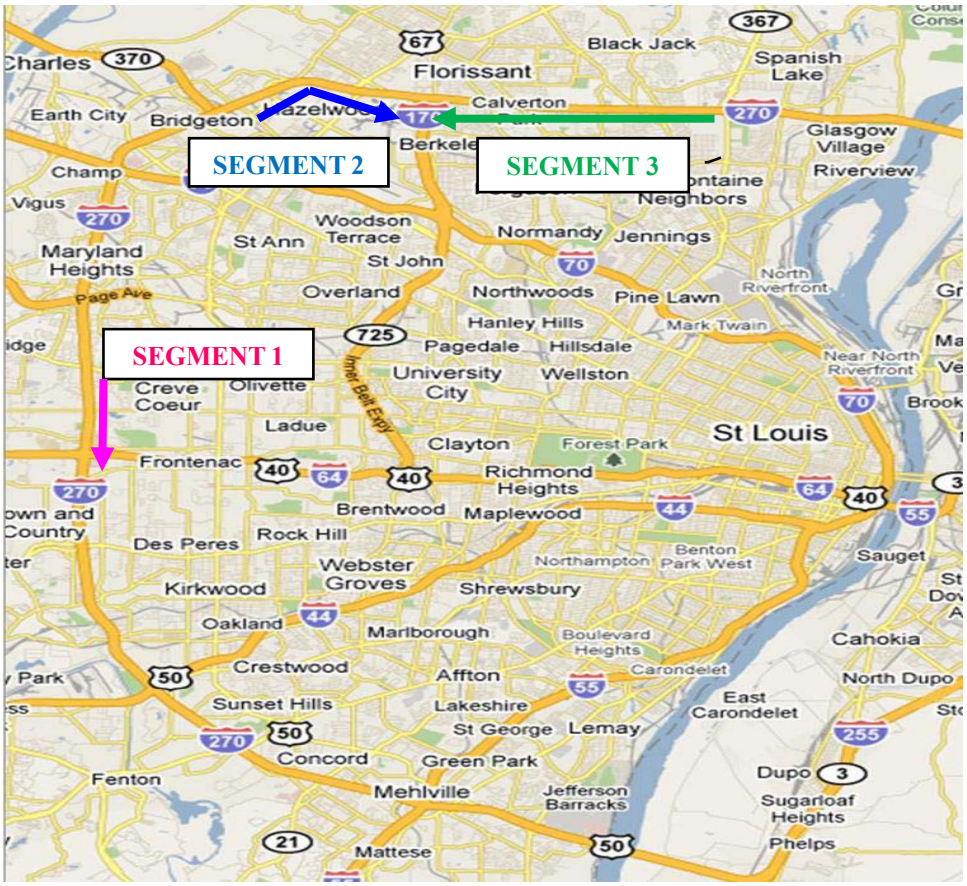



Figure 1. Segment of I-270. (Source: Google map)

Segments 1, 2 and 3 were evaluated using data from two, four and four detectors, respectively.. Segments 1 and 2 had evening peak periods (15:00 to 19:00 p.m.) but segment 3 had morning peak periods (6:00 to 10:00 a.m.).

The travel times were determined between detector stations as follows (Figure 2):

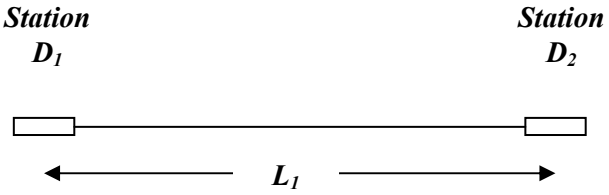



Figure 2. Detector Stations and Distance between them

$$T_{1-2} = \left[\frac{60}{2} \times \left[\frac{L_1}{V_{D_1}} + \frac{L_1}{V_{D_2}}\right]\right] \quad (1)$$

where:

$T_{1-2}$ = travel time from station $D_1$ to station $D_2$, minutes,

$L_1$ = distance between station $D_1$ and station $D_2$, miles,

$V_{D1}$, $V_{D2}$ = average speed measured at station $D_1$ and station $D_2$, respectively, mph.

## METHODOLOGY

This paper examined gamma, exponential, largest extreme value, smallest extreme value, logistic, Weibull, lognormal and log-logistic distributions to identify which best fits the distribution of travel time data in two cases: 24 hours and peak hours. Two skewed distributions: log-logistic and log-normal provided the best fit for travel time data.

In symmetrical distributions the mean and median are the same as in normal distribution. In asymmetrical distributions, by contrast, the mean and median are not the same. Such distributions skewed; that is, more than half of the data are either above or below the mean. The three measures of central tendency (mode, median, mean) converge for a distribution, to be symmetric. For skewed distribution the three measures of central tendency diverge.

Log-logistic distribution and log-normal distribution are positively skewed. The median is between the mode and the mean, and the mean of the highest estimate of central tendency (Figure 3).

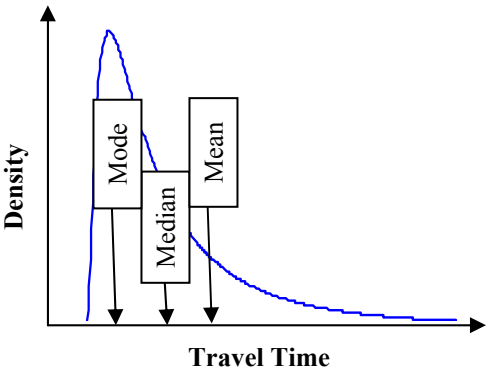



Figure 3. Measuring of Central Tendency for Skewed Distribution

Selecting an appropriate measure of central tendency ensures calculation of accurate indices of travel time reliability. Thus, median is preferable to the mean as an average to calculate the BTI, one of the most important measures of travel time reliability.

**Buffer Time Index**

The buffer time index represents the extra time (or time cushion) that travelers must add to their average travel time when planning trips to ensure on-time arrival (5). The BTI is calculated as:

$$\text{Buffer Time Index (\%)} = \left[\frac{\substack{\text{95th Percentile Travel Time} - \text{Average Travel Time} \\ \text{(minute)} \qquad\qquad \text{(minute)}}}{\substack{\text{Average Travel Time} \\ \text{(minute)}}}\right] \tag{2}$$

**Log-Logistic Distribution:** The probability density function (PDF) of this distribution is defined as:

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{-2} \qquad \gamma \le x < \infty \tag{3}$$

where:

$\alpha$ = shape parameter ($\alpha > 0$)
$\beta$ = scale parameter ($\beta > 0$), and
$\gamma$ = location parameter

**Lognormal Distribution:** The probability density function (PDF) of this distribution is defined as:

$$f(x) = \frac{\exp\left(-\frac{1}{2}\left(\frac{lnx-\mu}{\sigma}\right)^2\right)}{x\sigma\sqrt{2\pi}} \qquad 0 < x < \infty \tag{4}$$

where:

$\sigma$ = scale parameter ($\sigma > 0$), and
$\mu$ = location parameter

Statistical computer softwares Minitab and SAS were used to compute the shape and scale parameters of these distributions as well as to examine the best fit from the distributions to the data set. The distribution graphs of probability density functions (PDF) were used to visually compare the difference among the distributions. Additionally, the Anderson-Darling goodness-of-fit was applied to select the best fit distribution. This statistic measures how well the data follow a particular distribution. For a given data set and distribution, the better the distribution fits the data, the smaller value of the statistic.

Minitab calculates the Anderson-Darling statistic using the weighted squared distance between the fitted line of the probability plot (based on the selected distribution and using either maximum likelihood or least squares estimates) and the nonparametric step function. This calculation is weighted more heavily in the tails of the distribution.

## ANALYSIS OF RESULTS

Results of goodness-of-fit tests showed that for the first case (24 hours) log-logistic distribution fits better than log-normal distribution, whereas lognormal distribution provided best fit for peak hours.

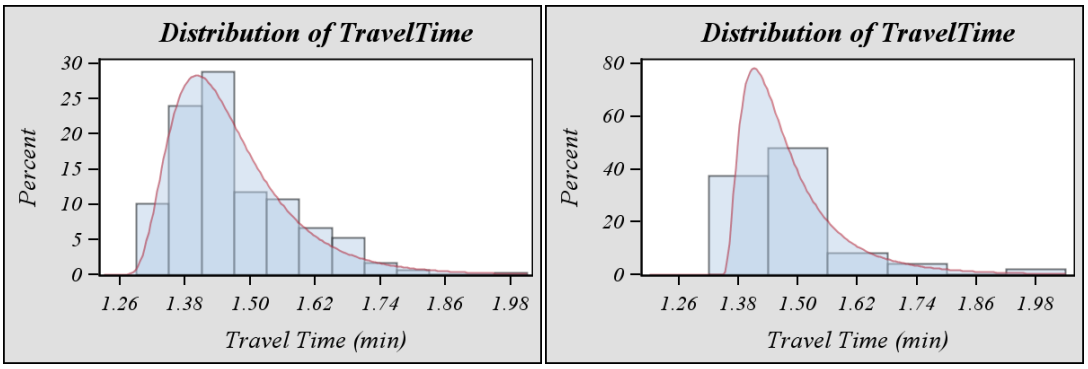
Table 1 presents the results of goodness-of-fit tests performed using the Anderson-Darling test for segment 1. The Anderson-Darling test values for loglogistic distribution were smaller than the lognormal distribution for 24 hours, which indicates that loglogistic distribution fits better for 24 hours travel time. For peak hours Anderson-Darling test values for lognormal distribution were smaller than the loglogistic distribution, which shows that the lognormal distribution fits better for peak hours of travel time. Figure 4 shows a three-parameter loglogistic curve fit for 24 hours and peak hours travel time of data. From the figure it is clear the loglogistic distribution better fits the data for 24 hours.

Table 1. Anderson-Darling Tests for Travel Time Distributions on Segment 1

| Dates | Lane | **24 hours** | | **Peak hours** | |
|---|---|---|---|---|---|
| | | Log-normal | Log-logistic | Log-normal | Log-logistic |
| 6th Oct. | 2 | 4.95** | 3.03** | 1.65** | 1.83** |
| | 3 | 3.70** | 2.09** | 1.77** | 1.98** |
| 23rd Oct. | 2 | 0.38* | 0.19* | 0.60* | 0.75* |
| | 3 | 0.98* | 0.58* | 0.46* | 0.60* |

* 5% significance level

** 1% significance level




**a) 24 hours** **b) Peak hours**

Figure 4. Three-Parameter Loglogistic Fit Curve for Travel Time Data

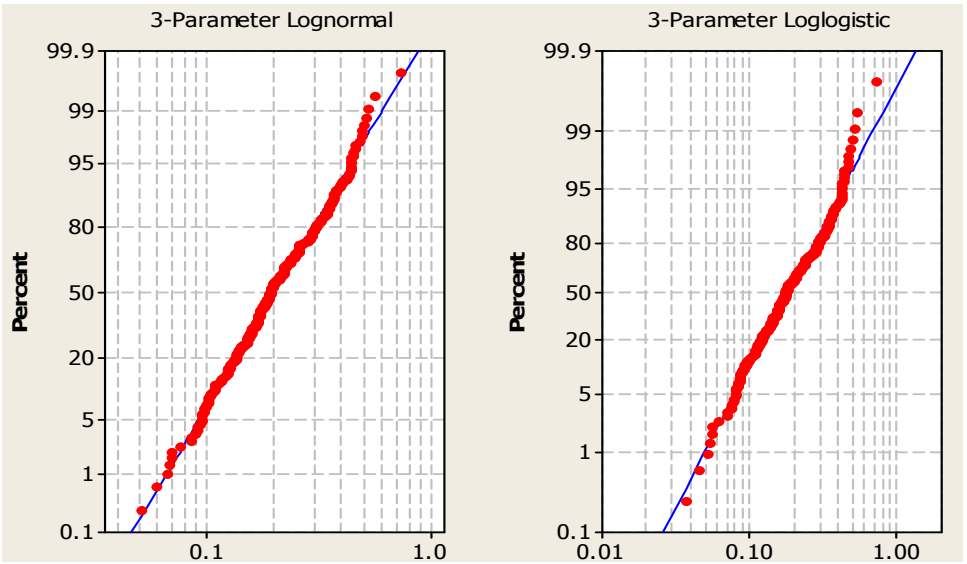



Figure 5. Probability Plot for Travel Time Distribution **(24 hours)**

Figures 5 and 6 show the probability plot for lognormal and loglogistic distribution for 24 hours and peak hours for 23rd October, 2008 travel time data, respectively. The points on this plot form a nearly linear pattern, which indicates that the lognormal and loglogistic distribution are good model for data sets.

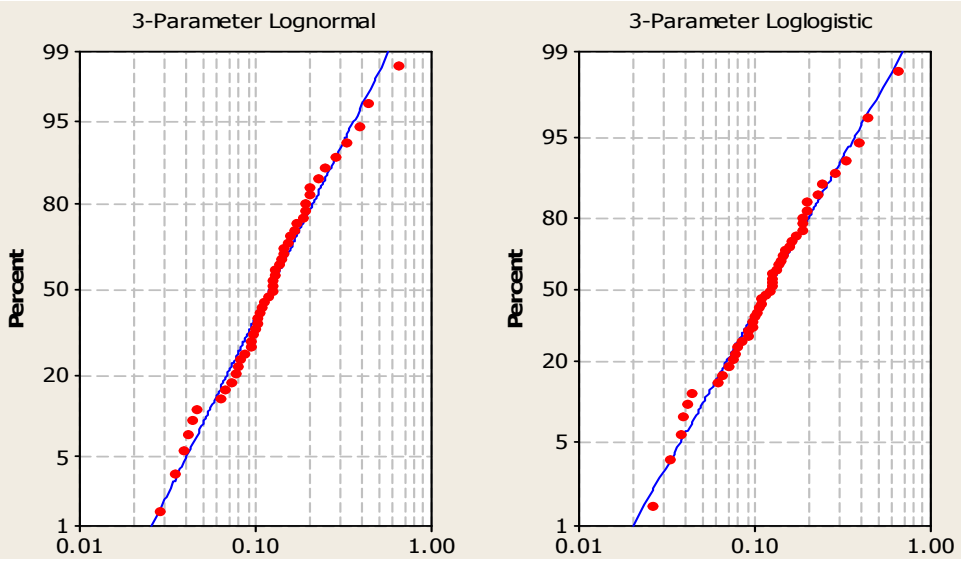



Figure 6. Probability Plot for Travel Time Distribution **(Peak hours)**

This research calculated mean and median travel times for 24 hours and peak hours. The results for 24 hours show the mean is higher than the median by 1% to 4% on all segments, but during peak hours the difference between mean and median was greater than the 24 hours period. The mean is higher than the median by 2% to 7%, by 2% to 11%, and by 2% to 15% on segment1, 2, and 3 respectively. Table 2 compares mean and median travel times for segment 1.

Table 2. Comparison of Mean and Median Travel Time on Segment 1 (minutes)

| Dates | Lane | **24 hours** | | | **Peak hours** | | |
|---|---|---|---|---|---|---|---|
| | | Mean | Median | Difference | Mean | Median | Difference |
| 6th Oct. | 2 | 1.42 | 1.39 | 2% | 1.52 | 1.41 | 7% |
| | 3 | 1.42 | 1.40 | 1% | 1.49 | 1.40 | 6% |
| 23rd Oct. | 2 | 1.47 | 1.44 | 2% | 1.53 | 1.48 | 3% |
| | 3 | 1.47 | 1.44 | 2% | 1.49 | 1.46 | 2% |

In this research, BTIs were calculated for two cases, the first using mean as average travel time and the second using median as average travel time. The results of comparison indicate that the BTI calculated with mean is lower than that calculated with median by 12% to 22%, by 1% to 36%, and by 8% to 30% on segments 1, 2, and 3 respectively (Table 3).

Table 3. Comparison of BTIs on Segment 1

| Dates | Lane | BTI (Mean) | BTI (Median) | Difference |
|---|---|---|---|---|
| 6th Oct. | 2 | 0.39 | 0.50 | 22% |
| | 3 | 0.45 | 0.54 | 17% |
| 23rd Oct. | 2 | 0.34 | 0.39 | 12% |
| | 3 | 0.16 | 0.18 | 13% |

For example, in lane 2 on October 6th, 2008 on segment 1, mean equaled 1.52 minutes and median equaled 1.41 minutes, and BTI with mean was found to be 39% which meant the traveler should allow extra (.39*1.52*60) 36 seconds to their average travel time for the trip in order to ensure on-time arrival 95 percent of the time, but BTI with median was 50% which meant that the traveler must add (.5*1.41*60) 42 seconds to their average travel time when planning trips to ensure on-time arrival 95 percent of the time.

Similarly, in lane 2 on segment 3, mean equaled to 6.95 minutes and median equaled 4.98 minutes, and BTI with mean was found to be 74% which meant the traveler should allow extra (.74*6.95) 5.1 minutes to their average travel time for the trip in order to ensure on-time arrival 95 percent of the time, but BTI with median was 143% which meant that the traveler must add (1.43*4.98) 7.1 minutes to their average travel time when planning trips to ensure on-time arrival 95 percent of the time.

## CONCLUSIONS

Improvements in travel time reliability can reduce uncertainty regarding the time required for passenger trips and movement of goods. This study which uses VSL travel time data on a congested segments of I-270 in St. Louis County.

The results of the goodness-of-fit tests indicated that the loglogistic distribution best represented the travel time data for peak periods, but lognormal distribution better fitted the travel time data for 24 hours.

The loglogistic and log-normal distributions are positively skewed and for skewed distributions, median is the best estimate of central tendency and the mean is always the highest estimate of central tendency. Hence, median should be used in lieu of mean as average travel time for calculation of BTI. For selected segments in this study mean was higher than the median travel time from 1% to 15%. This shows that mean travel time is up to 15% over estimate of average travel time.

This study evaluated the difference in BTI when mean and median travel time were used. It was found that when median was used, BTI was higher from 1% to 36% compared to using the mean, which means additional times needed to ensure on-time arrival can be up to 36% higher than extra time calculated using mean travel time for BTI.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Missouri Department of Transportation (MODoT) for providing the data used herein.

**REFERENCES**


1. Polus, A. (1979). "A Study of Travel Time and Reliability on Arterial Routes." Transportation, Vol. 8. pp. 141-151.
2. Florida Department of Transportation. (2000). "The Florida Reliability Method."In Florida's Mobility Performance Measures Program.
3. Chen, C., Skabardonis, A., and Varaiya, P. (2003). "Travel Time Reliability as a Measure of Service." In Transportation Research Record: Journal of the Transportation Research Board, 1855, pp. 74−79.
4. Lomax, T., Sharnk, D., Turner, S., and Margiotta, R. (2003). " Selecting Travel Reliability Measures." FHWA, U.S. Department of Transportation.
5. Federal Highway Administration (FHWA). (2006). "Travel Time Reliability: Making It There On Time."
6. Bogers, E. A. I., Van Lint, H.W.C. and Van Zuylen, H.J. (2008). " Reliability of Travel Time: Effective Measures from a Behavioral Point of View." In Transportation Research Record: Journal of the Transportation Research Board, 2082, pp. 27−34.
7. Saberi, M.K., and Bertini, R. L. (2010). "Beyond Corridor Reliability Measures: Analysis of Freeway Travel Time Reliability at the Segment Level for Hotspot Identification", Presented at the 89th Annual Meeting of the Transportation Research Board, Washington, D.C.
8. Lyman, K. and Bertini, R. L. (2008). " Using Travel Time Reliability Measures to Improve Regional Transportation Planning and Operation," In the Transportation Research Record: Journal of the Transportation Research Board, 2046, pp. 1-10.
9. Emam, E. B., and Al-Deek, H. M. (2006)."Using Real-Life Dual-Loop Detector Data to Develop New Methodology for Estimating Freeway Travel Time Reliability." In *Transportation* Research *Record: Journal of the Transportation Research Board,* 1959, pp. 140-150.
10. Hsing-Chung Chu. (2010). "Estimating Travel Time Reliability on Freight Corridors," Presented at the 89th Annual Meeting of the Transportation Research Board, Washington, D.C.

# SYSTEMATIC EVALUATION OF A VARIABLE SPEED LIMIT SYSTEM IN ST. LOUIS MISSOURI

Khushboo Lall and Ghulam H. Bham. PhD
Civil, Architectural and Environmental Engineering
Missouri University of Science & Technology
135 Butler Carlton Hall, 1401 N. Pine Street, Rolla, MO 65409
Email: ghbham@mst.edu, Phone: 573-341-6286, Fax: 573-341-4729

## ABSTRACT

This paper presents the systematic evaluation of Variable Speed Limit (VSL) system implemented on I-270 in St. Louis County. The system varies the speed limit in real time during periods of recurring and non-recurring congestion to prevent traffic from breaking down. The system was evaluated objectively and subjectively, and this paper describes the objective evaluation of the system. The objective evaluation was performed by comparing the traffic flow parameters during static speed limit (pre-VSL) with variable speed limit (post-VSL) by using five measures of performance: 1) speed, 2) volume, 3) occupancy, 4) travel time, and 5) congestion measures. Identification of congested segments and the time periods for the analyses were the initial steps for evaluating the system. Traffic data were provided by MoDOT collected using Remote Traffic Microwave Sensors. Average speed, volume and time occupancy profiles indicated that the VSL system change in traffic conditions including congestion. Preventing the highway from reaching its capacity and avoiding the traffic breakdown is one of the main objectives of VSL system initiation. It can be achieved by upstream drivers reducing speed, metering the upstream traffic and preventing traffic from breaking down. Congestion measures evaluation indicates change in Travel Time, Percentage of Congested Travel and reduction in duration, extent and intensity of congestion. The VSL system benefits were determined, which indicated that the system benefits will pay for the initial capital cost of the system over the years. This study can be used by Departments of Transportation for evaluation of active management systems.

## INTRODUCTION

Variable Speed Limit (VSL) signs were initiated on May 22$^{nd}$ 2008 along 38 miles of I-270/I-255 in St. Louis County to vary speed limit during recurring and nonrecurring periods to manage congestion. The system allows speed limits to be varied in real time based on actual traffic conditions. This is the first of its kind implementation in the state of Missouri and this study evaluates the system based on different measures of performance. The main purpose of the VSL system initiation was to improve traffic flow by preventing the highway from attaining capacity, thereby relieve congestion, and reduce travel time.

The objectives of this study were to evaluate the performance measures by comparing traffic data during pre-and post-VSL condition to quantify the system benefits by evaluating delay and cost savings. For the comparison data provided by the Missouri Department of Transportation were 30 sec average speed, occupancy and volume along I-270/I-255.

## LITERATURE REVIEW

Studies carried out in Europe, Canada and USA indicated that implementation of variable speed limit was effective. Speed-flow occupancy relationships, travel time estimation and congestion measures were carried out to analyze the different aspects of VSL system. These parameters were the measures of effectiveness for this new technology.

Papageorgiou et al. (2006) examined whether the speed limits modify the shape of the flow–occupancy diagram. A total of more than 800 km of VSL-equipped motorway in United Kingdom were used for the evaluation. Curves fitted to the flow-occupancy relationship were used to analyze the effect of speed limit. The VSL system shifted the critical occupancy to higher values in the flow–occupancy diagram and when applied at under-critical occupancies, speed limits have the effect of decreasing the slope of the flow–occupancy diagram. The difference in flow was found out to be small. Further research was recommended to obtain better results. Steel et al. (2005) carried out evaluation of the VSL system on I-90 in the US across the Snoqualmie Pass operated by the Washington State Department of Transportation. The comparison indicated decrease in the mean operating speed but increase in the deviation. They recommended increase in speed enforcement to minimize the deviation.

Riffkin et al. (2008) evaluated the VSL system in Utah by comparing three months' of daily average speed for day and night conditions at different locations. They tested the hypothesis whether the change in speed was due to randomness or variable speed limit initiation. Similar comparison was also carried out by Mott MacDonald from the United Kingdom. Data analyzed was for one year before and after the VSL implementation. The data were categorized into weekdays and weekends based on traffic flow. Graphical comparisons were carried out for speed compliance and capacity analysis. The level of non-compliance was calculated by taking the total number of offences at the Association of Chief Police Officers (ACPO) threshold for each speed limit (70 mph to 40 mph), at all sites, as a percentage of the sum of the total number of vehicles at each site per month. The results obtained were consistent. When considering the ACPO threshold of enforcement (speed limit + 10% + 2 mph), compliance on the main carriageway was on average 94%, or better, at the 70 mph, 60 mph, 50 mph speed limits and 84%, or better, at the 40 mph speed limit between January 2006 and September 2007. The $80^{th}$ percentile of the cumulative frequency for the 15 minute maximum flow observed over a 24 hours was used as a measure to estimate the change in capacity between the 'Before' and 'After' cases. It was assumed that the top 20% of the observed 15 minute maximum flows represent cases of unsustainable flow or unusual high throughput, due to a temporal change in traffic conditions. Increase in capacity and compliance in speed indicated improvement in traffic conditions.

Bertini et al. (2006) compared the variable speed limit (VSL) and traveler information presented to commuters with actual traffic conditions on the highway. Detector data used for comparison were extracted from a segment of German Autobahn near Munich, Germany. Results indicated strong correlation between the two data sets. Results indicated that when drivers were warned of approaching congested conditions, the speed limits were reduced before bottleneck activation. The system reduced the speed limit to control the congested but still flowing traffic, and traffic continued to flow during congested periods at low speeds (19 and 25 mph). Comparison of flow and speed data with fundamental diagrams of speed–flow and flow–density indicated that reducing the speed limits upstream reduced congestion downstream. This paper applies the findings from the literature above for the study presented in the paper.

## METHODOLOGY

VSL system evaluation is a step-wise process with three distinct steps. In the first step the congested segments and time periods for evaluation were identified. Next, was traffic data evaluated to determine the peak periods, and change in volume, speed and occupancy. Based on the results, if the system shows benefits, congestion measures were evaluated and benefits were calculated. This paper presents the systematic for evaluation of the VSL system on I-

255/I-270. This system is activated to vary speed limits during recurring and non-recurring congestion. The impact of this system on the aggregated traffic flow is reflected on the traffic data analysed.

**Selection of Segments and Time Periods**

The very first step was identification of congested segments. Congestion was defined as detector occupancy above 7% and average speed under 40 mph. Congested segments were identified from the archived data using the time space plots. Second step involved selection of time periods for evaluation. Time period can be divided based on weather conditions and the duration for which the VSL system was active.

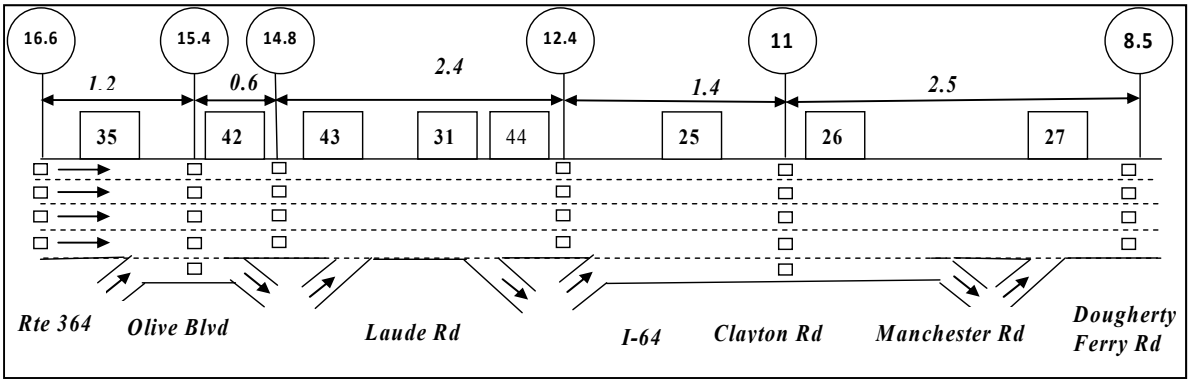



**Figure 1: Selected Segment for Evaluation, I-270 SB**

For this paper, the section of I-270 used for evaluation is displayed in Figure 1. This segment was found to be one of the most congested segments on the highway. In Figure 1, circles indicate the mile-marker (logmile)s for detector locations and squares present the locations of VSL signs along the highway. VSL system evaluation was carried out mainly for peak periods since the VSL system stays active during peak periods during recurring congested conditions. For off-peak condition VSL signs display speed limit of 60 mph. Days with clear and inclement weather conditions were separated. For this paper data clear weather days were used for comparison of pre-and post-VSL conditions. The data analysis was carried out for eight days in pre-and same in post-VSL for clear condition for the segment.

**Evaluation of VSL system**

Data available from one year of pre-and post-VSL traffic including average speed, vehicle count and average occupancy were used for the analysis. Data used for evaluation were averaged out for 15 minutes, 5 minutes and 1 minute for different evaluation purposes. To study the effect of volume, various volume profiles were plotted and the differences were recorded. Change in traffic volume was accounted to discern its effect on average speed, travel time and congestion measures. Traffic volume or flow versus occupancy curves were plotted for analysis of critical occupancy. The critical occupancy is defined as the occupancy at which highest flow (volume) is observed. One of the purposes of VSL system initiation is to reduce average speed upstream at under critical occupancy, since the arriving upstream demand is to be served without congestion.

For average speed performance measure, the objective was to compare the mean speeds before and after the initiation of the VSL system. The speed differential, the mean, standard deviation, and statistical analysis of speeds between lanes were evaluated. The objective of this step was the evaluation of the posted variable speed limit during peak periods. It concentrated in evaluating the speed during peak periods mainly because VSL system is most

active during peak periods. This evaluation determined the effectiveness of the system logic and the traffic flow parameters used for determining the appropriate speed limit. Evaluation of congestion measures is based on the results obtained from volume and speed analysis. If the results observed are positive from VSL system initiation, congestion measures are calculated to quantify the benefits.

The traffic flow characteristics before congestion, during congestion and after congestion were evaluated. More specifically increase or decrease in average speed, and the length of time before traffic reaches the congested state, how long congestion lasts, how long are the queues, what is the impact of the VSL system on the length of queue, would be evaluated. The performance measures that can be used are Total Delay, Congested Travel, Percent of Congested Travel, and Congested Roadway.

The congestion measures were based on travel times for the segment, delay as a result of change in travel time, and variables related to queue formation. For calculation of travel time aggregated archived data over 5-minute intervals were used. These variables quantify the effect on traffic congestion for pre-and post-VSL conditions. Total Delay and Percent of Congested Travel are important variables. The total delay for the roadway segment is the sum of time lost due to congestion. The percent of congested travel is an extension of the congested travel measure. It also measures the extent of congestion.

Queue clearance is defined as the time taken for traffic to pass through a bottleneck under congested conditions of the freeway. Queue extent is defined as the length of highway under the effect of congestion, the duration can be defined as the time for which the congestion lasted and intensity can be defined as values of average speed below 35mph. Average speeds aggregated for one minute were used for this step. A two dimensional time space plots were generated for the highway segment and average speeds were indicated using different shades to indicate congestion and free flowing traffic. The duration was measured on the time axis (x-axis), whereas the extent and intensity can be observed on the space axis (y-axis). For this paper, highway was considered congested when average speed went below 40 mph. Areas portraying low speed indicate congestion and from different shades indicate different intensities of congestion.

## COMPARISION RESULTS

Results from evaluation of VSL system impact on the segment is presented in this section. Traffic data from detector at mile-marker (logmile) 11 (Figure 1) on pre- (25th October, 07) and post (23rd October, 08) is used for demonstration of results. Figure 2 presents volume comparison for pre-and post-VSL conditions.

When average volume for pre-and post-VSL condition were compared (Figure 2), variation in average volume from -6% - 11% were observed, which is presented in Table 1. Since the percentage change observed in volume was small for pre- and post-VSL conditions, no adjustments were required for system evaluation. Also it was found that peak period occurred between 1500 hrs and 1730 hrs. Since the VSL system is active only during peak periods, therefore, the result for comparison of average volume is presented only for peak period at 15 minute interval.

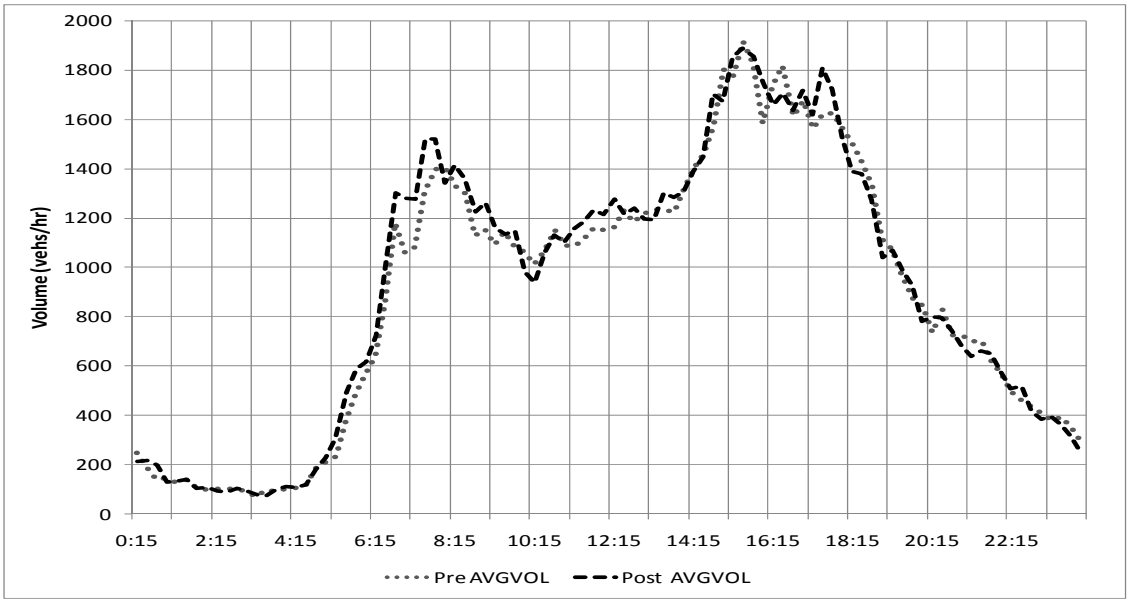



**Figure 2. Comparison of Pre (25th Oct, 2007) and Post-VSL (23rd Oct, 2008) Vol at mile-marker 11**

**Table 1: Average Volume at Detector 11D during Peak Periods**

| Time | 1500 | 1515 | 1530 | 1545 | 1600 | 1615 | 1630 | 1645 | 1700 | 1715 | 1730 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Pre-VSL Vol. (vph) | 1778 | 1916 | 1828 | 1584 | 1739 | 1817 | 1613 | 1675 | 1561 | 1624 | 1625 |
| Post-VSL Vol.(vph) | 1851 | 1886 | 1855 | 1745 | 1657 | 1705 | 1636 | 1715 | 1618 | 1807 | 1720 |
| Difference | 4% | -2% | 1% | 10% | -5% | -6% | 1% | 2% | 4% | 11% | 6% |

Figure 3 presents average occupancy for all five lanes at detector 11D for pre- and post -VSL system installation. Also, it can be noted that the maximum observed occupancy for pre-VSL condition was nearly 30 percent which reduced to 15 percent after VSL installation, which shows decrease in traffic congestion.

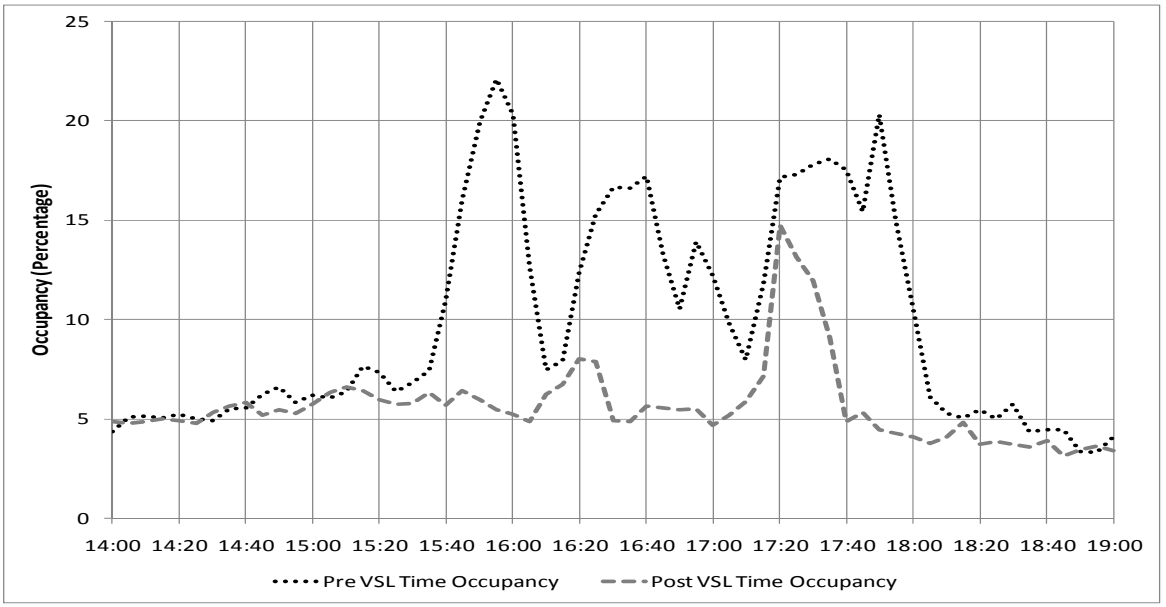



**Figure 3. Comparison of Pre (25th Oct, 2007) and Post-VSL (23rd Oct, 2008) Occup at mile-marker 11**

Figure 4 shows the comparison of pre- and post-VSL traffic volume and occupancy curves. Data used for flow occupancy plot is aggregated for 15 minutes. It can be observed from the figure that the VSL system installation has improved traffic flow as the occupancy is less than 10%. It can also be noticed that for post-VSL conditions, time occupancy reduced from a high value of nearly 20 percent to less than 12 percent. Also, fewer data points for the post-VSL conditions indicate better traffic flow indicating that the highway did not reach capacity. One of the objectives of the VSL system initiation was to prevent the highway from reaching capacity since after reaching capacity the volume breaks down and traffic flow can decrease significantly. Therefore, flow occupancy comparison indicates benefits of the VSL system initiation.

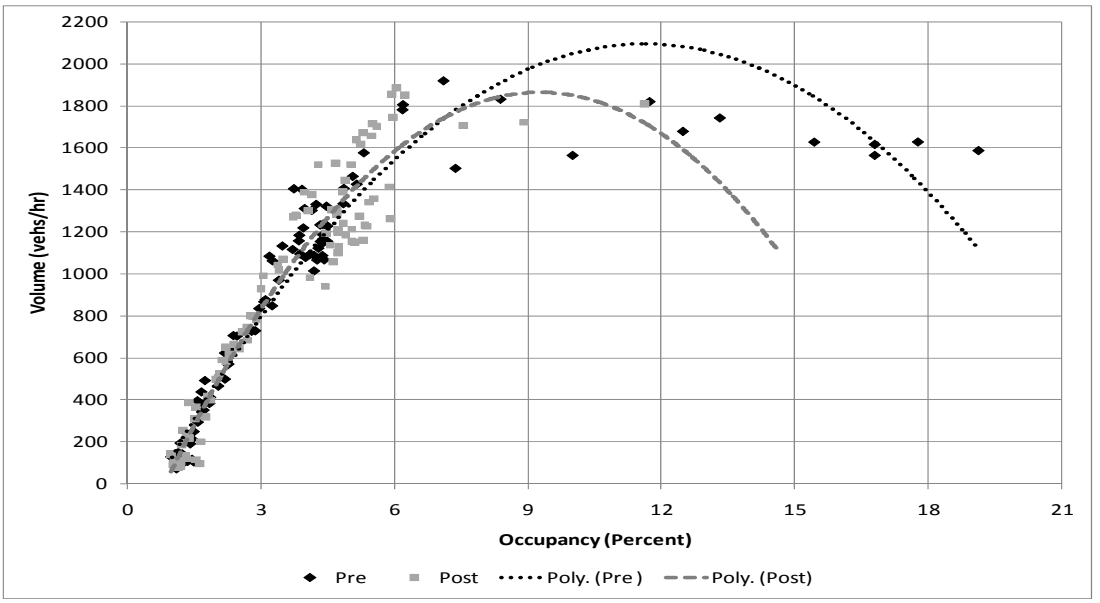



**Figure 4. Comparison of Pre (25th Oct, 2007) and Post-VSL (23rd Oct, 2008) Flow Occup at mile-marker 11**

Figure 5 presents the comparison of average speed profile for pre- and post-VSL system installation. Figure 5 indicates that the peak periods for pre-VSL lasted approximately two and half hours i.e. from 1545 hours to 1815 hours. During the peak period, the average speed reduced to below 30 mph in pre-VSL conditions but not in post-VSL conditions. Figure 5 also indicates gradual change in average speed due to the VSL system, compared to sharp decreases in average speed for the pre-VSL system. The average speeds were higher than 40 mph for most of the peak period because of which the variation in average speeds were lesser than pre-VSL conditions. The gradual decrease in speed for the post-VSL conditions may also reduce the speed variation across the lanes and along the highway segment. Figure 5 indicates that VSL has maintained the average speed between 40 mph and 60 mph which are within the minimum and posted speed limits for I-270.

Table 2 presents the standard deviation of speeds during the peak period for pre-and post-VSL conditions for detector data at 11D for pre-and post-VSL conditions, respectively. The difference indicates the change in post-VSL traffic conditions compared to pre-VSL conditions. Negative value indicates decrease in deviation of speed across the lanes, which indicates increase in speed homogeneity. Increase in speed homogeneity can cause reduction in traffic crashes which is a positive outcome of the VSL system.

Figure 6 presents the comparison of average speeds aggregated for the peak period of 1545 hours to 1815 hours averaged for all four/five lanes for pre- and post-VSL system installation at the detector locations on segment 1. The line plots compare the average volume during the

peak periods. The figure shows detectors upstream, on and downstream of segment 1. Figure 6 also presents the difference in average speed during peak periods at every detector location.

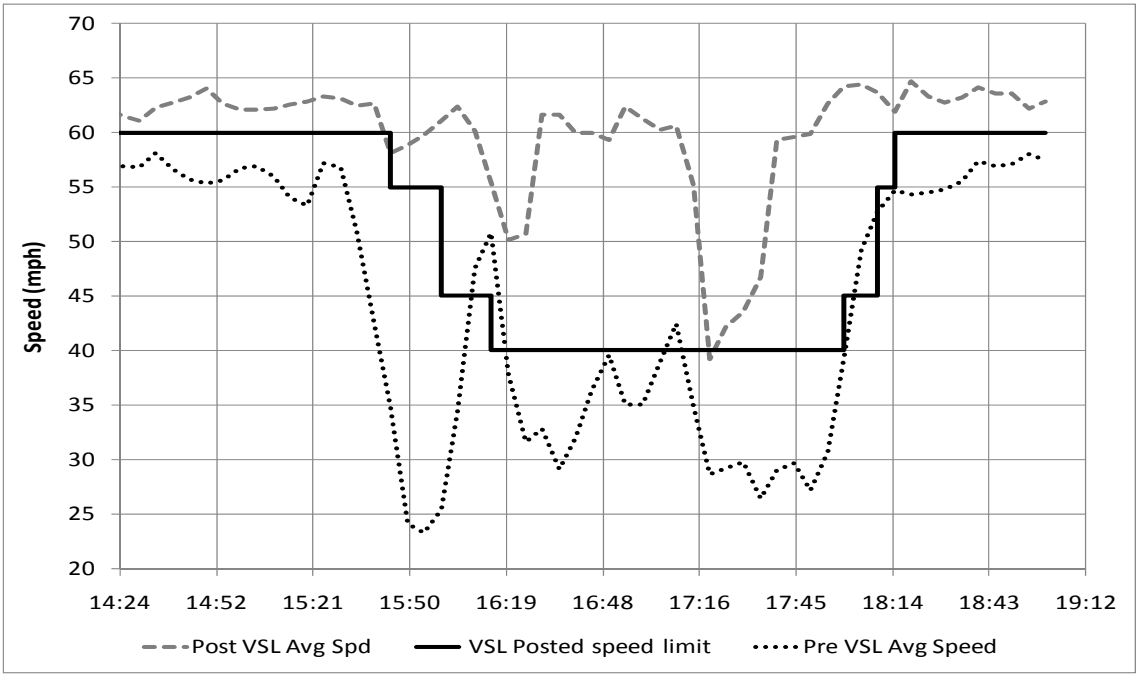



**Figure 5. Comparison of Pre (25th Oct, 07) and Post-VSL (23rd Oct, 08) Spd data at mile-marker 11**

**Table 2: Speed Deviation at Detector mile-marker (mile-marker (logmile)) 11 during Peak Periods**

| Time | 15:15 | 15:45 | 16:15 | 16:45 | 17:15 | 17:45 | 18:15 | 18:45 |
|---|---|---|---|---|---|---|---|---|
| Pre-VSL SD (mph) | 5.07 | 9.30 | 10.26 | 15.58 | 11.59 | 14.44 | 8.11 | 5.73 |
| Post-VSL SD (mph) | 4.31 | 5.38 | 5.30 | 5.42 | 5.80 | 12.63 | 5.65 | 5.93 |
| Difference (mph) | -0.75 | -3.92 | -4.96 | -10.15 | -5.79 | -1.81 | -2.46 | 0.20 |

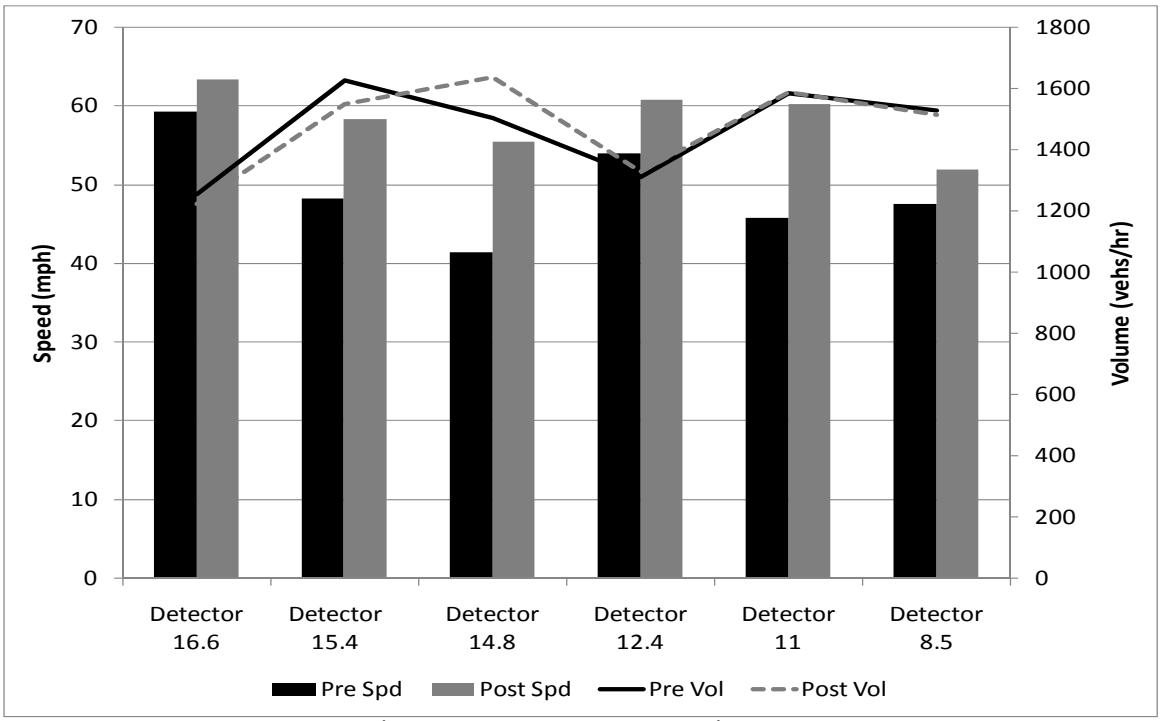



**Figure 6. Comparison of Pre (25th Oct, 07) and Post-VSL (23rd Oct, 08) Average Speed and Volume**

Figure 7 presents the speed flow data for pre-and post-VSL system installation. Data used were aggregated for 15 minute interval for this step. Congested and uncongested data points are indicated separately. It can be observed from the figure that VSL system installation has maintained the speed flow curve on the un-congested state providing better traffic conditions. Fewer data points were observed in the congested regime of the speed flow curve indicating better traffic flow conditions for post-VSL conditions. It can also be inferred that after VSL system installation, the segment did not reach capacity; hence no breakdown in traffic flow was observed. From the above discussion it can be stated that VSL system initiation prevented the highway from reaching capacity, though the maximum flow in both conditions were around 1900 vehicles per hour.

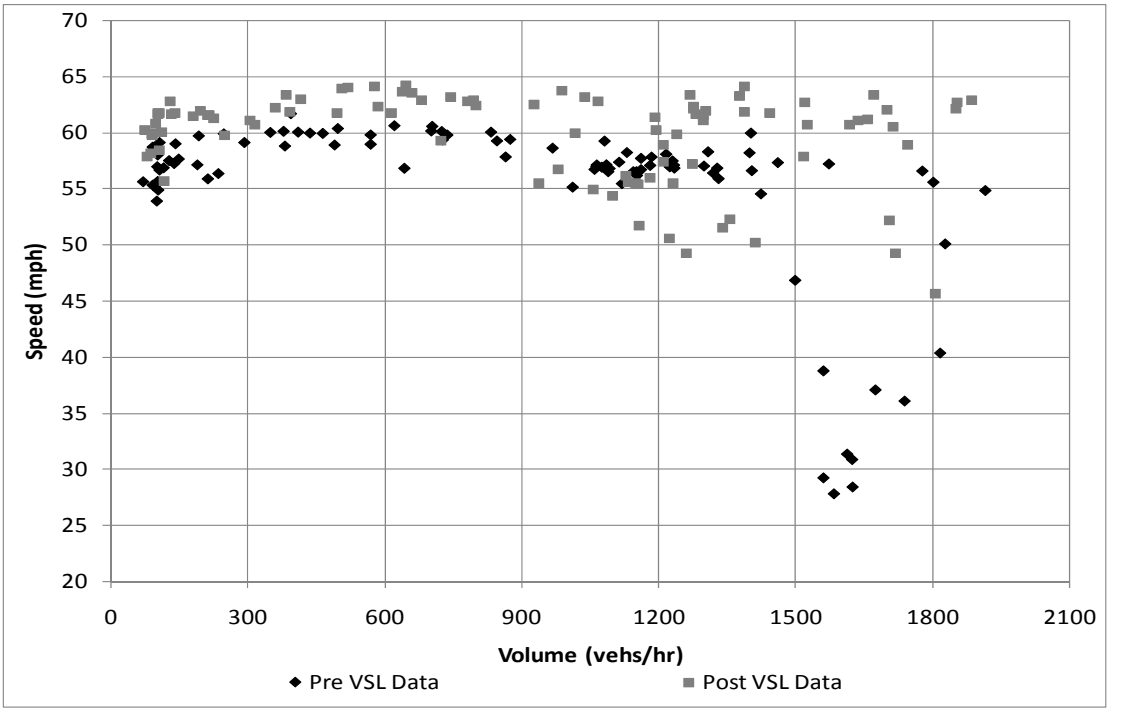



**Figure 7. Speed Flow at 11D Pre- and Post-VSL at mile-marker 11**

Figure 8 presents the calculated travel time along the segment. For calculation of travel time aggregated archived data over 5-minute intervals were used. Travel time was calculated as time taken to travel through the segment at the average speed. Delay was calculated as the difference between the travel times in pre-and post-VSL conditions and percentage of congested travel was calculated as the percent of ratios of delays over their respective travel times. Figure indicates significant decrease in travel time in post VSL conditions. The difference in travel time is calculated as delay and is presented in Table 3.

**Table 3. Travel Time, Delay and Percentage of Congested Travel during peak periods**

| Detector ID | MP 16.6 to MP 15.4 | MP 15.4 to MP 12.4 | MP 12.4 to MP 11 | MP 11 to MP 8.5 |
|---|---|---|---|---|
| Pre VSL Travel Time (minutes) | 2.00 | 6.20 | 2.20 | 4.85 |
| Post VSL Travel Time (minutes) | 1.29 | 4.03 | 1.51 | 3.26 |
| Delay (minutes) | 2.36 | 6.55 | 2.52 | 5.18 |
| Percent Change in Congested Travel | -35.4% | -35.0% | -31.6% | -32.9% |

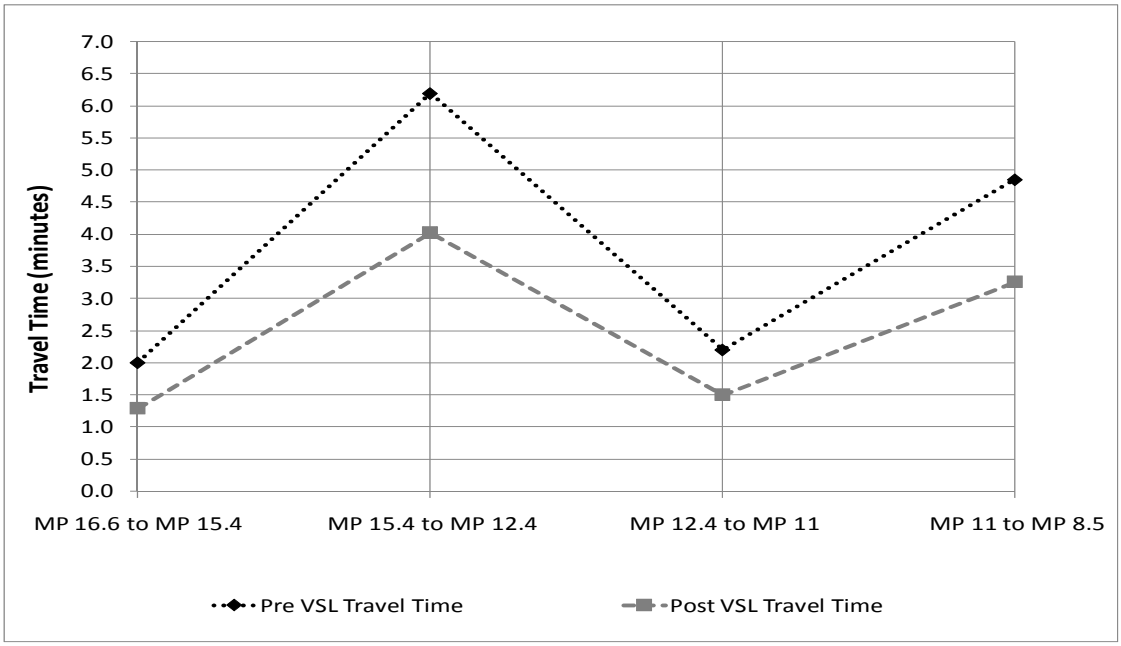
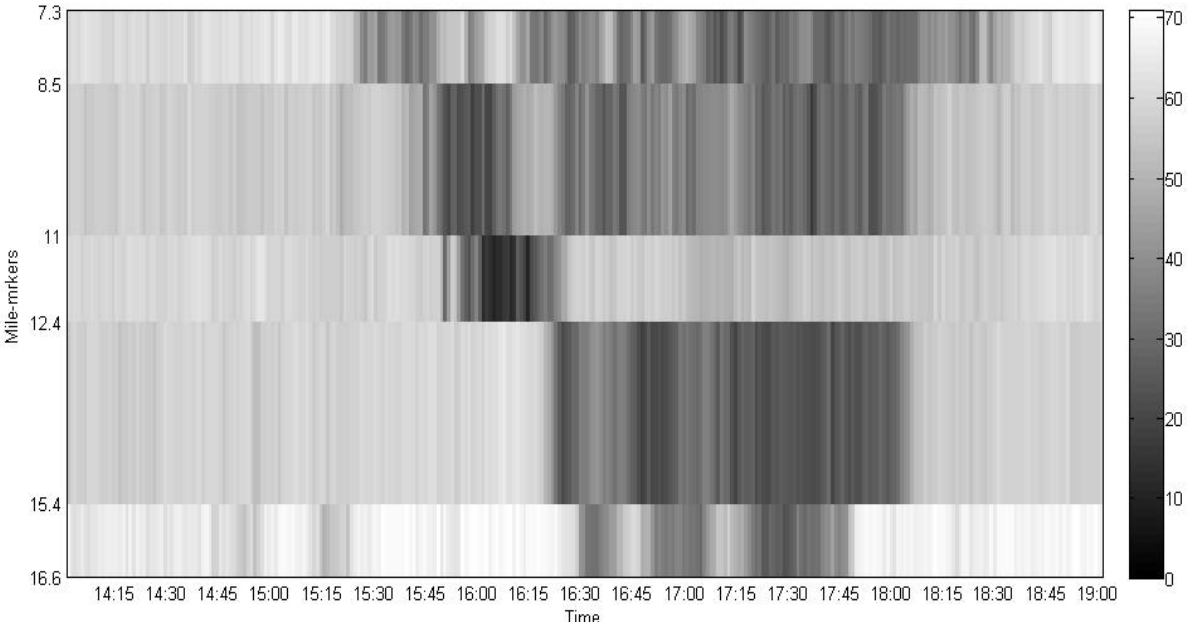



**Figure 8. Comparison of Travel Time for Pre (25th Oct, 07) and Post-VSL (23rd Oct, 08)**

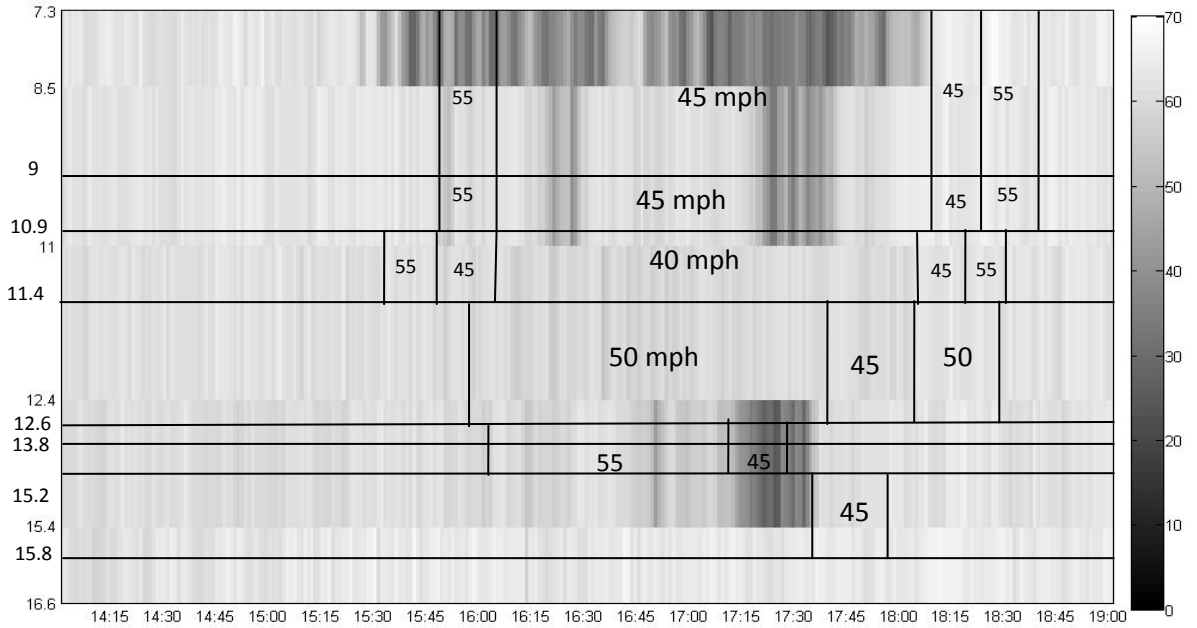
From Figure 9 the duration, extent and intensity of congestion can be quantified. During post-VSL, the duration of congestion reduced appreciably upstream of detector at mile-marker (logmile) 8.5. For detectors between mile-marker (logmile) 11 and 8.5, the duration of congested state of traffic was reduced approximately from 110 minutes (pre-VSL, 16:00-17:50) to 25 minutes (post–VSL, 17:30-17:55). Similarly, for mile-marker (logmile)s from 8.5 to 7.3, the duration reduced by 45 minutes (pre-VSL, 15:40-18:40) and (post-VSL, 15:45-18:00) which indicates improvement in the state of traffic.

This step takes into account the entire highway segment of 9.3 miles from detectors at milepost 16.6D to 7.3D. The time-space plots, presented in Figure 9 compared two different dates for pre- and post-VSL conditions. Comparing the pre- and post-VSL data plots in Figure 9, it can be stated that the VSL system initiation has reduced the duration, extent and intensity of queue and congestion noticeably. In terms of the queue extent (length), it decreased from 9.3 miles (mile-marker (logmile) 16.6 to 7.3) for pre-VSL to 3.3 miles (mile-marker (logmile) 11 to 7.3) for post-VSL conditions. For pre-VSL conditions, Figure 9 (a) clearly indicates the propagation of congestion upstream as average speeds are lower, starting from detector at mile-marker (logmile) 7.3 to 16.6 over time. The pre-VSL conditions indicate that congestion started around 15:30 between mile-marker (logmile)s 8.5 to 7.3 and propagated upstream to mile-marker (logmile) 11 by 15:50, and further to mile-marker (logmile) 12.4 by 16:05. This trend continued upstream and reached mile-marker (logmile) 15.4 around 16:20 and further to mile-marker (logmile) 16.6 by 16:30. In contrast, for post-VSL conditions it was found that the extent of congestion was considerably shorter. Lower speeds were observed at 15:45 between mile-marker (logmile)s 8.5 and 7.3 but congestion did not propagate upstream till 16:15, and the recovery was short. Lower speeds were observed second time (still higher compared to pre-VSL conditions) between 17:30 to 17:55. No congestion was observed between mile-marker (logmile)s 11 and 12.4, however between mile-marker (logmile)s 15.4 and 12.4, congestion was observed for 25 minutes. No congestion was observed upstream of mile-marker (logmile) 15.4. Comparing the two plots significant reduction in congestion was observed for post-VSL conditions.

In terms of intensity of congestion, the pre-VSL condition also indicated lower values of speed compared to post-VSL conditions. The post-VSL plot show fewer yellow and blue patches. Additionally, no dark blue patches can be observed. Hence, congestion intensity reduced considerably in post-VSL conditions.




**Pre-VSL data (25th Oct 07)**




**Post-VSL data (23rd Oct 08)**

**Figure 9. Time space plot**

## CONCLUSIONS AND RECOMMENDATIONS

From the results presented in this paper, it can be concluded that VSL system was effective for the segment evaluated. Similar tasks can be carried out for the evaluation of any existing Variable Speed Limit System. It is also suggested that the evaluation framework used in this study is an effective tool for optimizing the algorithm structure and parameter values.

## REFERENCES

1. Papageorgiou, M., E. Kosmatopoulos, M. Protopappas, and I. Papamichail. (2006) Effects of Variable Speed Limits (VSL) on Motorway Traffic. Internal Report 2006-25. Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece.
2. Steel, P., McGregor, R.V., Guebert, A.A., and McGuire, T.M. (2005). Application of Variable Speed Limits along the Trans Canada Highway in Banff National Park. Annual Conference of the Transportation Association of Canada Calgary, Alberta.
3. Riffkin M., McMurtry. T., Heath.S., and Saito. M., (2008) Variable Speed Limit Signs Effects on Speed and Speed Variation in Work Zones. Prepared For: Utah Department of Transportation Research and Development Division.
4. Mott MacDonald (2008): M42 ATM Monitoring, 4-Lane Variable Mandatory Speed Limits 12 Month Report (Primary and Secondary Indicators) November 2006 and January 2007.
5. Bertini, R. L., S. Boice, and K. Bogenberger (2006) Dynamics of Variable Speed Limit System Surrounding Bottleneck on German Autobahn Transportation Research Record: Journal of the Transportation Research Board, No. 1978, Transportation Research Board of the National Academies, Washington, D.C., pp. 149–159.
6. NCHRP Report 618, Cost-Effective Performance Measures for Travel Time Delay, Variation and Reliability, National Cooperation Highway Research Program, Transportation Research Board, Washington, DC. 2008.

# Use of Micro Unmanned Aerial Vehicles in Roadside Condition Surveys

W. Scott Hart[1] and Nasir G. Gharaibeh[2]

Submitted for Publication in the Proceedings of the 1st Transportation and Development Institute (T&DI) Congress, Chicago, IL, March 13-16, 2011

**Abstract**

Micro unmanned aerial vehicles (MUAVs) that are equipped with digital imaging systems and Global Positioning System (GPS) provide a potential opportunity for improving the effectiveness and safety of roadside condition and inventory surveys. This paper provides an assessment of the effectiveness of MUAVs as a tool for collecting condition and inventory data for roadside infrastructure assets using a field experiment. The field experiment entails performing a level of service (LOS) condition assessment on 10 roadway sample units on IH-20 in Tyler, Texas. The condition of these sample units was assessed twice: onsite (i.e., ground truth) and by observing digital images (still and video) collected via a MUAV. The results of these surveys are analyzed to determine if there are statistically significant differences in the standard deviation and mean values of the condition ratings. Additionally, the operational performance of the MUAV was observed in various weather and field conditions. The results of this study will help transportation agencies to decide if MUAV technology can be adopted for inventory and condition surveys of roadside assets and maintenance activities.

[1] Graduate Research Assistant Student, Department of Civil Engineering, Texas A&M University, College Station, TX 77843-3136, e-mail: scottferd1986@neo.tamu.edu

[2] (Corresponding Author) Assistant Professor, Department of Civil Engineering, Texas A&M University, College Station, TX 77843-3136, e-mail: ngharaibeh@civil.tamu.edu

## Introduction

The majority of state Departments of Transportation (DOTs) collect inventory and condition data for asset management and maintenance quality assurance (MQA) purposes (Pantelias et al. 2009). These programs require periodic and systematic inventory and condition surveys of all roadway assets (culverts, ditches, signs, pavement, guardrails, vegetation, etc.) to determine the roadway's level of service (LOS), set maintenance priorities, and make tradeoff decisions.

Current methods for assessing roadside condition involve manual field inspections that tend to be time consuming, labor intensive, and lack adequate visual digital recording of condition information. Digital inspection can potentially address these shortcomings. While vehicle-mounted digital imaging systems are commonly used for pavement condition and roadway inventory surveys, roadside condition surveys are typically performed manually due to limited accessibility. Micro unmanned aerial vehicles (MUAVs) (see Figure 1) that are equipped with digital imaging systems (still and video cameras) and Global Positioning System (GPS) provide a potential opportunity for improving the effectiveness and safety of roadside condition and inventory surveys. These MUAV systems are commercially available and have been used in areas such as crime scene investigation, cinematography, building inspection, and wind turbine inspection.

The contribution of this paper lies in evaluating the feasibility of using MUAV systems for roadside condition assessment using a field experiment. In this field experiment, an MUAV is used to digitally record roadside sections on Interstate Highway 20 (IH-20) in Tyler, Texas for later assessment of roadside condition and inventory. The results of this study will help transportation agencies to decide if MUAV technology can be adopted for roadside condition and inventory surveys. Ultimately, it is hoped that with the adaption of MUAV technology, the data collection process for roadside assets will become safer, more accurate, and more cost-effective, and will provide visual digital records of these assets.

To provide the reader with a background on this subject, we begin with a review of relevant literature, followed by an overview of the roadside condition assessment method used in this study. The remainder of the paper discusses the field experiment and its results.

## Background

### *Roadway Condition Assessment Methods for Maintenance Quality Assurance*

Most existing MQA programs use a form of the infrastructure condition assessment method that was originally developed in 1985 by Florida DOT and then refined under NCHRP Project 14-12 (Highway Maintenance Quality Assurance) (Stivers et al. 1999). This method allows maintenance contractors and agencies to periodically measure how well maintenance forces are achieving certain performance standards and LOS targets. It also allows for benchmarking of current LOS and for measuring increase or decline in LOS over time.

Automated inventory and condition surveys of pavement assets have come a long way. Vehicle-mounted sensors (digital imaging systems, laser, acoustic, etc.) are able to capture accurate

pavement surface condition that can then be analyzed via computer software, which then output a quantitative value for the condition of that pavement asset. While advances have been made in developing these technologies for pavement, roadside assets are not as accessible and therefore currently require manual inspection methods. A recent survey of 48 transportation agencies from 40 different states in the U.S. showed that 34 agencies use manual methods for collecting roadside and drainage condition data (Pantelias et al. 2009). The same survey showed that only three agencies use manual methods for collecting pavement condition data. Manual methods for conducting roadside condition and inventory surveys involve certain safety issues, ranging from traffic crashes to natural hazards such as washouts, sharp changes in elevation, or hidden objects. Additionally, these manual inspection methods lack an accurate record of the roadside's true condition. Inadequate data records make it virtually impossible to re-evaluate previously inspected roadside sections without having to travel back to the same site.

MUAVs outfitted with digital imaging systems and GPS technology can capture digital videos and still-frame images of roadside assets. These digital images can later be analyzed in a safe, non-stressful work environment and forever stored for later visualization. The purpose of this study is to evaluate the feasibility of using MUAV systems for roadside condition assessment.

### *Overview of Unmanned Aerial Vehicle Systems*

Unmanned Aerial Vehicles (UAVs) were first designed to act as decoys to distract opposing military forces from what was occurring on the ground. Later, UAVs were modified to perform surveillance missions. After the Vietnam War, military science agencies set out to find a more "soldier safe" method for reconnaissance (Levinson 2010). This led to the development of UAVs that could be flown unmanned, but have the functionality of a manned aircraft. Remote sensing combined with computer and GPS technologies led to the development of the present day omniscient UAV. Current military UAVs are fully autonomous and can perform multiple tasks, such as seek and destroy, pre-determined flight, and supply and reinforcement (Taylor 2004).

To improve mobility of UAVs, smaller UAVs that can be carried and operated by a single person were developed and are currently known as micro unmanned aerial vehicle, or MUAV. This has opened many doors for civilian applications to take advantage of this state-of-the-art technology. MUAVs are currently being used in civilian applications such as firefighting, search and rescue, law enforcement, monitoring of oil and gas pipelines, monitoring of rivers and canals, and private surveillance. Limited research efforts have begun to explore the feasibility of using UAV systems in infrastructure management such as bridge condition inspection (Menti and Hamel 2007), pavement condition inspection (Herold et al. 2004, Zhang 2008a, Zhang 2008b, Jengo et al. 2005), and collection of roadway traffic data (Coifman et al. 2006 and Srinivasan and Latchman 2004). Rathinam et al. (2008) developed a detection algorithm that enables UAVs to identify and localize linear infrastructures such as canals, roads, and pipelines.

Generally, there are two major types of MUAVs: plane-configured and helicopter configured. Examples of these MUAV types are shown in Figure 1.

**Fig. 1.** Helicopter-configured (Left) and Plane-configured (Right) MUAVs

A plane-configured MUAV mimics a single-propeller driven aircraft. These MUAVs have the ability to fly in a straight-line path and must be designed to obey the same laws of aero-dynamics that apply to regular aircrafts. The wingspan on this type of MUAV can vary from 12 inches up to four feet, depending on application.

A helicopter-configured MUAV utilizes upward thrust induced by a single or multiple propellers to maneuver in flight. The typical size for helicopter-configured MUAV is approximately 2-3 ft diametrically. However, recent research have used nanotechnology to produce an insect-sized helicopter-configured MUAV (Newcome 2004).

Advantages of plane-configured over helicopter-configured MUAVs include:

1. greater speed
2. ability to carry larger payloads, and
3. ability to glide while in flight (which reduces fuel or battery consumption).

Advantages of helicopter-configured over plane-configured MUAVs include:

1. greater maneuverability (which allows for making immediate and sharp changes in flight direction),
2. ability to loiter in place (which, when coupled with GPS, allows for programming the MUAV to hover at predetermined coordinates)
3. smaller size, and
4. ability to takeoff from a standing position

## Field Experiment

The field experiment entailed performing a level of service (LOS) condition assessment on 10 roadway sample units on IH-20 in Tyler, Texas. Each sample unit is 0.1-mile long. The condition of each sample units was assessed twice:

a. Onsite (i.e., ground truth): Three inspectors rated the roadside assets and maintenance activities within each sample unit directly in the field, and
b. MUAV video: A fourth inspector rated the same sample units by observing digital images (still and video) collected via the MUAV.

The LOS condition assessment method (discussed in the following section of this paper) is used in both surveys (manual and digital). The results of these surveys are analyzed to determine if there are statistically-significant differences in the standard deviation and mean values of the condition ratings obtained through manual inspection (performed directly in the field) and digital inspection (performed in the office using digital videos obtained from the MUAV).

***Roadway Condition Assessment Method***

A roadside condition assessment method that is being developed for the Texas Department of Transportation (TxDOT) is used in the field experiment of this study. In this method, condition assessment field surveys are based on random sampling. Random sampling is used to ensure realistic and affordable data collection requirements. Under a random sampling scheme, once the rating zones (e.g., a 10-mile highway segment) are established, sample units of equal length (typically 0.1-mi long) are chosen from within these zones using random sampling techniques. The sample size [i.e., number of sample units needed to be surveyed to achieve a desired confidence level (e.g., 95%)] is computed as follows (de la Garza et al. 2008):

$$n = \frac{s^2 N}{s^2 + \frac{(N-1)e^2}{Z^2}} \quad (1)$$

where n = sample size, N = total number of sample units in the rating zone; e = tolerable error allowed by the specifications (i.e., allowable difference between sample mean and true mean), Z = standard normal distribution value for the required confidence level that the error doesn't exceed *e*, and s = standard deviation of rating from past experience (e.g., previous pilot projects).

This condition assessment method consists of the following steps:

1) The rating zone (a stretch of highway, in this study) is divided into *N* sample units (typically 0.1-mi long)

2) *n* sample units are selected randomly for field survey (n is computed using Eq. 1).

3) The randomly-selected sample units are inspected and rated on a "Pass/Fail/Not Applicable" basis using the inspection form shown in Figure 3-1. The form includes a total of 55 performance standards for 12 roadside elements (i.e., asset types and maintenance activities).

4) A 0-100 sample score (SUS) is computed as a weighted average score for all elements within the sample unit, as follows:

$$SUS = \frac{\sum_{i=1}^{k} \frac{PS_i}{AS_i} \times PM_i}{\sum_{i=1}^{k} 100 \times PM_i} \quad (2)$$

where *PS* is the number of passing performance standards; *AS* is the number of applicable performance standards; *PM* is an agency-specified priority multiplier (or

weight) for each roadside element, and *k* is the total number of roadside elements within the sample unit.

5) A roadside average LOS for the for the rating zone is computed, as follows

$$LOS = \overline{SUS} = \frac{\sum_{j=1}^{n} SUS_j}{n} \quad (3)$$

where $SUS_j$ is the sample score for sample unit *j* and *n* is the total number of inspected sample units (i.e., sample size).

As mentioned earlier, this condition assessment method was applied on 10 randomly-selected roadway sample units (each is 0.1-mi long) on IH-20 in Tyler, Texas (approximately, 100 miles east of Dallas). The site begins at Mile Marker 556+0.0 and extends for 10 miles in generally rural areas.

### *MUAV used in Field Experiment*

The Dragan Fly X6 helicopter-configured MUAV (see Figure 3) was used in the field experiment. The selection of this particulate MUAV model was based on the following criteria:

1. Loiter capabilities
2. Ability to takeoff/land in confined spaces
3. Carry state-of-the-art imaging devices
4. Equipped with GPS capabilities
5. Onboard and satellite media storage devices
6. Able to maintain continuous flight for at least 15 minutes
7. Reasonably priced compared to other commercial MUAVs
8. Easily piloted
9. Compact, simple, and durable

Table 1 shows the specifications of this MUAV.

| Inspector's Name: | | | | Inspection Date: | Time: | |
|---|---|---|---|---|---|---|
| District: | | Highway: | Milepoint: | Sample Unit No.: | Urban/Rural: | |

| Roadside Asset Type/Mainten | No. | Performance Standard | Grade (Pass, Fail, NA) |
|---|---|---|---|
| Mowing and Roadside Grass | 1 | Any use of herbicide requires advance approval of the Engineer. | |
| | 2 | Paved areas (shoulders, medians, islands, slope, and edge of pavement) shall be free of grass | |
| | 3 | Unpaved areas (shoulders, slopes, and ditch lines) shall be free of bare or weedy areas | |
| | 4 | Roadside vegetation in the mowing area shall be at least 85% free of noxious weeds (undesired vegetation) | |
| | 5 | In rural areas, roadside grass height shall be maintained below 24 inches and shall not be cut to below 7 inches. | |
| | 6 | In urban areas, roadside grass height shall be maintained below 18 inches and shall not be cut to below 7 inches. | |
| Landscaped Areas | 7 | Any use of herbicide requires advance approval of the Engineer. | |
| | 8 | Landscaped areas shall be maintained to be 90 percent free of weeds and dead or dying plants. | |
| | 9 | Grass height in landscaped areas shall be maintained at a maximum height of 12 inches. | |
| Trees, shrubs and Vines | 10 | No trees or other vegetation shall obscure the message of a roadway sign. | |
| | 11 | No leaning trees presenting a hazard shall remain on the roadside. | |
| | 12 | Vertical clearance over sidewalks and bike paths shall be maintained at 10 feet or more. | |
| | 13 | Vertical clearance over roadways and shoulders shall be maintained at 18 feet or more. | |
| | 14 | Clear horizontal distance behind guardrail shall be at least 5 ft for trees | |
| | 15 | No dead trees shall remain on the roadside. | |
| Ditches and Front Slopes | 16 | Ditches and front slopes shall be maintained free of eroded areas, washouts, or sediment buildup that adversely affects water flow. | |
| | 17 | Erosion shall not endanger stability of the front slope, creating an unsafe recovery area. | |
| | 18 | Front slopes shall not have washouts or ruts greater than 3 inches deep and 2 feet wide. | |
| | 19 | No part of the ditch can have sediment or blockage covering more than 10% of the depth and width of the ditch | |
| | 20 | Concrete ditches shall not be separated at the joints, misaligned, or undermined. | |
| | 21 | Front slopes shall not have holes or mounds greater than 6 inches in depth or height. | |
| Culvert and Cross-Drain Pipes | 22 | A minimum of 75% of pipe cross sectional area shall be unobstructed and function as designed. There shall be no evidence of flooding if the pipe is obstructed to any degree | |
| | 23 | Grates shall be of correct type and size, unbroken, and in place. | |
| | 24 | Installations shall not allow pavement or shoulder failures or settlement from water infiltration. | |
| | 25 | Culverts and cross-drain pipes shall not be cracked, have joint failures, or show erosion. | |
| Drain Inlets | 26 | Grates shall be of correct size and unbroken. Manhole lids shall be properly fastened. | |
| | 27 | Installation shall not present a hazard from exposed steel or deformation. | |
| | 28 | Boxes shall show no erosion, settlement, or have sediment accumulation. | |
| | 29 | Outlets shall not be damaged and shall function properly. | |
| | 30 | Inlet opening areas shall be a minimum of 85% unobstructed. | |
| | 31 | Installations shall have no surface damage greater than 0.5 square feet. | |
| Chain Link Fence | 32 | Installations shall have no open gates. | |
| | 33 | Installations shall have no openings in the fence fabric greater than 1.0 square feet. | |
| | 34 | Installations shall have no openings in the fence fabric with a dimension greater than 1.0 feet. | |
| Guard Rails | 35 | Installations shall be free of missing posts, offset blocks, panels or connection hardware. | |
| | 36 | End sections shall not be damaged. | |
| | 37 | Rails shall not be penetrated. | |
| | 38 | Panels shall be lapped correctly. | |
| | 39 | No more than 10% of guard rail blocks in any continuous section shall be twisted. | |
| | 40 | No 25-foot continuous section shall be more than 3 inches above or 1 inch below the specified elevation. | |
| | 41 | No more than 10% of wooden posts or blocks in any continuous section shall be rotten or deteriorated. | |
| Cable Median Barrier | 42 | Installations shall be free of missing or damaged post, cable, or connections | |
| | 43 | Installations shall be free of missing or damaged end sections | |
| | 44 | Installations shall be free of loose cable or cable with incorrect weave | |
| Attenuators | 45 | Each device shall be maintained to function as designed. | |
| | 46 | Installations shall have no visually observable malfunctions (examples – split sand or water containers, compression dent of the device, misalignment, etc.) | |
| | 47 | Installations shall have no missing parts. | |
| Litter and Debris | 48 | 1. No litter or debris that creates a hazard to motorists, bicyclists, or pedestrians is allowed. | |
| | 49 | 2. No 0.1 mile roadway section shall have more than 50 pieces of fist-size or larger litter or debris on either side of the centerline of the highway. | |
| | 50 | Litter volume shall not exceed 3.0 cubic feet per 0.1 mile roadway section on both sides of the pavement. | |
| | 51 | In rural areas, traffic lanes shall be free of dead large animals. | |
| | 52 | In urban areas, traffic lanes and right of way shall be free of dead animals. | |
| Graffiti | 53 | No graffiti is allowed | |
| | 54 | Surfaces and coatings shall not be damaged by graffiti removal. | |
| | 55 | Surfaces from which graffiti has been removed shall be restored to an appearance similar to adjoining surfaces. | |

**Fig. 2.** Roadside Inspection Form used in Field Experiment

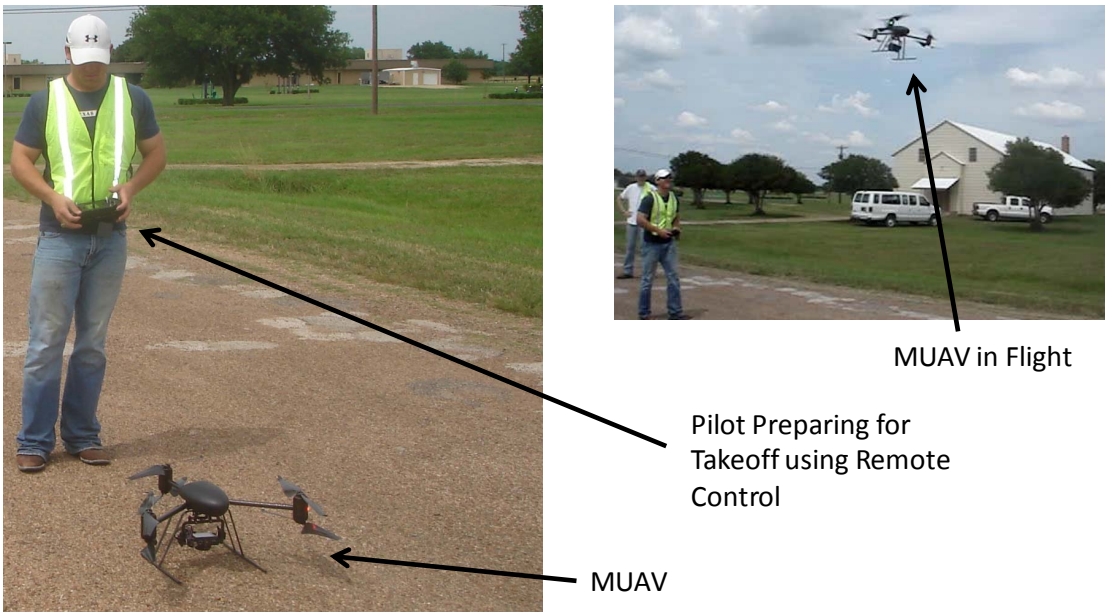



**Fig. 3.** MUAV used in Field Experiment

**Table 1.** Dragan Fly X6 Helicopter Technical Specifications

| Aspect | Characteristic | Value |
|---|---|---|
| Helicopter Size (Fully Assembled) | Width | 36 in |
| | Length | 33 in |
| | Height | 10 in |
| Weight and Payload | Helicopter Weight | 2.2 lbs. |
| | Payload Capacity | 1.1 lbs. |
| | Maximum Gross Takeoff Weight | 3.3 lbs. |
| Flight | Unassisted Visual Reference Required | Path Entered Flight Capabilities |
| | Max Climb Rate | 23 ft/s |
| | Max Descent Rate | 13 ft/s |
| | Max Turn Rate | 90 °/s |
| | Approximate Max Speed | 30 mph |
| | Minimum Speed | None |
| | Launch Type | Vertical Take Off and Landing |
| | Maximum Altitude | 8,000 ft. |
| | Max Flight Time | 25 min. |
| Camera Type | Still Camera | 10 MP Digital Still |
| | Motion Camera | 720p High-Definition |
| | Max Storage | 2 GB |
| GPS | Satellites Used | 16 |
| | Position Update Rate | 4 Hz |
| | GPS Capabilities | Position Hold, Location Data |

### *Discussion of Results*

Figure 4 shows the level of agreement between the performance standards ratings (Pass, Fail, or Not Applicable) obtained by monitoring MUAV videos and corresponding ratings obtained directly in the field by three different inspectors. Considering all performance standards, 72-95 percent of the time, the ratings assigned by the MUAV video rater matched those assigned by the field raters. On average, these ratings matched 81% of the time.

Figure 5 shows an example of false MUAV reading, where an existing drain inlet (identified by the onsite inspectors) was not visible in the MUAV image. Such false readings can increase or decrease the sample unit scores, depending on the condition of the missed asset. This is discussed in the following paragraph.

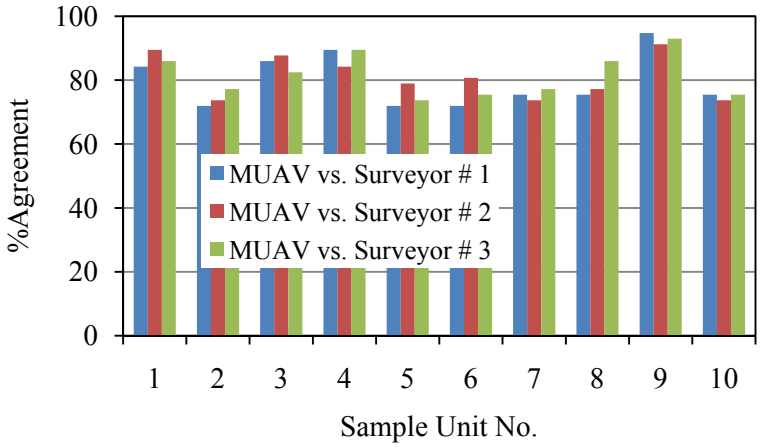



**Fig. 4.** Percent Agreement between MUAV and Onsite Ratings (Pass/Fail/Not Applicable)

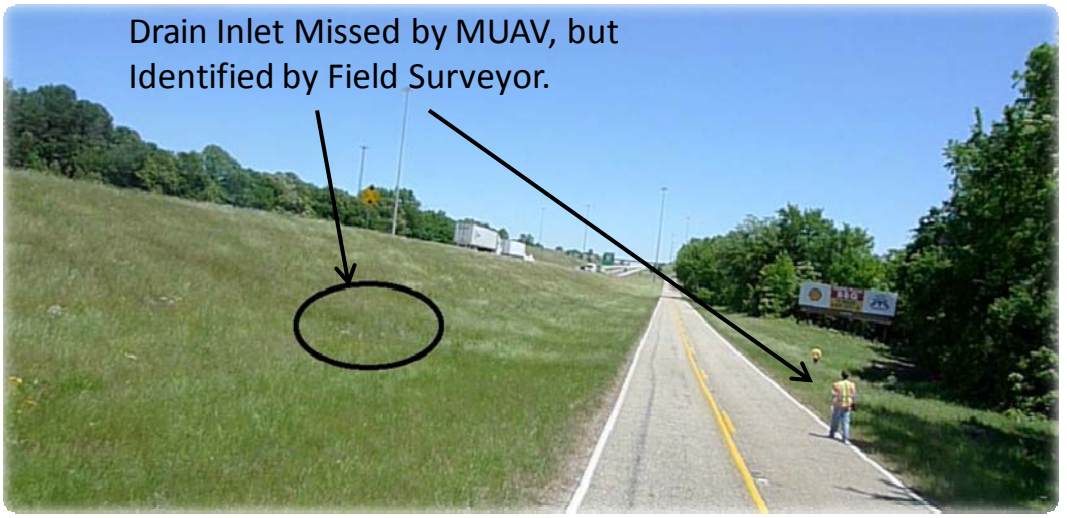



**Fig. 5.** Example False Reading by MUAV (Image Captured by MUAV)

Equation 2 was used to compute a SUS for each sample unit. Figure 6 shows the sample unit scores computed using ratings obtained from the onsite (field) raters and the corresponding scores computed using ratings obtained from the MUAV.

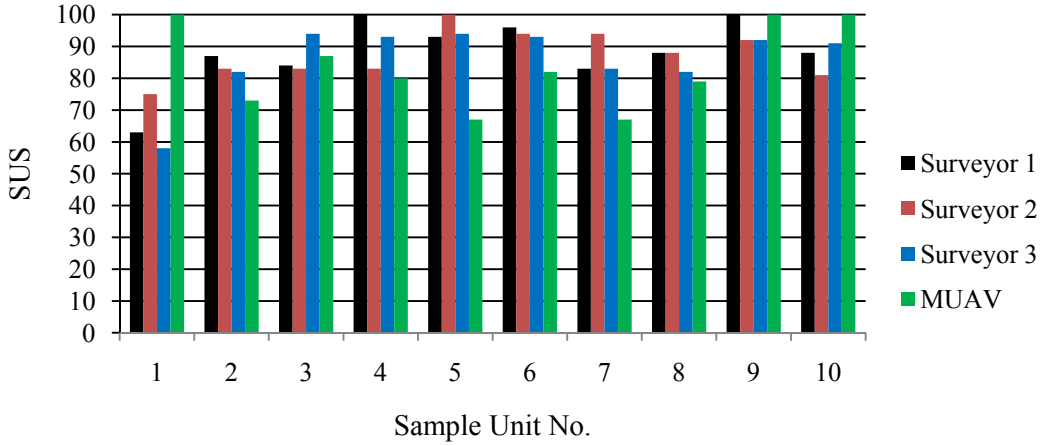



**Fig. 6.** Onsite vs. MUAV-based Condition Scores

The SUSs vary from one sample unit to another. For example, for Sample Unit # 1, the SUS computed based on MUAV data is noticeably higher than that computed based on onsite data (100 vs. 58-73). This is because of false readings taken by the MUAV. The MUAV video showed that this sample unit should pass all ditch and front slope performance standards (which resulted in a SUS of 100). The onsite (ground truth) ratings, however, showed that this sample unit failed to meet the required performance standards for ditch and front slope (which resulted in a SUS of 58-73, depending on the field inspector).

Two statistical tests were conducted on the SUS results. The first was a two-tailed t-test in which the onsite SUS data sets were compared to the corresponding MUAV SUS data set, under the null hypothesis that true mean values are equal. The second statistical test was the F-test which was conducted on the same data sets under the null hypothesis that the variances are equal. Table 2 shows the results of these two statistical tests. The results show that, at a 95% confidence level, there is no statistical evidence that the null hypothesis in either case is false.

**Table 2.** Statistical Results Comparing Onsite vs. MUAV-based Sample Unit Condition Scores (95% Confidence Level)

| Comparison | Sample Size (number of sample units) | T-Test p-value | F-Test p-value | Evidence of Difference in SUSs (Reject Null Hypothesis?) |
|---|---|---|---|---|
| Surveyor # 1 vs. MUAV | 10 | 0.390 | 0.585 | t-Test: No<br>F-Test: No |
| Surveyor # 2 vs. MUAV | 10 | 0.437 | 0.126 | t-Test: No<br>F-Test: No |
| Surveyor # 3 vs. MUAV | 10 | 0.437 | 0.650 | t-Test: No<br>F-Test: No |

**Operational Performance**

The operational performance of the MUAV was observed in the field under three conditions: time of day, wind speed, and flight speed. These observations are summarized in the following

paragraphs.

The MUAV was flown at three different times throughout the day in order to find the optimum window to collect best quality images. This is not truly a test of the MAUV's capabilities, but rather the camera mounted on the MUAV. The specific digital camera that was used in this study was a LUMIX DMC-LX3 manufactured by Panasonic. The camera captured 720p high-definition video images and 10.2 megapixel still images. It was observed that the most optimum time of day to capture images was between 8:00 A.M. until 12 noon. In the afternoon, there is excessive glare off of adjacent pavement surfaces, which reduced the quality of the captured images.

Weather was the most restricting parameter in the entire data collection process. While the MUAV was not flown in rainy weather, wind was found to be the most restricting weather condition. Generally, the MUAV performed well and was easy to control in 0-5 mile per hour winds. In 5-10 mile per hour winds, the MUAV became more difficult to control, but with some training, data could be collected. Wind speed greater than 10 miles per hour interfered in operating the MUAV and resulted in "shaky" video that was difficult to analyze. The MUAV was not operational (could not be controlled) in 15-mile per hour (or more) winds.

Flight speed affects the quality of video and images that the MUAV captures as well as endurance of the MUAV (i.e. maximum flight time). The slower the MUAV travels, the higher the quality of data becomes. However, slower flight speed (i.e., longer flight times per sample unit) reduces the number of sample units surveyed per battery. Approximately, 1.5 minutes of flight time per 0.1 mile sample unit (allowing 4 sample units to be collected per battery), appears to be most practical.

**Summary and Conclusions**

This paper provides an assessment of the effectiveness of MUAVs as a tool for collecting condition and inventory data for roadside infrastructure assets based on a field experiment. The motivation of this study is to improve the safety, accuracy, and time efficiency of roadway condition assessment surveys, and to identify technologies that can provide visual digital records of these surveys. The cost-effectiveness of this approach is not addressed in this paper, since it is likely to change over time as the MUAV technology matures.

The field experiment entails performing a level of service (LOS) condition assessment on 10 roadway sample units on IH-20 in Tyler, Texas. The condition of these sample units was assessed twice: onsite (i.e., ground truth) and by observing digital images (still and video) collected via the MUAV. Statistical analyses of the field data showed that there are no statistically significant differences in the standard deviation and mean values of onsite (ground truth) condition scores and corresponding scores obtained from observing MUAV videos.

Weather was the most restricting parameter in the data collection process. While the MUAV was not flown in rainy weather, wind was found to be the most restricting weather condition. The MUAV was easy to control and produced the highest quality images in 0-5 mile per hour winds. The MUAV was not operational (could not be controlled) in 15 mile per hour (or more) winds.

Overall, the field experiment described in this paper shows that MUAV is a promising technology for improving current data collection methods for roadway inventory and condition assessment. However, false readings and limitations on the operational performance of MUAVs show that there is still a need for improving this technology before it can be adopted in the field.

**References**

Coifman, B., McCord, M., Mishalani, R.G., Iswalt, M., and Ji, Y. (2006). "Traffic Flow Data Extracted from Imagery Collected Using a Micro Unmanned Aerial Vehicle." *Conference Proceedings*, *Applications of Advanced Technology in Transportation, 9th International Conference, ASCE,* pp. 298-303.

de la Garza, J.M. Pinero, J.C., and Ozbek, M.E. (2008) "Sampling Procedure for Performance-Based Road Maintenance Evaluations." *Journal of the Transportation Research Board*. No. 2044, pp. 11-18.

Herold, M., Roberts, D.A., Smadi, O., and Noronha, V. (2004). "Road condition mapping using hyperspectral remote sensing." *Proceedings of the 2004 AVIRIS Workshop*, March 31 - April 2, Pasadena, CA.

Levinson, Charles (2010). "Israeli Robots Remake Battlefield". *The Wall Street Journal*.. Retrieved January 13, 2010. pp. A10.

Metni, N. and Hamel, T. (2007). "A UAV for bridge inspection: Visual servoing control law with orientation limits." *Automation in Construction*, Volume 17, Issue 1, pp. 3-10.

Newcome, Laurence R. (2004). "Unmanned aviation: a brief history of unmanned aerial vehicles". *Library of Flight* Series. ISBN 1563476444.

Pantelias, A., Flintsch, G.W., Bryant, J.W., Jr, and Chen, C. (2009). "Asset Management Data Practices for Supporting Project Selection Decisions." Public Works Management and Policy, Vol. 13, No. 3, pp. 239-252

Rathinam, S., Kim, Z.W., Sengupta, R. (2008). "Vision-Based Monitoring of Locally Linear Structures Using an Unmanned Aerial Vehicle." *Journal of Infrastructure Systems*, Vol. 14, No. 1, pp. 52-63.

Srinivasan, S., and Latchman, H. (2004). "Airborne traffic surveillance systems—Video surveillance of highway traffic." *Proc., ACM 2nd Int. Workshop on Video Surveillance and Sensor Networks*, pp. 131–135.

Stivers, M. L., Smith, K. L. Hoerner, T. E. and Romine A. R. (1999). *Maintenance QA Program Implementation Manual*. NCHRP Report 422, TRB, National Research Council, Washington, D.C.

Taylor, A. J. P. (2004). *Jane's Book of Remotely Piloted Vehicles ,* 1977,Revised 2004

Zhang, C. (2008a). "An UAV-based photogrammetric mapping system for road condition assessment." *Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, ISPRS Congress*, Beijing, China, XXXVII. Part B5, pp. 627-631.

Zhang, C. (2008b). "Development of a UAV-Based Remote Sensing System for Unpaved · Road Condition Assessment." *Proceedings of the American Society for Photogrammetry & Remote Sensing (ASPRS) 2008 Annual Conference,* Portland, Oregon, April 28–May 2.

Jengo, C.M., Hughes, D., LaVeigne, J.D., Curtis, I. (2005). "Pothole Detection and Road Condition Assessment using Hyperspectral Imagery." *Proceedings of the American Society for Photogrammetry & Remote Sensing (ASPRS) 2005 Annual Conference,* Baltimore, Maryland, March 7-11.

# THE CALCULATION OF THE REMAINING SERVICE LIFE BASED ON THE PAVEMENT DISTRESS DATA

By

G. Y. Baladi, T. A. Dawson, C. M. Dean, S. W. Haider, and K. Chatti[1]

Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824, Baladi@egr.msu.edu

## ABSRACT

The distress points, which are the numerical values assigned to pavement distress are not universal or standardized. Most agencies assign these points based on past experience with the intention to periodically calibrate them. The distress points significantly affect the distress indices, the life cycle cost, the remaining service life (RSL), and the accuracy of the pavement decisions.

Distress data were obtained from four State Departments of Transportation. The data were analyzed and the RSL was calculated and equalized to the dollar values of the network and its rate of depreciation. Results of the analyses are presented in this paper. It is shown that the RSL could be calculated without the need to assign distress points. It is also shown that the weighted average RSL of the network reflects the dollar value of the network and its rate of depreciation. The RSL could be used to support accurate pavement decisions.

## INTRODUCTION

Over the last 30 years, many important tools were developed for transportation asset management or pavement management. These include distress points, distress index (DI), remaining service life (RSL), life cycle cost (LCC), and so forth. The DI is typically based on a scale whose limiting values vary from one highway agency to another. Some agencies use a scale from zero to one hundred where one end of the scale represents an excellent pavement condition; a new pavement or any pavement with no surface distress. Others use a scale from zero to ten. In addition, along the DI scale, one or more threshold values are established where preventive maintenance or major rehabilitation actions must be taken. In this paper and for discussion purpose only, a DI scale from zero to one hundred is used as a standard scale where the value of 100 represents an excellent pavement condition as shown in Figure 1.

The distress point system (the numerical values assigned to each pavement defect or combination of defects based on their severity and extent) plays a major role in pavement management (Anderson, 1998, Baladi et al, 1991, Kuo et al, 1991, Novak et al, 1991, Baladi et al, 2004 and 2002). The numerical values of the distress points significantly affect the DI, the LCC outputs, the RSL, and the accuracy of the pavement management decisions. Since most state highway agencies manage similar

Distress index = DI = 100 = excellent pavement, DI = 60 = threshold value where the pavement is in need of a major rehabilitation action.

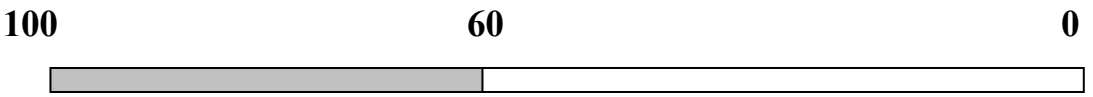



Figure 1a a distress index (DI) scale used for discussion in this paper

The remaining service life (RSL) = DSL = the assumed design service life, which corresponds to a DI of 100 and asset dollar value of $DCC; RSL = 0.0 corresponds to a DI of 60 and certain percent of $DCC; RL = 0.0 corresponds to a DI of 0.0 and asset value of zero ($ZDV).

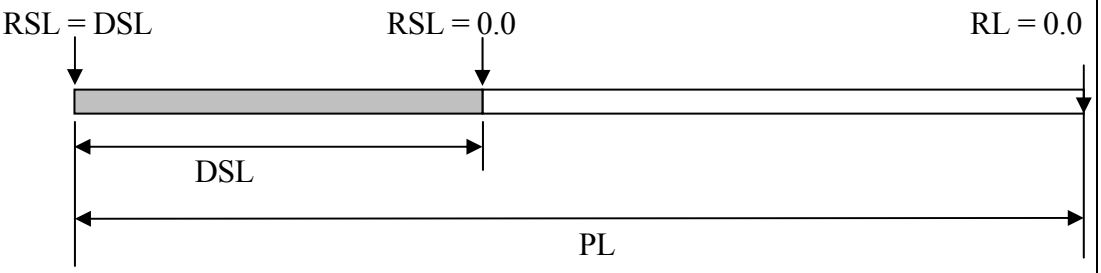



Figure 1b The remaining service life (RSL) and the remaining life (RL)

$DCC = the pavement design and construction costs, $ZDV = 0.0 = the pavement as a structure has zero dollar value.

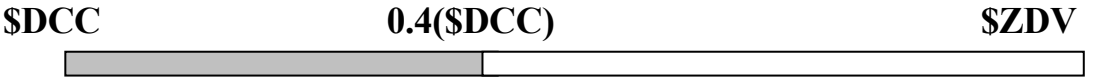



Figure 1c The two universal limiting dollar values of a pavement section

Figure 1 Relationships between the DI, the $DCC, the RSL, the RL, the DSL and the PL

pavement networks, conceptually, their distress point systems for each road class should be similar. In reality, they vary significantly from one highway agency to another. Stated differently, for a given class of road (e.g., interstates, primary, etc) no universal distress point system has been developed or adopted. This paper shows that the pavement distress points could be standardized if their values are based on the true dollar value of the pavement network and its rate of depreciation or deterioration. The paper also shows that since the two generic limiting pavement values are universal in nature, the standardized distress point system should address the two

limits, and should express the pavement condition and its rate of deterioration/depreciation.

Since this paper uses pavement management as illustration for the bigger picture of asset management, two basic, but well known, definitions must be stated or reviewed herein as follows:

a) **Pavement Design Service Life (DSL)** – The DSL expresses the number of years where the pavement structure is designed to provide a level of service equal to or higher than the specified minimum standard level of service. Relative to a given DI, the DSL is the assumed number of years from construction or rehabilitation to the time when the pavement distress or the DI would reach the threshold value as shown in figure 1b. Most pavement design methodologies are mainly based on designing the service life (not the pavement life) of pavements to serve traffic within a range of serviceability level. For example, the AASHTO Design Guide recommends initial and terminal serviceability indices of about 4.2 and 2.5. Whereas the serviceability level in mechanistic-empirical design procedures is based on certain percent of cracking and rut depth.

b) **Pavement Life (PL)** – The PL is the number of years from the original construction to the time when the pavement is in need of reconstruction as shown in figure 1b. The PL may consist of one or more DSL such as the DSL of a newly designed and constructed pavement and the DSL of one or more overlays. The PL of a pavement structure is a function of the pavement preservation practice used by the responsible agency.

Finally, for the benefits of the reader, some of the basic tools that were developed for pavement management are briefly reviewed in the next subsections. These basic tools include the DI, the two generic and universal limiting dollar values of the pavement, the remaining service life, the remaining life, and the distress points.

a) **The Distress Index (DI) -** Most highway agencies have established some type of DI or DIs based on certain algorithms. The various types of indices used today include: International Roughness Index (IRI) or simply roughness index or ride quality index (RQI), structural index (SI), distress index (DI), overall quality index (OQI), pavement condition index (PCI), and so forth. Some indices (such as the roughness index or the rut index) are based on one type of distress whereas others (such as the SI, OQI, and PCI) are based on a combination of distresses. In this paper, a DI scale of 0.0 to 100 is used with a threshold value of 60 indicating pavement sections in need of major rehabilitation or reconstruction. The 60 threshold value expresses the pavement surface conditions at which the pavement section is providing the minimum acceptable service level. The threshold value does not represent a failed pavement.

b) **The Two Generic and Universal Limiting Dollar Values of Pavement -** When a pavement section is designed and constructed properly, its upper dollar value after construction (neglecting the real estate value) is equal to its design and construction

costs ($DCC). On the other hand, if a pavement section is subjected to no maintenance and/or rehabilitation, it will deteriorate to the point where the pavement needs a total reconstruction (from the roadbed soil and up). At that point in time, the pavement has zero dollar value ($ZDV). At any other time, the dollar value of a pavement section depends on its conditions and the level of services it provides to the users. Hence, the two generic and universal limiting dollar values (see figure 1b) of any pavement section are the $DCC and the $ZDV. It should be noted that the $ZDV expresses the value of the pavement as a structure not its salvage value as recycled materials (such as recycled concrete by crushing or recycled asphalt pavements (RAP)). The dollar value of a pavement network is the weighted average values of all uniform pavement sections within the network. Given that any pavement section has two limiting dollar values, and given that the DI scale shown in figure 1a also has two limiting values (100.0 and 0.0), one could relate a DI of 100 to the $DCC and a DI of 0.0 to the $ZDV. Such relationship requires that the distress points assigned to each distress type and its severity and extent are calibrated so that when the pavement structure is failed and its dollar value is equal to the $ZDV, the DI value would be 0.0. The distress point calibration issue is addressed further later in this paper.

c) **The Remaining Service Life and the Remaining Life -** The remaining service life (RSL), not the remaining life, of a given pavement section is the estimated number of years from any given date (usually from the last distress survey date) to the time when the pavement DI reaches a threshold value or the minimum acceptable service level as shown in figure 2. The RSL is typically calculated using the distress data or the DI. To illustrate, assume that, for each 0.1 mile of pavement, the established threshold value of the DI is 60, which corresponds to 500 feet longitudinal crack (LC), and that the equation of the DI and that of the LC as a function of time are known such as DI = f(t) and LC = G(t), the RSL can be calculated using the following equation:

$$RSL = \{t(DI = 60) - SA\} = \{t(LC = 500) - SA\} \leq DSL$$

Where t(DI = 60) is the time at which the value of the DI is 60 (the threshold value); T(LC = 500) is the time when the longitudinal cracks become 500 feet long; SA is the pavement surface age in years, and DSL is the pavement design service life.

The RSL of a given pavement network can be calculated as the weighted average RSL of the "n" pavement sections within the network using the following equation:

$$RSL_{(network)} = \frac{\sum_{i=1}^{n} (RSL_i)(SL_i)}{\sum_{i=1}^{n} SL_i}$$

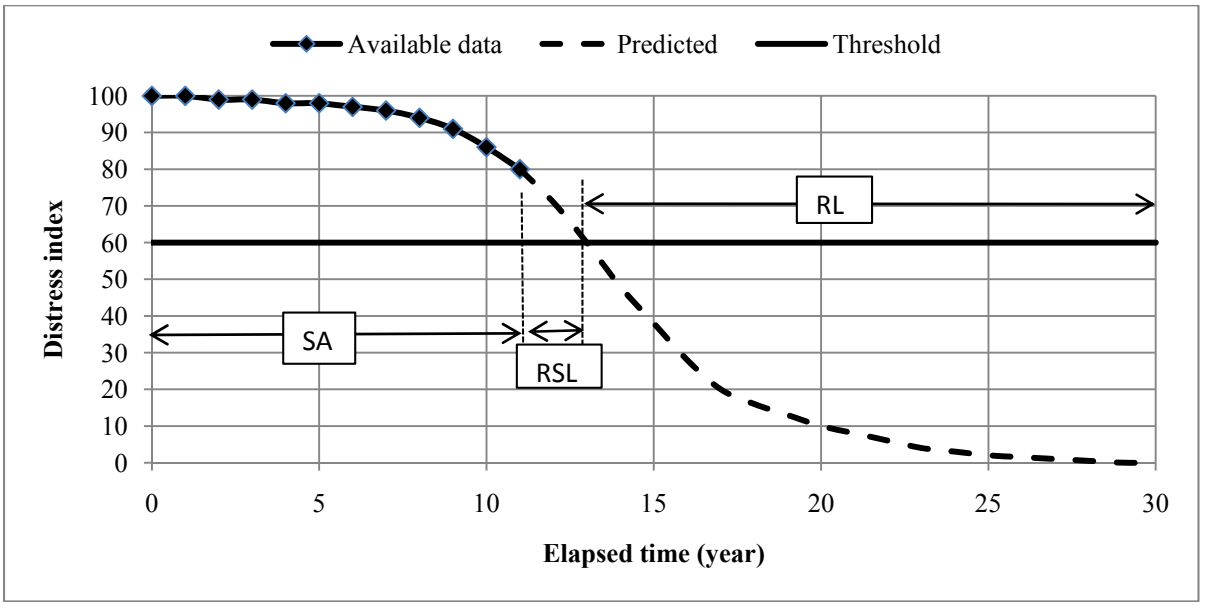



Figure 2 Measured and predicted DI over time showing the remaining service life (RSL), the remaining life (RL) and the pavement surface age (SA)

Where $RSL_i$ is the RSL of pavement section "i", and $SL_i$ is the length of pavement section "i".

Like the pavement values, the RSL also has two universal limiting values, maximum, which is equal to the pavement design service life (DSL) and minimum, which is equal to zero. At any given time, the RSL is somewhere between these two values. When the RSL is the same as the DSL, the asset dollar value is equal to its upper limit, the $DCC. On the other hand, when the RSL is zero, the pavement dollar value is not zero, it is between the $DCC and the $ZDV.

The remaining life of a pavement section is the estimated number of years, from any given date (usually from the last distress survey date) to the time when the pavement section must be reconstructed (a DI of zero or LC = 1250) as shown in figure 2. The calculations of the RL of a pavement section and the pavement network are similar to those of the RSL as follows:

$$RL = \{t(DI = 0.0) - SA \text{ and}\} = \{t(LC = 1250) - SA \text{ and}\} \geq RSL$$

$$RL_{(network)} = \frac{\sum_{i=1}^{n} (RL_i)(SL_i)}{\sum_{i=1}^{n} SL_i} \geq RSL$$

d) **The Distress Points -** From the above scenarios it is clear that the values of the distress points assigned to each type of distress and its severity level and extent have significant impact on the calculation of the DI, the RSL, and the remaining life of any pavement section and of the pavement network (Kuo et al, 1991, Novak et al,

1991, Baladi et al, 2002 and 2004). Stated differently, the assigned values of the distress points are the foundation of asset management and they affect all asset management decisions. Therefore, establishing distress point matrices is a long term exercise that should be periodically calibrated. Otherwise, the maximum acceptable amount of distress (such as 500 and 1250 ft) can be specified and the RSL is calculated on that basis without the need of DI.

## PAVEMENT DOLLAR VALUES

There are several requirements to good pavement management practices; these include a full knowledge of the dollar value of each pavement section and the entire network as well as their rates of depreciation and dollar values. Such knowledge would assist managers to properly estimate the required budget level and the rate of expenditures. The current dollar values of each pavement section and their respective rates of depreciation can be accurately estimated from the time series pavement distress data, from the historical DI values, or from the RSL. The degree of accuracy of the results is a function of the accuracy of the distress data, the distress point systems, the DI or the remaining service life. The relationships between the pavement dollar value and the DI and between the former and RSL are presented below.

### Pavement Dollar Value and the DI

It is assumed herein that the pavement roughness is not included in the calculation of the DI. As stated previously, the relationship between the dollar value of the pavement and its DI at the two DI limits (100 and 0.0) are known. When the DI is equal to 100 or to 0.0, the dollar value of the pavement is respectively equal to the $DCC or to the $ZDV as shown in figure 1b. If the DI threshold value truly represents pavements in need of major structural rehabilitation, the pavement dollar value has already depreciated substantially from its $DCC value. Data from various agencies suggest that when the pavement is in need of major structural rehabilitation action, the average cost of the action is about 60 percent of its $DCC (discounting the inflation rate). This could be interpreted as the pavement dollar value has depreciated by about 60 percent from its $DCC and its value before rehabilitation is about 40 percent of the $DCC. The forty percent will certainly vary from one pavement section to another depending on the distress type and pavement types, location and class. Since, for a pavement section in need of major structural rehabilitation, the DI is equal to about 60, one can associate the 60 percent depreciation with a DI of 60. Such an association is based on the assumption (which is supported by data from many roads under the jurisdiction of many State Highway Agencies) that, the average pavement rate of deterioration after major rehabilitation actions are completed is about 10 percent higher than the average rate of deterioration of a newly designed and constructed pavement. Stated differently, on average, the expected service life of a pavement section subjected to a major structural rehabilitation (such as full-depth patching and overlay) is about 90 percent of that of a newly constructed pavement. To illustrate, consider that a newly constructed pavement section provided 16 years of service by the time the DI reaches a value of 60, the maximum expected service life of that section after a major structural

rehabilitation is taken would be about 90 percent or about 14 years. Certainly, the service life after major structural rehabilitation actions are taken will vary substantially from one pavement section to another depending on the type of rehabilitation action and the pre-rehabilitation repair actions. The above scenario implies that, discounting inflation, the dollar value of the pavement after major rehabilitation is completed is about 90 percent of the $DCC. This would drop to about 80 percent of the $DCC after a second major rehabilitation action is taken and to about 70 percent after the third major rehabilitation as shown in Figure 3.

It is very important to note that if the DI is equal to 60 due to pavement roughness only, then the dollar value of the pavement is much higher than the forty percent stated above. For this reason, in this discussion, pavement roughness is excluded from the DI. Relative to pavement roughness, the ride quality index or the International Roughness Index (IRI) can also be used to estimate the pavement rate of deterioration, the dollar value of the pavement and its rate of depreciation. However, the relationship is much more complex and it is based on the causes of pavement roughness. For example, if the IRI of a newly constructed asphalt or concrete pavement section is high due to construction (no pavement surface distress), the dollar value of the pavement structure is only slightly lower than the $DCC. The structural capacity of the pavement is not affected by the high IRI and the pavement surface can be smoothed at a relatively low cost. On the other hand, if the high IRI values are due to pavement surface distress (e.g., transverse cracks, faulting, etc.), then the structural DI should be used as stated above.

One other point on the DI scale can also be associated with the pavement dollar value, the triggering DI value for preventive maintenance actions. Since preventive maintenance actions are typically taken at an early stage (for example, 2 to 5 percent cracking); the triggering DI value should be high. For the purpose of this discussion, the triggering DI value for preventive maintenance action is taken at 85 (assuming the same DI scale used above). On average, when the DI value is 85, the dollar value of the pavement is about 80 percent of the $DCC. That is the pavement value has depreciated by about 20 percent. It is important to note that a well designed and executed preventive maintenance program would restore the pavement conditions and would slow the pavement rate of deterioration. Hence, preventive maintenance actions would increase the pavement value to about 95 percent of its original $DCC and would increase the remaining service life of the pavement at relatively low costs.

**Pavement Dollar Value and the RSL**

Once again, it is assumed herein that the pavement roughness is not included in the calculation of the RSL. As stated previously, the relationship between the dollar value of the pavement and its RSL at the two RSL limits of DSL and 0.0 are known. When the RSL is equal to the DSL, the dollar value of the pavement section is the $DCC. When the RSL is zero, which correspond to a DI of 60, major structural rehabilitation is required and the pavement dollar value is about 40 percent of its $DCC as shown in figures 1b and 1c. Once again, distress data obtained from various SHAs indicate that the average pavement rate of deterioration after major rehabilitation actions are completed is about 10 percent higher than the average rate of a newly designed and

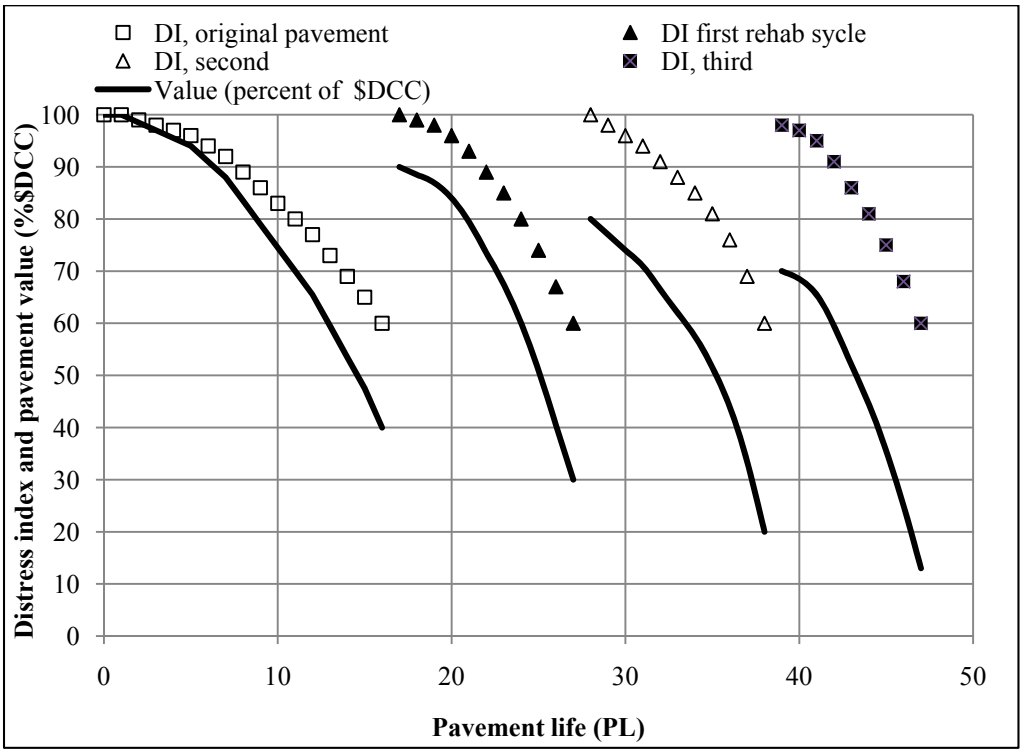



Figure 3 Distress index and the dollar value of a pavement section versus pavement life

constructed pavement. That is the expected service life of a pavement section subjected to the first cycle of a major structural rehabilitation is about 90 percent of that of a newly constructed pavement. For example, if a newly constructed pavement section provided 16 years of service by the time the DI reaches a value of 60, the maximum expected service life of that section after major structural rehabilitation actions are taken would be about 90 percent or about 14 years. Certainly, the service life after major structural rehabilitation actions are taken will vary substantially from one pavement section to another depending on the type of rehabilitation action and the pre-rehabilitation repair actions. The above scenario implies that, discounting inflation, the dollar value of the pavement after major rehabilitation is completed is about 90 percent of the $DCC. This would drop to about 80 percent of the $DCC after a second set of major rehabilitation actions are taken as shown in Figure 4.

**SUMMARY**

The distress data could be used to calculate distress indices and the RSL of pavement sections and the network. The average cost data for new construction and for major structural rehabilitation were analyzed along with the distress indices and the RSL of pavement sections. Results of the analyses indicate that the dollar values of pavement sections are directly related to the distress indices, if the assigned distress points are correct. The dollar values of pavement sections are also directly related to the RSL, which can be calculated using distress data without the need to calculate distress indices.

**ACKNOWLEDGEMENT**

The authors wish to express their sincere thanks for the US Federal Highway Administration, the Arizona Department of Transportation, the Michigan Department of Transportation, the Louisiana Department of Transportation and development, the Washington State Department of Transportation, and to the Colorado Department of Transportation for their support in obtaining the data in support of the analyses. The valuable inputs of various personnel of the Michigan Department of Transportation are deeply appreciated.

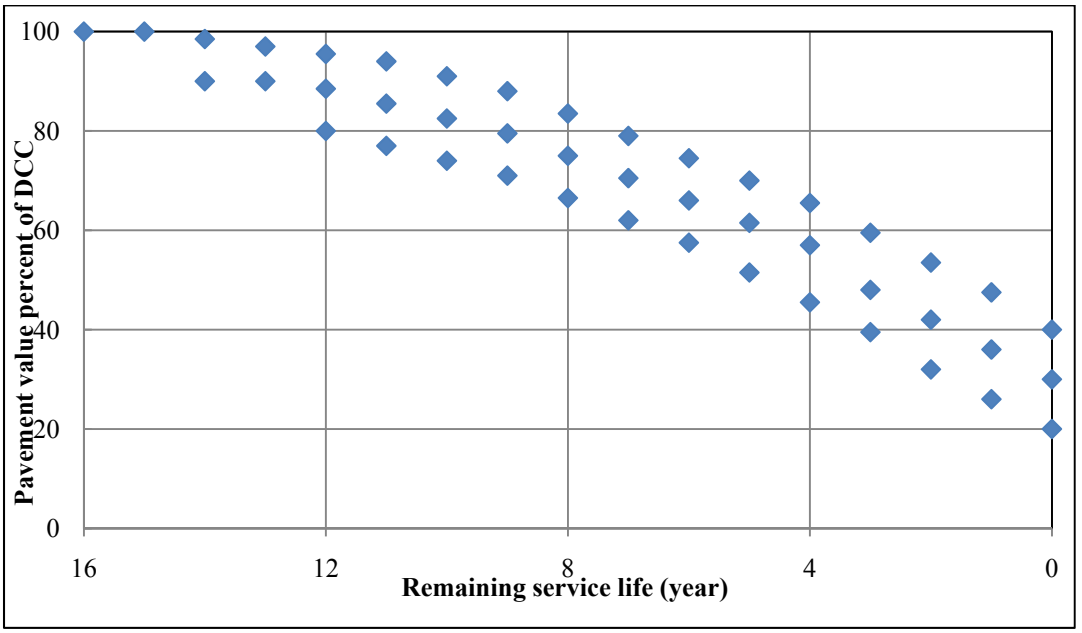



Figure 4 RSL and the dollar value of a pavement section with 2 major rehabilitation actions

## REFERENCES

Anderson J. (1988), *Pavement Management System (PMS) - Distress Survey manual,* Michigan Department of Transportation, Lansing, Michigan

Baladi, G.Y., Novak, E.C., and Kuo, W.H. (1991), *Pavement Condition Index - Remaining Service Life,* the American Society for Testing and Materials, STP 1121

Cooper, K. (1990) *Manual Methods for Using Program Strategy,* Michigan Department of Transportation, Lansing, Michigan, October

Kuo, W.H., E.C. Novak, and G.Y. Baladi (1991), *Inclusion of Pavement Remaining Service Life Concept in PMS Methodology Standardization*, the American Society for Testing and Materials, STP 1121

Novak, E.C., G.Y. Baladi, and W.H. Kuo (1991), *Toward Standardization of a PMS Analysis Method*, the American Society for Testing and Materials, STP 1121

Baladi, G.Y., K. C. Chatti, and N. Buch, "An Excellent Asset Management Tool, The Remaining Service Life", proceedings, the 6th International Conference on Pavement Management, Australia, 2004.

Baladi, G. Y., T. Svasdisant, T. Van, N. Buch, and K.C., Chatti, "Cost-Effective Preventive Maintenance", Transportation Research Record, Journal of the Transportation Research Board No. 1795, pp 17-27, 2002

# THE REMAINING SERVICE LIFE A GOOD PAVEMENT MANAGEMENT TOOL

By

G. Y. Baladi, T. A. Dawson, C. M. Dean, S. W. Haider, and K. Chatti

Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824, Baladi@egr.msu.edu

## ABSTRACT

The remaining service life (RSL) of a given pavement section is the estimated number of years between now and the time when the pavement conditions reach certain amount of distress. The calculation of RSL is based on two steps; fitting the proper mathematical function to the time series condition data, and predicting the time at which the pavement reaches the threshold conditions.

Time series pavement condition data of several miles of roads were obtained from four State Highway Agencies (SHAs). The data were analyzed and the RSL of each 0.1-mile section of each road was calculated based on the International Roughness Index, (IRI), rut depth and cracking data. Results of the analyses are presented in this paper. It is shown that, the RSL is a good communicating and strategy planning tool. It could be used to address engineers, the public, legislators and managers.

Keywords: pavement management system, pavement distress data, remaining service life, pavement management decisions.

## INTRODUCTION

The basic measurements of the conditions of a pavement section are its existing distresses. A pavement distress is defined herein as any conditions that adversely affect the pavement structural capacity, and the pavement serviceability, ride quality, and/or safety. In general, there are two classes (structural and functional) and several types of distress that are associated with each pavement type (ERES Consultants, 1987, Canadian Good Road association, 1961, and Baladi and Snyder 1890). Structural distresses are associated with the ability of the pavement to carry the design load. Functional distresses deal mainly with ride quality (smooth and comfortable ride) and safety issues (skid resistance, rut, bleeding and hydroplaning). Pavements that exhibit structural distresses (e.g., severe alligator cracking, transverse or longitudinal cracks) will also exhibit functional distress. On the other hand, functionally distressed pavements (e.g., very rough) may be structurally sound. In addition, each class of pavement distress (functional or structural) contains several types of distresses. During the pavement distress survey, each distress type is typically identified by its severity and extent. Further, each distress type is caused by one or more variables (e.g.,

moisture, drainage, environment, load) which, when known, provide great insight into the causes of pavement deterioration and the selection of cost-effective pavement preservation alternatives. Hence, the proper pavement evaluation program and procedure should include the identification of the type, severity, and extent of the distresses as well as any abnormal drainage problem (e.g., standing water in the drainage ditch).

Since the structures of the various pavement networks are made of various materials and are located in different environmental regions, there is no standard distress identification manual that have been universally adopted by highway authorities. However, some standardization exists and various but similar distress identification manuals were written and have been published. The most similar manuals are the Highway Pavement Distress Identification Manual (published by the Federal Highway Administration (FHWA)) and the Distress Identification Manual for the Long Term Pavement Performance (LTPP) Studies, which was published by the Strategic Highway Research Program (SHRP), (Smith et al, 1979, 1987, and AASHTO Design Guide, 1986). Both of these manuals address the four conventional highway pavements listed below.

1. Asphalt surfaced (including asphalt overlays of concrete)
2. Jointed plain concrete
3. Jointed reinforced concrete
4. Continuously reinforced concrete

## DISTRESS POINT SYSTEMS

The general state-of-the-practice includes the assignment of distress points for each type of pavement surface distress, severity level and extent. The assigned values of the distress points affect the values of the calculated distress index (DI) or indices and the pavement remaining service life (RSL). Hence, the exercise of establishing distress point values or systems is very critical to the success of pavement management. Such an exercise must be based on the true conditions of the pavement, the upper and lower limits of the DI scale, and the trigger values for major pavement rehabilitation and for pavement preservation and preventive maintenance actions. The assigned distress points must also address the severity level and the extent of each distress. Stated differently, the distress point system must be designed to transfer real distress data to a scalar that can be used to express the health of the network. This scalar is either a single DI or multiple indices. Typical uses of the indices include:

- Establishment of uniform pavement sections based on conditions
- Calculation of RSL of pavement sections
- Selection of candidate project boundaries

The accuracy of the above functions depends on the accuracy of the collected distress data and on the assigned distress points. The accuracy of the data collection activities could be improved by establishing a better quality control protocol and by improved training. The potential problems in the accuracy of the distress point system

could be totally eliminated or neutralized if the above decisions are made using the actual distress data. To illustrate, a typical time series transverse crack data are listed in Table 1 in terms of low, medium and high severity levels. The straight sum of all cracks (low, medium and high) for each year is also included in the table. The transverse crack index was calculated based on the priority factor for low, medium and high severity levels of 0.1, 0.2 and 0.7. Figure1depicts the numbers of low, medium and high severity transverse cracks as a function of time. It can be seen that although the numbers of high and medium severity transverse cracks increases over time, the trend for the number of low severity cracks versus time is not as clear. The reasons for these observations are:

1. The number of cracks in the database that are correctly labeled (by the people who digitize the images) as low, medium and high severity levels is unknown.
2. The effects of pavement temperature during the videotaping process on the crack opening are not known. Low pavement temperatures cause the pavement to shrink and the cracks to open wider, and this artificially increases the number of cracks in the medium and high severity levels. Likewise, high pavement temperature causes the crack opening to tighten and hence, the number of cracks in the medium and high severity levels could decrease substantially from one year to the next.
3. The number of transverse cracks that change severity level from one year to another is unknown. Hence, some low severity cracks may change to high severity, some may change from low to medium severity, and some may stay in the same severity level.

Figure 2 depicts the transverse crack index versus time. The index was calculated based on priority factors of 0.1, 0.2 and 0.7 for low, medium and high severity cracks as shown in Equation 1.

$$TCI = 0.1(LTC) + 0.2(MTC) + 0.7(HTC) \qquad \text{Equation 1}$$

Where TCI = transverse crack index;
LTC = number of low severity transverse cracks;
MTC = number of medium severity transverse cracks; and
HTC = number of high severity transverse cracks.

Figure 3 shows the total number of transverse cracks as a function of time. The total number of transverse cracks (TTC) was calculated by simply adding the numbers of low, medium and high severity transverse cracks as stated in Equation 2.

$$TTC = LTC + MTC + HTC \qquad \text{Equation 2}$$

Examination of the data presented in Figures 2 and 3 indicate that the variability of the total number of cracks versus time is much lower than that of the transverse crack index. Although the data in Figure 3 do not differentiate between the numbers of transverse cracks in the various severity levels, the procedure of adding the numbers of cracks can be justified and it makes more engineering sense. The mechanics of the problem are that the pavement integrity and structural capacity is jeopardized when it

Table 1 Number of transverse cracks, total number of cracks, and transverse crack index

| EMP | Year | Number of transverse cracks | | | | Transverse crack index priority | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Low | Medium | High | Total cracks | Low | Medium | High | Index |
| | | | | | | 0.1 | 0.2 | 0.7 | |
| 1 | 2000 | 10 | 9 | 4 | 23 | 1 | 1.8 | 2.8 | 94.4 |
| 2 | 2001 | 24 | 7 | 1 | 32 | 2.4 | 1.4 | 0.7 | 95.5 |
| 3 | 2002 | 23 | 8 | 1 | 32 | 2.3 | 1.6 | 0.7 | 95.4 |
| 4 | 2003 | 21 | 11 | 8 | 40 | 2.1 | 2.2 | 5.6 | 90.1 |
| 5 | 2004 | 26 | 12 | 10 | 48 | 2.6 | 2.4 | 7 | 88 |
| 6 | 2005 | 36 | 12 | 3 | 51 | 3.6 | 2.4 | 2.1 | 91.9 |
| 7 | 2006 | 22 | 16 | 16 | 54 | 2.2 | 3.2 | 11.2 | 83.4 |
| 8 | 2007 | 37 | 14 | 8 | 59 | 3.7 | 2.8 | 5.6 | 87.9 |
| 9 | 2008 | 26 | 19 | 22 | 67 | 2.6 | 3.8 | 15.4 | 78.2 |
| 10 | 2009 | 22 | 23 | 24 | 69 | 2.2 | 4.6 | 16.8 | 76.4 |

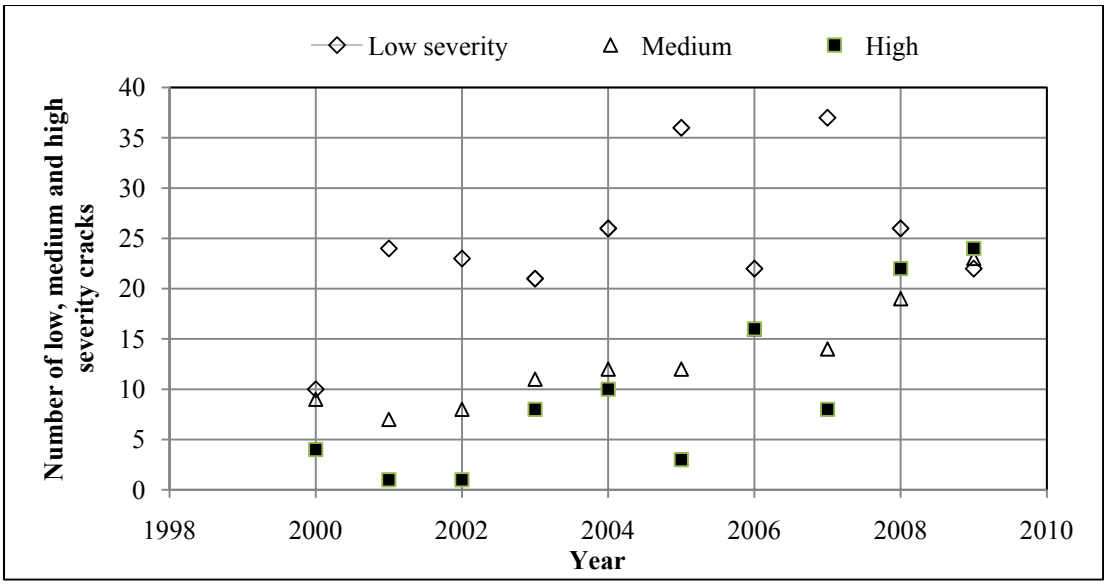



Figure 1 Numbers of low, medium and high severity transverse cracks

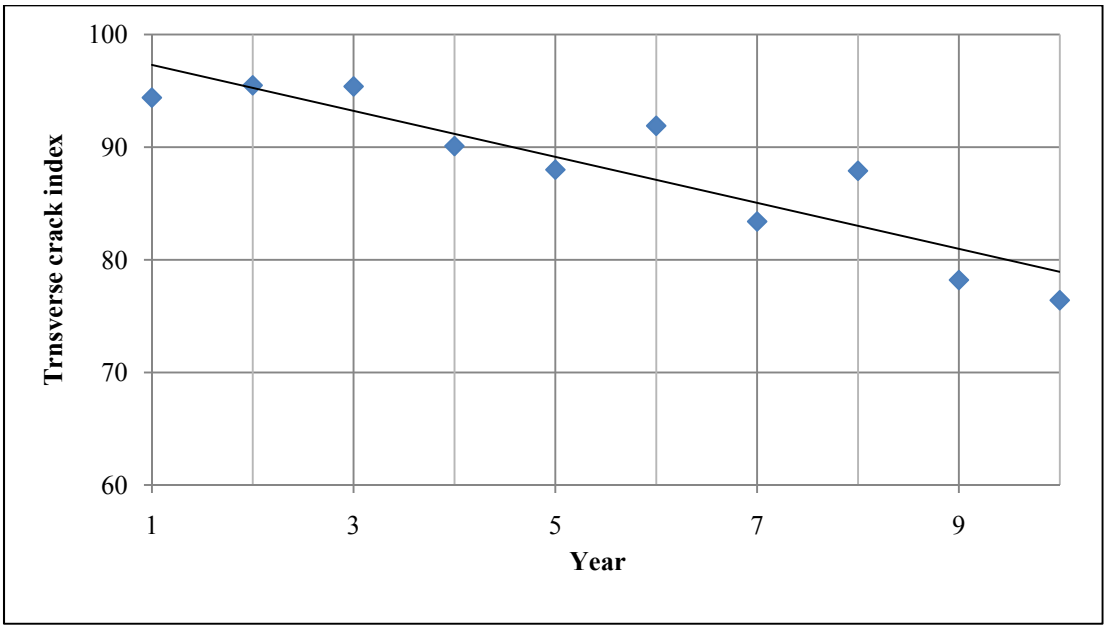



Figure 2 Transverse crack index versus time

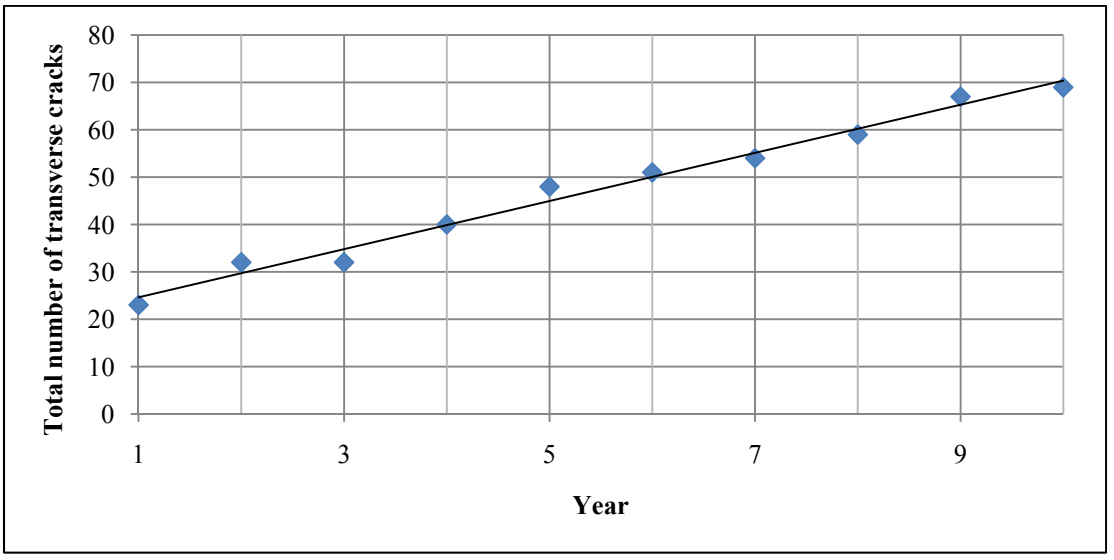



Figure 3 Total number of transverse cracks versus time

cracks. Tighter cracks may not allow as much water to penetrate the surface layer and damage the lower materials as wide open cracks. However, the integrity of the surface layer and its capability to spread the load to a wider area is broken. Hence, straight addition of the numbers of cracks in all three severity levels would be more meaningful to the decision makers than the transverse crack index. Analyses of distress data obtained from four State Highway Agencies showed this is to be the case for each road and for longitudinal and alligator cracks as well. In all roads, the variability of the transverse crack index data over time was so severe that no mathematical function could describe the data with good accuracy. When the numbers of low, medium and high

severity cracks were simply added, the data followed good trend against time. Some of the better transverse crack data (least variable data) obtained from the Washington State Department of Transportation were used as example and are shown in Figures 1 through 3. The important point herein is that, one does not need to transfer the number of transverse cracks or the cumulative length of longitudinal cracks to an index. Many people argue that the index has no units and could be easily communicated to other interested parties. Although this could be true, the value of the index could be tampered with on the basis of the priority factors between the various severity levels.

Nevertheless, the RSL of a pavement section could be calculated on the basis of the total number of transverse cracks without the need to calculate transverse crack index (Kuo et al, 1991, Novak et al, 1991, and Baladi et al, 1991). This could be accomplished if the absolute maximum number of acceptable transverse cracks is determined. To illustrate such calculation and the role of the RSL as a highly significant tool in the pavement management system, suppose that:

- The maximum allowable number of transverse cracks at which a 0.1 mile pavement section is in need of major rehabilitation is 90 (average transverse crack spacing of about 6 feet).
- The corresponding threshold value on the transverse crack index is 60.

Consider the transverse crack data of the 0.1 mile pavement section located in the State of Washington. The databank contains ten data points from 2000 to 2009. During this period, the pavement section was not subjected to any pavement preservation or maintenance actions (do nothing section). The total number of transverse cracks and the transverse crack index data were fitted to linear mathematical functions of the form shown, respectively in Equations 3 and 4.

$$TTC = \alpha(t) + \beta \qquad \text{Equation 3}$$

$$TCI = \eta(t) + \gamma \qquad \text{Equation 4}$$

Where TTC = total number of transverse cracks
TCI = transverse crack index
t = elapsed time or surface age in years
α, β, η and λ = regression constants

The above equations can be used to solve for the time t at which the threshold value of TTC of 90 cracks is reached and the threshold value of TCI of 60 is reached. The RSL can then be calculated by subtracting the surface age (SA) or the elapsed time stated in Equations 5 and 6 below.

$$RSL = \frac{[(TTC = 90) - \beta]}{\alpha} - SA \qquad \text{Equations 5}$$

$$RSL = \frac{[(TCI = 60) - \gamma]}{\eta} - SA \qquad \text{Equation 6}$$

The calculated RSL values can then be used to check the accuracy of the transverse cracking index and the total number of transverse cracks using the following steps:

1. Use the data points from the years 2000, 2001, and 2002 and calculate the RSL to the threshold value.
2. Repeat step 1 by using the 2000, 2001, 2002, and 2003 data.
3. Continue repeating step 1 by adding each additional year's data point.

The above three steps were taken to produce the seven RSL values listed in Table 2 for the TTC and TCI. Since the road section was not subjected to any preservation or maintenance actions, the value of RSL calculated in step 1 should decrease by one year in each subsequent step. Figure 4 depicts the calculated RSL values against the surface age (elapsed time) of the pavement section. The two solid lines in the figure represent the reference line along which the RSL value decreases by one year for every year increase in the pavement surface age. It can be clearly seen that the RSL values based on the total number of transverse cracks closely follow the solid line. On the other hand, the RSL values based on the transverse crack index deviate significantly from the corresponding solid line. Indeed, the first RSL value was negative 65 years.

The above illustration demonstrates the significance of the RSL and its important role as a powerful tool in managing pavements. The other role of the RSL as a quality control tool is presented elsewhere (Baladi et al, 2011).

## SUMMARY AND CONCLUSIONS

The distress data could be used to calculate distress indices and the RSL of pavement sections and the network. It is shown that a direct calculation of the RSL from the distress data yields more accurate results. The combination of low, medium and high severity transverse cracks based on priority factors produces high variability in the distress indices. If the distress data are accurate, the value of the RSL must decrease one year for every elapsed physical year unless preservation, rehabilitation and/or maintenance activities were applied.

## ACKNOWLEDGEMENT

The authors wish to express their sincere thanks to the US Federal Highway Administration (FHWA) for their financial support in sponsoring this study. Many thanks to the staff of the Arizona Department of Transportation, the Michigan Department of Transportation, the Louisiana Department of Transportation and Development, the Washington State Department of Transportation, and to the Colorado Department of Transportation for their support in obtaining the data in support of the analyses. The valuable inputs of various personnel of the Michigan Department of Transportation are deeply appreciated.

Table 2 Calculated RSL based on the total number of transverse cracks and the transverse crack index

| Surface age (year) | RSL (years) | |
|---|---|---|
| | Total | Index |
| 3 | 12.6 | -65 |
| 4 | 9.9 | 26.5 |
| 5 | 7.5 | 16.0 |
| 6 | 6.8 | 25.8 |
| 7 | 6.5 | 15.4 |
| 8 | 5.9 | 18.3 |
| 9 | 4.6 | 12.1 |
| 10 | 3.9 | 9.3 |

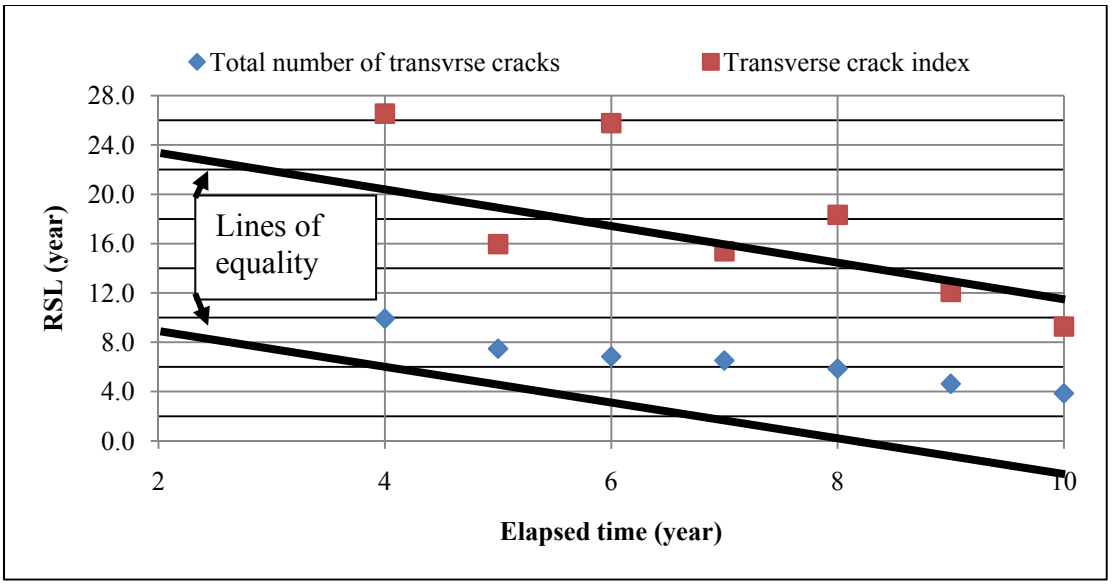



Figure 4 Remaining service life versus elapsed time

## REFERENCES

American Association Of State Highway And Transportation Officials, "AASHTO Guide for Design of Pavement Structures" Washington, D.C., 1986.

Baladi, G.Y., "In Service Performance of Flexible Pavements" volume II, proceedings International Air Transportation Conference, American Society for Civil Engineering, New Orleans, 1979.

Baladi, G.Y., T. Dawson, C.M. Dean, S.W. Haider, K. Chatti, " The Remaining Service Life, the Most Powerful Tool in Transportation Asset management", to be submitted for presentation and publication at the 90th Transportation Research Board Meetings in Washington DC, January 2011.

Baladi, G.Y., Novak, E.C., and Kuo, "W.H. Pavement Condition Index - Remaining Service Life" Special Technical Publication, ASTM Symposium on Standardization in Pavement Management Implementation, June 26-27, 1991, Atlantic City, New Jersey

Baladi, G. Y. and M. Snyder, "Highway Pavements", A 4 week short course, the U.S. Department of Transportation, the Federal Highway Administration, National Highway Institute, Volumes I, II, and III, Washington, DC, 1890.

Canadian Good Road Association, proceedings, Symposium on Pavement Design and Evaluation, 1960 - 1961.

ERES Consultants, Inc. "Pavement Design Principles and Practices", Champaign, Illinois, National Highway Institute, Washington, D.C., 1987

Kuo, W.H., E.C. Novak, and G.Y. Baladi, "Inclusion of Pavement Remaining Service Life Concept in PMS Methodology Standardization" Special Technical Publication, ASTM Symposium on Standardization in Pavement Management Implementation, June 26-27, 1991, Atlantic City, New Jersey.

Novak, E.C., G.Y. Baladi, and W.H. Kuo, "Toward Standardization of a PMS Analysis Method" Special Technical Publication, ASTM Symposium on Standardization in Pavement Management Implementation, June 26-27, 1991, Atlantic City, New Jersey.

Smith, R. E., M. I. Darter, and S. M. Herrin, "Highway Pavement Distress Identification Manual" Federal Highway Administration Report No. FHWA-RD-79-66. 1979. Transportation Research Board No. 1795, pp 17-27, 2002Transportation Research Board No. 1795, pp 17-27, 2002

Smith, X D., M. I. Darter, J. B. Rauhut, and W T. Hall, "Distress Identification Manual for the Long Term Pavement Performance (LTPP) Studies" Strategic Highway Research Program (SHRP), March 1987.

# Optimization of Step-Frequency Ground Penetrating Radar Protocols for Detection of Near-Surface Features

Scott, Michael L.,[1] Chintakunta, Satish R.,[2] Mekemson, James R.,[3] Gagarin, Nicolas[4]

[1]Starodub, Inc., 3504 Littledale Rd., Kensington, MD 20895, USA; PH (202) 493-3124; e-mail: mscott-ctr@dot.gov
[2]Starodub, Inc., 3504 Littledale Rd., Kensington, MD 20895, USA; PH (202) 493-3276; e-mail: satish.chintakunta@dot.gov
[3]Starodub, Inc., 3504 Littledale Rd., Kensington, MD 20895, USA; PH (202) 493-3016; e-mail: jim.mekemson.ctr@dot.gov
[4]Starodub, Inc., 3504 Littledale Rd., Kensington, MD 20895, USA; PH (301) 929-0964; e-mail: nicolas.gagarine.ctr@dot.gov

## Abstract

This article discusses the results of a research and development project to detect and evaluate inductive loops used in intersections for traffic detection and control using a step-frequency ground-penetrating-radar (SF-GPR) array. The frequency bandwidth, the dwell time, the transmitter-receiver patterns within the array, and the processing options defined the domain of an optimization effort to maximize the detection of the inductive loops. Descriptions of the application and the SF-GPR system and method are provided. A summary of the research effort and final results are presented. The inductive loops were detected and mapped at increasing levels of success with the optimization of data acquisition controls and data reduction options.

## Introduction

A nondestructive evaluation method for in-pavement sensor inspection was developed, tested and optimized in this study using data from a new integrated sensor system called the Advanced Pavement Evaluation (APE) vehicle and subsequent analysis of data from the system. The test method was evaluated using a range of several SF-GPR test parameters that were the focus of the optimization work in the study, including the dwell time, transmitter-receiver array scan pattern and data analysis processing options. The motivation for developing and optimizing this new test method comes from the high rate of in-pavement loop wire sensor failure due to deterioration, combined with the common use of these sensors throughout the United States, as described by Klein, 2006 in a Federal Highway Administration (FHWA)

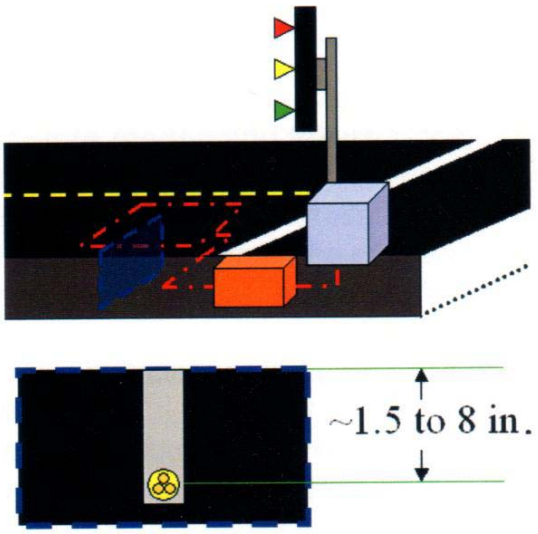



Figure 1. Loop wire schematic diagram

report. In the US and abroad these loop wire sensors are used for traffic actuated signal control, traffic responsive signal control, freeway surveillance and traffic management, and data collection systems, (Klein, 2006). A schematic drawing of an embedded in-pavement loop wire sensor system is provided in Figure 1, where the loop wire is the dot-dashed line in the upper drawing. The lower drawing in Figure 1 shows a cross section of the loop wire embedded in a saw cut, (subsurface).

State of the art Impulse GPR (I-GPR) technology does not currently have the resolution or the rapid three dimensional scanning capability required to effectively image loop wire sensors from a moving vehicle platform. Therefore, a nondestructive in-pavement loop wire evaluation method was developed and implemented by the research team for this study using an integrated SF-GPR sensor system together with three dimensional imaging and analysis software. After initial system integration and analysis development was completed, a systematic test plan was carried out using a laboratory test apparatus specifically devised for the study. Subsequently, field testing was conducted on embedded in-pavement loop wire sensors located on the grounds of Turner Fairbank Highway Research Center (TFHRC), a FHWA research laboratory, using the optimized configuration.

**Background**

Subsurface imaging and evaluation of buried infrastructure, (including in-pavement sensors that are the focus of this study), has had significant limitations in the past due to time consuming data acquisition requirements, problems with geometric alignment of data collection scans, sub-optimum system adjustability, and data processing inefficiencies. The subsurface evaluation method used in this study takes steps to address each of these problems with a new approach implemented using a new prototype system. This system has been configured to allow optimized performance to be achieved for subsurface imaging of in-pavement sensors.

A brief discussion of recent Ground Penetrating Radar (GPR) technology progress follows here. Until recently, GPR systems have typically been implemented using I-GPR technology. This I-GPR technology has improved substantially during the past few decades through advances in electronics and system design (Annan, 2002). I-GPR continued to be refined in recent years for applications to infrastructure through improvements in data processing techniques and incremental improvements in system hardware, (Al-Qadi, 2004; Scott, 2004; Saarenketo, 2006). From a functional perspective, I-GPR transmits impulses containing microwave frequencies across an ultra-wideband (UWB) range of frequencies. For civil infrastructure applications, I-GPR typically uses antennas that cover a fraction of the practical GPR operating bandwidth. Low, medium and high center frequencies for I-GPR antennas are often approximately 500 MHz, 1 GHz, and 2 GHz or higher, respectively (Daniels, 1996). In many I-GPR systems, the fractional bandwidth of interest must be anticipated before it is used so that the correct antenna with the appropriate fractional bandwidth can be selected.

In recent years, SF-GPR technology has become available for applications to subsurface infrastructure evaluation. New SF-GPR systems that have a superior signal to noise ratio to I-GPR, multiplexed antenna arrays with as many as 47 antenna pairs, fast scanning capabilities, and customizable system controls are beginning to be manufactured and applied to infrastructure evaluation problems, (Scott, 2006, Dec. 2006; Eide, 2002; Weedon, 2000; Binningsbo, 2000).

The SF-GPR sensor used in this study can be programmed for multi-offset data collection, such as Common Midpoint (CMP) data that can be analyzed to determine quantitative dielectric material properties of pavements without contacting the material or extracting a core. This SF-GPR sensor was the only one used for this project. Synchronized data collection and dielectric material calibration capabilities of the APE vehicle were used in this study to maximize the quality of images.

**Objectives**

The objectives of this study were to optimize the embedded in-pavement loop wire sensor evaluation process by following the steps described below:

1. Test and evaluate SF GPR subsurface imaging of embedded pavement loop wires using a systematic optimization of key test parameters during laboratory testing.
2. Evaluate capability to detect and image loop wire sensor embedded in pavement materials using optimized system configuration.
3. Future objective – use SF GPR analysis together with an inductive loop wire probe measurement to obtain an image of the subsurface loop wire and a probe based evaluation of the functionality of the loop wire. Together, these two measurements have the potential to quantify and locate loop wire problems. Loop wire problems include breaks or gaps in the wires and moisture intrusion into the loop wire saw cut.

## Test Design

Laboratory testing and field testing of embedded loop wires were conducted in this study. Many configurations were tested rapidly in the laboratory and selected configurations were then tested under field conditions. Figure 2 is an image of the laboratory configuration, including a yellow Step Frequency GPR antenna array.

Loop wires are positioned underneath or between the asphalt shingles, (used to closely simulate the properties of asphalt pavement in the laboratory), in an appropriate geometric configuration to simulate the effect of embedding a loop wire in an asphalt pavement.

The following parameters were varied during laboratory testing in order to provide a range of system configurations that was anticipated to include an optimum:

1. Transmitter-receiver array scan pattern, (transmitter receiver antenna array element combinations)
2. Antenna array height offset relative to the ground plane, (inches)
3. Frequency step size of the SF GPR data acquisition, (MHz)
4. Dwell time of the SF GPR data acquisition, (μsec time at each frequency step)
5. Data analysis processing options, (application of filters)

During subsequent field testing on TFHRC grounds, pictured in Figure 3, embedded in-pavement loop wires were scanned using the APE system SF GPR sensor. The resulting data was then analyzed to produce subsurface images. As Figure 3 illustrates, both diamond shaped loop wires, Figure 3a, and square shaped loop wires, Figure 3b, were scanned. The difference between the orientation of the loop wire in the diamond configuration and the square configuration is anticipated to change the response of the SF GPR due to the polarization of the antenna array elements, (polarized in the direction of travel of the van in Figure 3).


Figure 2. Laboratory test apparatus used to conduct parametric performance study.

(a) (b)

Figure 3. APE system scanning embedded in-pavement (a) diamond shaped loop wire seen as adhesive filled saw cut marks in the pavement surface in the foreground of the photograph and (b) square shaped loop wire.

**Analysis**

Laboratory data was collected using a range of optimization parameters for each data set and then each data set was analyzed using the procedure defined in the general processing flow diagram provided in Figure 4. The procedure starts with two sets of frequency domain data. One data set, frequency domain CF, is always collected using the calibration format corresponding to common midpoint data. The other synchronized data set, frequency domain DR, can be collected using common offset, alternating polarization, or cross polarization scan patterns.

After the data has been collected it must be processed to transform the original SF GPR data from the frequency domain, (which it was collected in), into the time domain. This is accomplished by carrying out an inverse Fourier transform procedure. Next, the time domain data is filtered, if necessary, to reduce spurious features. Then, a coherence analysis is carried out on the CF data to determine the radar wave propagation velocity and real dielectric material property value in the asphalt pavement layer that contains the loop wire. The basic coherence analysis method is described by equation 1. The coherence analysis is performed over a range of propagation velocities, which results in a two dimensional plot with peaks corresponding to the root mean squared (rms) velocity in each layer. Interval velocities for each pavement layer, (which can be computed from rms velocities using the Dix equation), allow real dielectric values and thickness values to be computed for each pavement layer (Robinson, 1983). Subsequently the real dielectric values can be used in imaging calculations that use data migration (Binningsbo, 2000) or wave field backpropagation (Mast, 1993) methods. For the current analysis, a three dimensional frequency-wavenumber (f-k) migration of the data is carried out (Binningsbo, 2000).

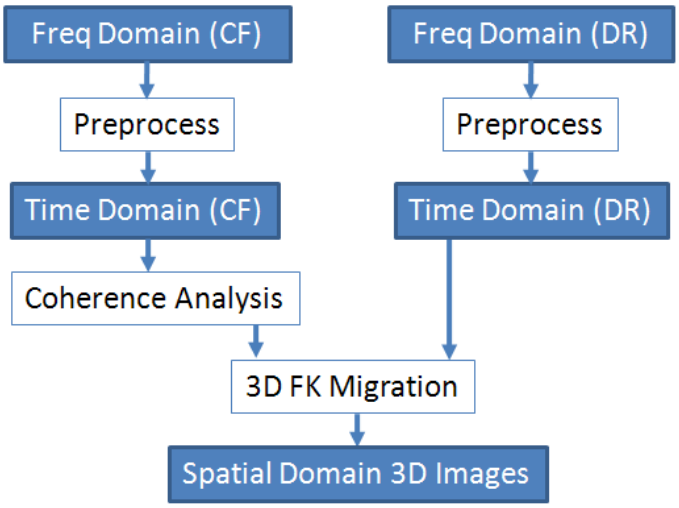



Figure 4. Data processing flow diagram.

$$C_t = \frac{\sum_{q=t-\tau}^{t+\tau}\left(\sum_{i=1}^{M} f_{iq}\right)^2}{M\sum_{q=t-\tau}^{t+\tau}\sum_{i=1}^{M} f_{iq}^{\,2}} \tag{1}$$

Where:

$C_t$ = Coherency computed at a given time "t"
$f_{iq}$ = Amplitude of the individual trace "i" at time "t"
$t$ = Time "t"
$\tau$ = Defines width of time gate which has a total width of $2\tau$
$i$ = Individual waveform trace index "i"
$M$ = Total number of waveform traces

The equations used to derive this f-k migration procedure for two-way plane wave propagation in the direction, $k$, of any target of interest within the synthetic aperture, are the following:

$$k = |\mathbf{k}| = \sqrt{k_x^2+k_y^2+k_z^2} = \frac{2\omega}{c} \tag{2}$$

$$k_z = \sqrt{\left(\frac{2\omega}{c}\right)^2 - k_x^2 - k_y^2} \tag{3}$$

$$F_m(k_x, k_y, k_z) = \frac{k_z}{|\mathbf{k}|} F_u(k_x, k_y, \frac{c}{2}\sqrt{k_x^2 + k_y^2 + k_z^2}) \tag{4}$$

Where:

ω = Frequency of SF GPR wave
c = Propagation velocity of SF GPR wave in a vacuum
$F_u$ = Unfocused frequency domain response data
$F_m$ = Migrated frequency domain response data

This migration procedure works via a substitution of variables, where the expression in equation 3 is solved for ω and substituted into $F_u$. This change of variables, and the use of equations 2 and 3 to cast this equation in terms of available information from the collected data, allows the migrated result $F_m$ to be computed. Finally, an inverse Fourier transform is carried out to produce final migrated output in the spatial domain. This output is a three dimensional volume of data.

**Results**

Laboratory results illustrate significant improvements in system performance due to the optimization procedure implemented in the study, while field results show the performance of the system in the optimized configuration. Example parametric output is presented graphically in this section and additional optimization parameters are subsequently discussed in terms of qualitative results. One of the most important optimization parameters from the laboratory study was the antenna array scan pattern used to evaluate the half diamond loop wire test configuration shown in Figure 5.

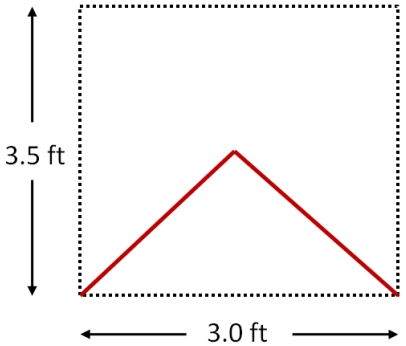



Figure 5. Plan view laboratory loop wire geometry scanned using various SF GPR configurations. 3 inches of asphalt material simulated pavement above the wire.

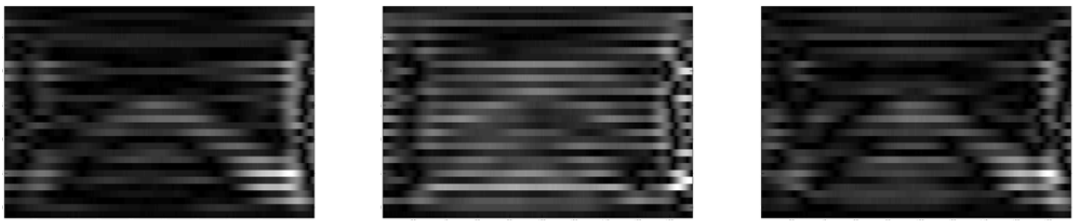



Figure 6. Plan view SF GPR images obtained from preprocessed data collected over the laboratory loop wire illustrated schematically in Figure 6. Scan pattern = CO.

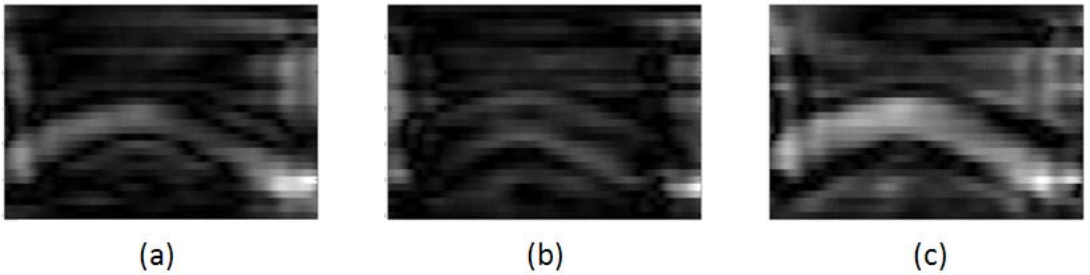



Figure 7. Plan view SF GPR images, a through c, obtained from preprocessed data collected over the laboratory loop wire illustrated schematically in Figure 6. Scan pattern = AP.

Figures 6 and 7 each illustrate a series of plan view preprocessed images of the Figure 5 loop wire, (before migration processing), where Figure 6 corresponds to common offset (CO) scan pattern data and Figure 7 corresponds to alternating polarization (AP) scan pattern data. In Figures 6 and 7, images a through c are each separated in time by 0.12 nanoseconds, (where the time associated with each plan view image increases from a to c). In Figure 6, the loop wire is detected, but differences in antenna characteristics among CO antenna array elements cause horizontal bands to appear in the data. Figure 7 data detects the loop wire more clearly via a direct reflection in Figure 7a and sharp diffraction patterns in Figures 7b and 7c. The AP configuration uses antenna array elements with more consistent frequency characteristics than the CO configuration. In addition, antenna array element polarizations that alternate between two orientations reduce the directional bias of imaged features in the AP data. Finally, the smallest available frequency step size,

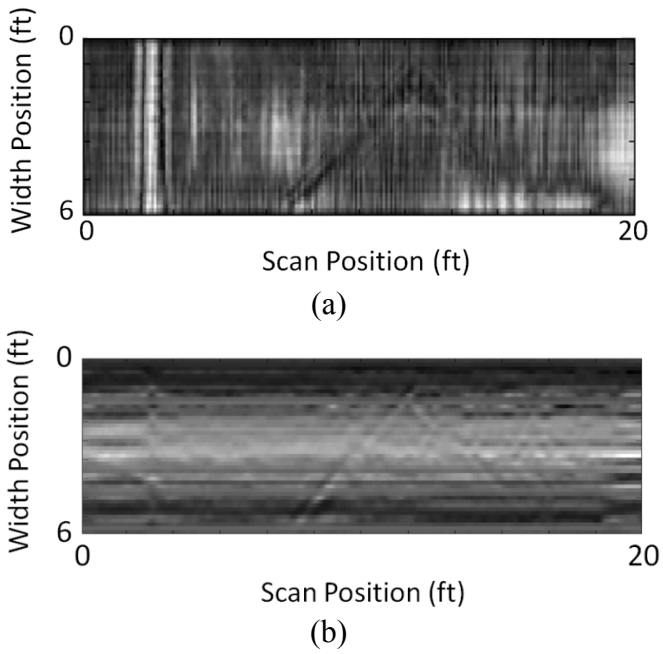



Figure 8. Migrated field images of (a) surface layer and (b) loop wire depth layer.

2.0 MHz, an optimum dwell time of 4.5 μsec, (longer dwell times did not significantly reduce noise) and an antenna height of 5 inches were used to produce the optimized Figure 7 results. In one other laboratory test scenario, a simulated loop wire discontinuity (a break in the wire) was detected and imaged using the optimized SF GPR configuration.

Figure 8 shows example results from processed field data migrated using a calibrated real dielectric value of 1.8. The data was collected over the left half of a diamond shaped loop wire using the APE vehicle in the same SF GPR configuration optimized in the laboratory. The Figure 8a image (a plan view layer corresponding to the pavement surface) includes a response to an aluminum tape marker placed at the 3 ft. scan position and spanning the width of the image. More importantly, the saw cut in the pavement surface is imaged in Figure 8a. The saw cut is observed as the left half of a diamond shaped response feature, (oriented based on a left to right scan). The parallel response features in the image correspond to the two sides of the pavement saw cut that the loop wire was installed in. Figure 8b shows the migrated loop wire depth layer image, where the loop wire is located inside the boundaries of the saw cut imaged in Figure 8a. Images 8a and 8b provide information that could allow loop wire discontinuities and defects or nearby pavement issues to be detected.

## Conclusion

A new method to systematically evaluate in-pavement loop wire sensors using an SF GPR sensor in the APE vehicle system was successfully optimized and demonstrated. After laboratory based optimization was completed, field data was successfully analyzed and presented. In the future, a complementary inductive loop wire measurement probe will be demonstrated together with the subsurface imaging system to measure the functional status of loop wires imaged using the new method.

## References

Al-Qadi, I.L. (2004). “Ground Penetrating Radar: State of the Practice for Pavement Assessment,” *Materials Evaluation*, Vol. 62, No. 7, 759-763.

Annan, A.P. (October 2002). “GPR – History, Trends, and Future Developments,” *Subsurface Sensing Technologies and Applications*, Vol. 3, No. 4, 253-270.

Binningsbo, J., Eide, E.S., Hjelmstad, J.F. (May 2000). “3D Migration of GPR Array-Antenna Data,” *Proceedings of the 8th International Conference on Ground Penetrating Radar*, Gold Coast, Australia.

Daniels, D.J. (1996), Surface Penetrating Radar, IEEE Press, London, 1-300

Eide, E.S., Hjelmstad, J.F. (May 2002). “3D Utility Mapping Using Electronically Scanned Antenna Array,” *Proceedings of the 9th International Conference on Ground Penetrating Radar*, Santa Barbara, CA.

Klein, L.A., Mills, M.K., Gibson, D.R.P. (October 2006). Traffic Detector Handbook: Third Edition – Volume 1, US DOT Federal Highway Administration Publication No. FHWA-HRT-06-108.

Mast, J.E. (1993). "Microwave Pulse-Echo Radar Imaging for the Nondestructive Evaluation of Civil Structures," Electrical Engineering Department Thesis, University of Illinois at Urbana-Champaign, 1-105.

Robinson, E.A. (1983). Seismic Velocity Analysis and the Convolutional Model, International Human Resources Development Corporation, Boston, pp. 1-290.

Saarenketo, T. (2006). "Electrical Properties of Road Materials and Subgrade Soils and the use of Ground Penetrating Radar in Traffic Infrastructure Surveys," Doctoral Dissertation, University of Oulu, A 471, 1-121

Scott, M. (2004). Nondestructive Testing Handbook, Vol. 5, Electromagnetic Testing, Chapter 17, Part 2, Infrastructure Applications of Electromagnetic Testing, American Society for Nondestructive Testing, 430-436.

Scott, M., Gagarin, N., Mills, M.K., Oskard, M. (December 2006). "Step Frequency GPR Evaluation of the Natchez Trace Parkway: Pavement and Infrastructure Measurement and Assessment for FHWA Digital Highway Measurement Vehicle Applications," *Highway Geophysics NDE Conference*, Saint Louis, MO, 325-336.

Scott, M.L, Gagarin, N., Mills, M.K., Oskard, M. (2006). "Step Frequency Ground Penetrating Radar Applications to Highway Infrastructure Measurement and System Integration Feasibility with Complementary Sensors," *Review of Quantitative Nondestructive Evaluation*, Vol. 25, American Institute of Physics, 1624-1631.

Weedon, W.H., Chew, W.C., Mayes, P.E. (2000). "A Step-Frequency Radar Imaging System for Microwave Nondestructive Evaluation," *Progress in Electromagnetics Research*, PIER 28, 121-146.

**Acknowledgement**

The researchers would like to acknowledge project sponsors at the Federal Highway Administration (FHWA) for resources to conduct this project under contract number DTRT57-06-R-SBIR. Contracting Officer Technical Representative (COTR) David Gibson and Contracting Officer (CO) Darren Shaffer from FHWA are thanked for their effort and dedication to supporting this work. In addition, Thomas Yu from the Federal Highway Administration is thanked for his collaborative work. Finally, 3D-Radar is acknowledged for manufacturing the SF-GPR sensor that was used in the study as part of the integrated, synchronized Advanced Pavement Evaluation (APE) data collection system prototype developed by Starodub, Inc. and used in this study.

# Adaptive Flow Control for Traffic Data Collection based on Low Bandwidth iDEN Wireless Network

Liping Zhang[1], Meng Li[2], Wei-bin Zhang[1], Kai Leung[3]

[1]California PATH, University of California, Berkeley, 1357 S46th st., Richmond, CA94804;

[2]Department of Civil Engineering, Tsinghua University, Bejiing, China

[3]Caltrans, Traffic Operations

## ABSTRACT

This paper describes a cost-effective traffic data collection system which is based on low bandwidth iDEN mobile network. An adaptive flow control protocol over the TCP/IP stack was developed to ensure a "best-effort" style data transmission over the low bandwidth network while the data link status is monitored to avoid unnecessary drop-offs caused by excessive demand of the wireless bandwidth. The iDEN handsets communicate via serial port with Caltrans field master signal controller using AB3418 standard. Signal phase and timing (SPAT) data from up to 10 signal controllers could be collected with one client at a maximum sampling rate of 200ms. Lab testing at California PATH showed that the data collection system based on the adaptive flow control provides an average upload data rate of 2.68kps per remote client for over 95% of the time, which could support SPAT sampling rate up to 200ms. Overall for more than 98% of the time the measured system availability is greater than 98%. The lab testing of the data collection showed significant improvement in the availability of the data collection system when compared with the transmission without our adaptive flow control protocol.

## INTRODUCTION

Traffic data has become more and more critical to the success of effective traffic operations and planning (Wolshon, Taylor, 1999). The traffic data could be obtained by either installing new infrastructure based sensors for high quality data (Hoogendoorn, 2003) or using the existing data sources such as the inductive loop detectors (ILD) which are widely deployed.

Using the existing low-bandwidth wireless commercial network as a communication means has become an attractive approach for the data collection. It reduces the equipment cost by using the lost cost programmable Commercial off the Shelf (COTS) devices. It also reduces the operational cost as well, due to the competitive price of the service. The low throughput of the wireless network, however, limits the application to only low data rate sources such as the ILD data or sampled high quality sensor data when real-time data collection is needed.

This paper describes a specific design of a cost-effective real-time arterial traffic data collection system which is based on existing mobile communication networks and Motorola® iDEN mobile handsets. The purpose of the system is to collect signal phase and timing (SPAT data and the ILD data) from Caltrans 170E signal controllers in real-time (Kell and Fullerton, 1982).

# SYSTEM OVERVIEW

## Data Collection System

The data collection system consists of three layers: field layer, data server layer, data application layer.

At the field layer, the Motorola iDEN series handsets with customized Java programs are capable of polling traffic signal status and traffic detection data from the local controllers through an RS232 connection using AB3418 protocol (Caltrans, 1995) and forwarding the data to the remote traffic data server via wireless communication. The developed field set-up does not require any hardware changes in the existing infrastructure in control cabinets. The entire data collection process is automatic and in real time. Field setup is illustrated in Figure 1.

Using the mobile telecommunications network enables a very cost effective way to interconnect the distributed traffic controllers. The major disadvantage of the wireless communication is the highly variable nature of the link quality, especially for the data service. Majority of the outages and latency incurs over the wireless link (Sinha, 2002). Therefore we adopt an adaptive flow control protocol to cope with the link variation problem. The achieved system performance measurements, including the throughput, data loss percentage, etc. are presented in the Section PERFORMANCE CHARACTERISTICS.

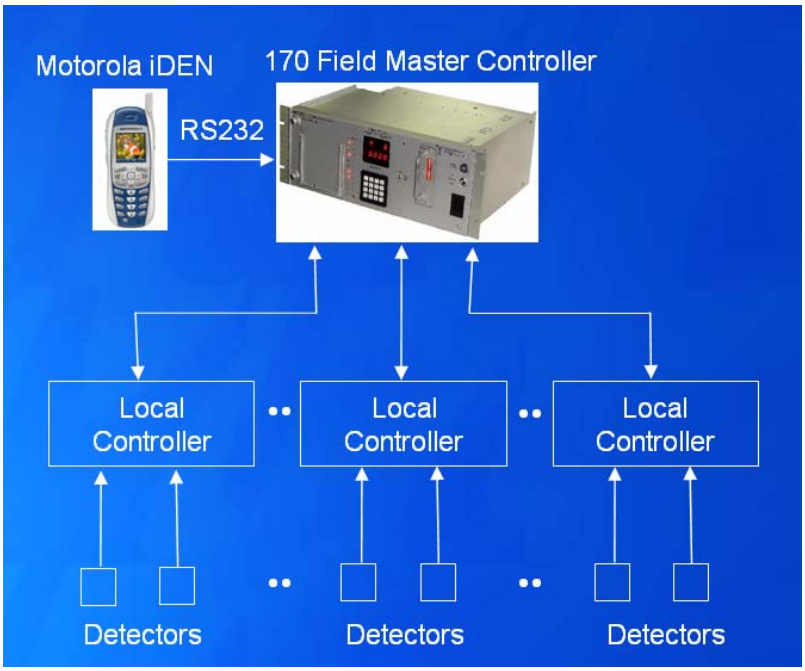



**Figure 1 Field Set-up between Motorola iDEN wireless device and Caltrans 170E Controller**

## SPAT and ILD Data

The master signal controller polls the local controllers either in *focus mode* or *normal mode.* The data collection system needs to handle data from master controller at an interval of as short as 200ms.

The length of messages varies from several bytes to as long as 67 bytes (Caltrans 170E AB3418). When all sentences are 67 bytes long and the master controller is working in focus mode, the maximum requirement for air data rate will be 335byte per seconds which equals to 2680bps. So the system requirement of the continuous data throughput has been set as 2680bps. The lab testing showed that the system meets this requirement over 90% of the time. Details of the test results can be found in Section PERFORMANCE CHARACTERISTICS.

# DATA COLLECTION OVER WIRELESS NETWORK

## Characteristics of Wireless Channel and Real-time Traffic Data

There are two factors related to the variations of the data rate. One is that the demand (data from the signal controller) variation from time to time and from site to site. The other factor is the variable nature of the wireless channel. There is an outage probability, such that during some short period, no data can be transmitted. Even when there is no outage, the wireless data communication is still vulnerable to multi-

path, rain degradation and other factors. These are factors common to wireless networks, and need to be taken into account for the continuous data collection system.

Occasional loss of the real-time traffic data is acceptable. Occasional loss of the traffic data could be considered as a loss of the sampling rate, if handled properly. Therefore we have also taken into account (and taken advantage of) this characteristic of the traffic data to selectively discard data when throughput rate is over the channel capacity.

To ensure a continuous, high throughput data communication over the wireless iDEN network using TCP/IP, we developed an adaptive flow control protocol. On the one hand, the remote device transmits more when the signal controller is sending more data to the remote device. On the other hand, the remote device delivers data in a "best effort" way, which means it sends as much as the channel currently allows during a given period and discards the data which fails to be delivered after a few seconds.

## Adaptive Flow Control of Traffic Data Collection

At the remote device end, every several data packets, it will send to the data center a label packet (L) which serves as a virtual timestamp and indicates the relative sequence and absolute numbering of the data packets sent. At the data center end, the server program maintains a slide window buffer and monitors label sequence. The server sends acknowledgement label packets back to the remote device, so that it could tell the latency from the labels and selectively discarding data from its internal slide-window buffer whenever necessary. The flow control mechanism tries to avoid channel saturation caused by greedy client programs without awareness of the instantaneous link quality measurements.

The wireless flow control protocol is for adaptive flow rate control. This keeps the wireless link operating at a close-to-maximum-allowed data rate. Package loss may occur due to the flow rate control.

A more detailed illustration is shown below in Figure 2 of this flow control procedure. Also Figure 3 is a flow chart of the remote device program flow control feature.

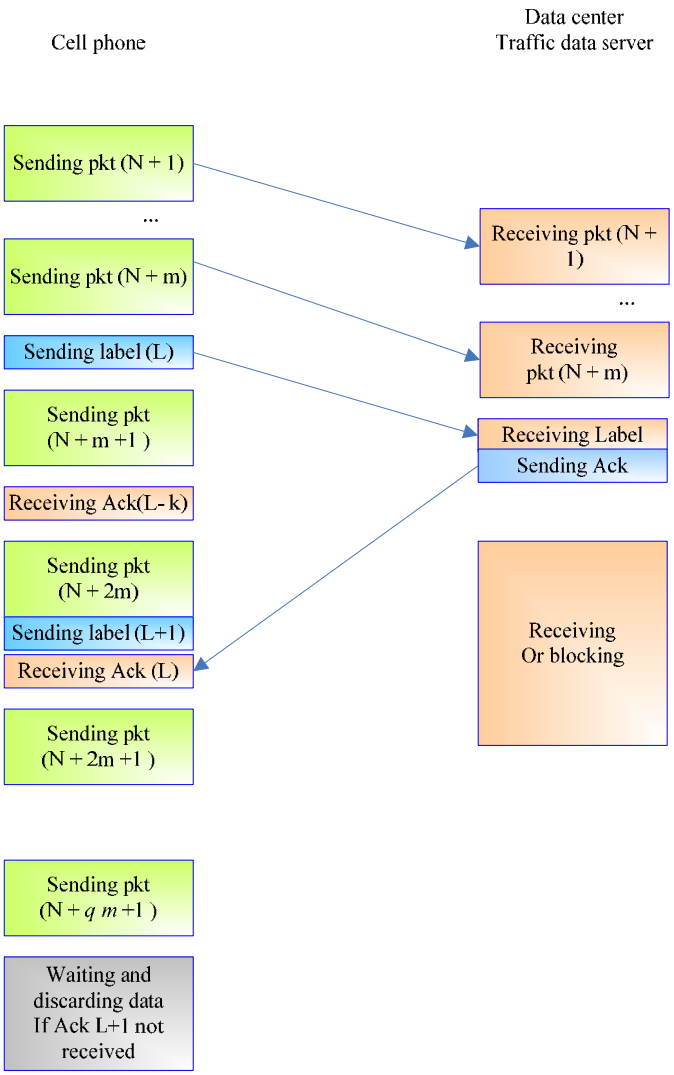



**Figure 2 Adaptive flow control mechanism**

Every *m* packets are grouped to form macro packets, which are indexed and used as "hand-shake" labels between the remote devices and the traffic data server. The remote device then determines the instantaneous latency from this pseudo timestamp label by comparing the received acknowledgement label and internal outgoing label. Under good communication conditions, this difference can be kept low, but with system outages, the difference will increase quickly. The macro packet size *m* and the threshold of label difference *q* are all selected according to empirical data obtained through experiments. The loss due to flow control is kept to a minimum by carefully selecting parameters and using a windowed buffer.

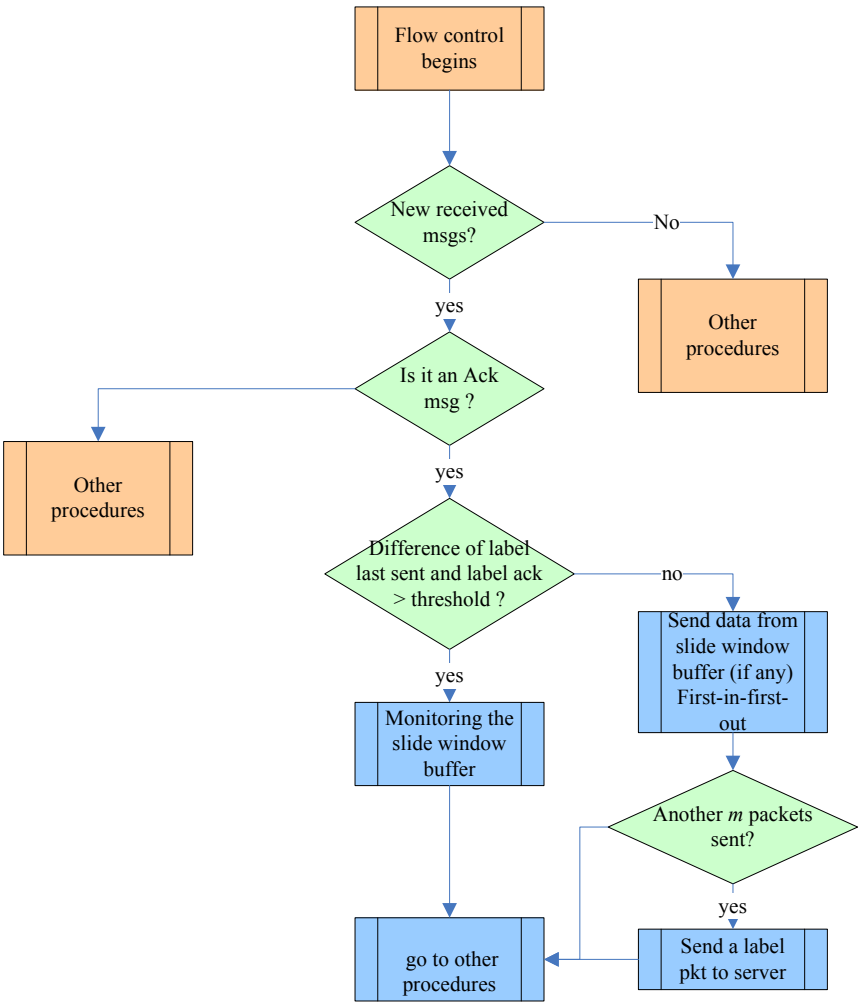



**Figure 3 Flow control procedure in the remote device**

# PERFORMANCE CHARACTERISTICS

## System Performance Indexes

The system performance characteristics described in this report are the throughput and the instantaneous link quality (which shows the service availability). Preliminary results on the reliability of the system are also reported based on tests conducted at University of California at Berkeley Richmond Field Station.

The statistics of eight remote devices and one data center over 10 days were averaged to form the following average performance indexes as shown in Table 1. Distributions of the system performance indexes are listed in the following subsections.

**Table 1 Average system performance indexes**

| Performance | Average | Definition |
|---|---|---|
| Instantaneous throughput | 619Bytes/s | Number of bytes received per second by the data center from one remote device, measured every 10 seconds. *Note: these statistics do not include measurements taken when there is a communication outage.* |
| Hourly throughput | 533Bytes/s | Number of bytes received per second by the data center from one remote device. Measured every hour. |
| Instantaneous System Availability | 99.5% | The number of bytes received by the data center divided by the number of original bytes sent by the signal controller to the remote device, measured every 10 seconds |
| Hourly system availability | 99% | The number of bytes received by the data center divided by the number of original bytes sent by the signal controller to the remote device, measured every hour |
| Latency | 2 s | The time a packet takes to travel from the source (only the GPS message has its original time stamp, so the source originates from the GPS satellites) to the data center.<br><br>Due to a lack of high resolution timestamp, the latency is estimated to be roughly 2s in most observations. |

**Throughput** Both the hourly average throughput and instantaneous throughput at 10 seconds period were obtained for all the clients under test. The tests were carried out at Richmond Field Station, Richmond, CA, where the communication network condition was worse than at several of the arterial intersections we tested, including one in Palo Alto, CA. The measured system throughput and service availability at the field test locations were higher than those obtained at the Richmond Field Station.

Figure 4 and Figure 5 show the cumulative distributions of instantaneous and hourly throughput, respectively. It shows that, the instantaneous rates (regardless of the communication outage) are highly probably greater than 335B/s most of the times. That rate is the throughput required when the master controller works in focus mode and is polling with a 200ms period. This is accomplished with a probability of over

96%, while rates higher than 335B/s over 90% of the time can be sustained over the long term when outage and other losses are taken into account.

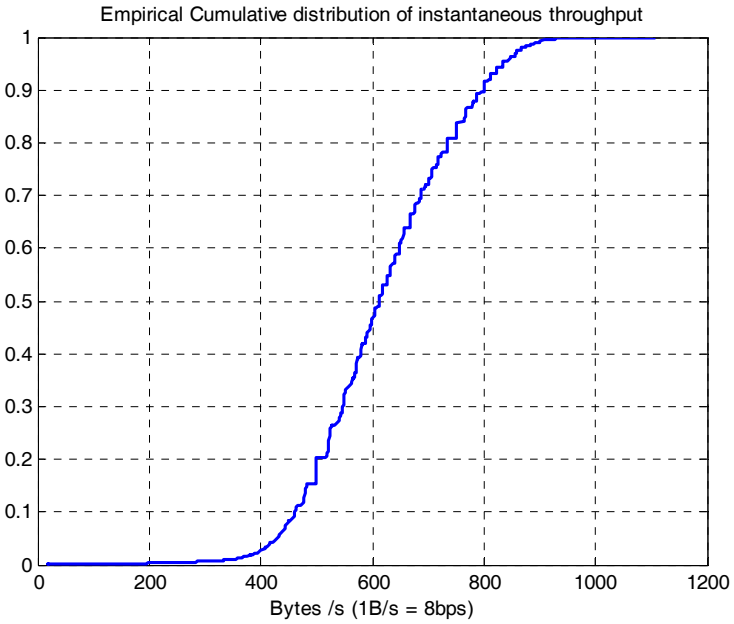



**Figure 4 Cumulative distribution of the instantaneous throughput (Bytes/s)**

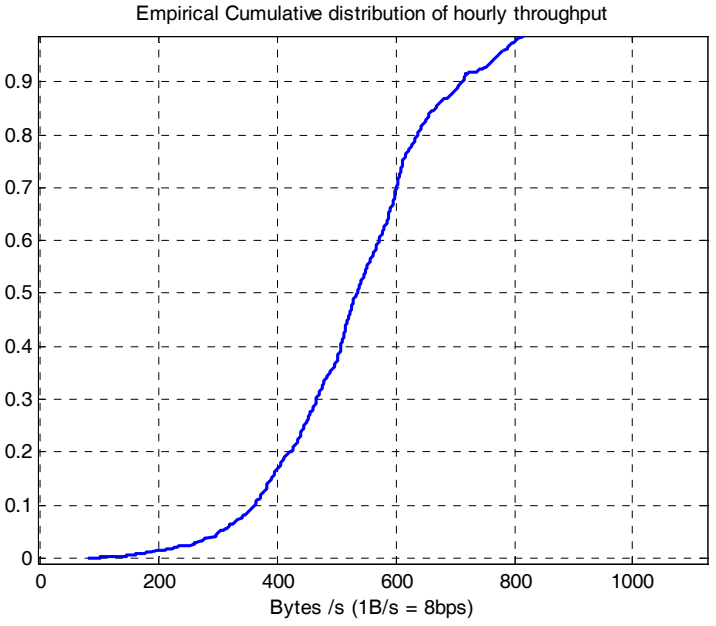



**Figure 5 Cumulative distribution of hourly throughput (Bytes/s)**

**The service availability** is defined as the number of bytes received by the data center divided by the total number of bytes the original signal controller sent to

the client. It is always less than 1.0, so hereafter service availability is presented in percentage.

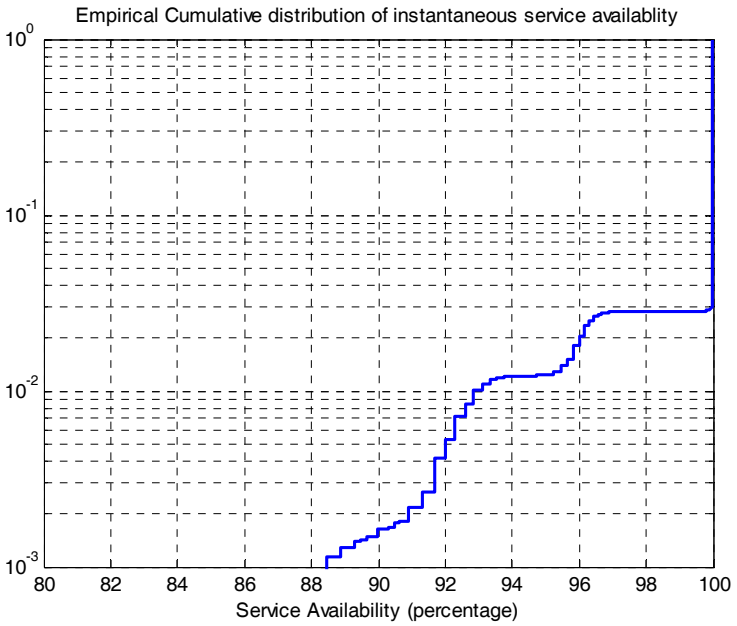



**Figure 6 Instantaneous system service availability**

From Figure 6, the probability of data lost due to flow control being greater than 0 is only 2%. **Figure 7** illustrates the data loss due to flow control and outage being greater than 2% is about 2%.

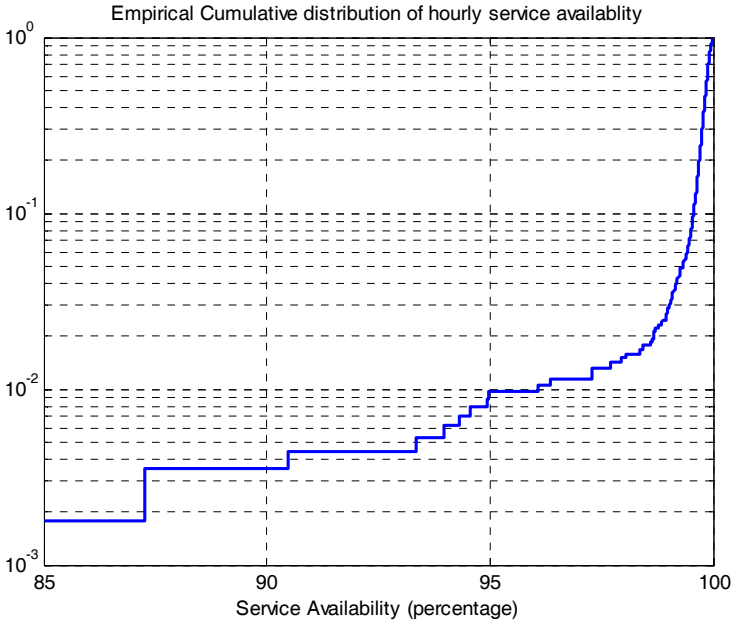


**Figure 7 Hourly system service availability**

## CONCLUSIONS

A cost effective traffic data collection system has been developed at California PATH at UC Berkeley. The system is designed for the purpose of collecting real-time Caltrans 170E signal controller data, including the SPAT and the ILD data. To improve the overall system performance over the outage-prone low cost wireless network, we have developed an adaptive flow control protocol that measures the real-time latency of the client server communication and selectively discard data packages when the link is congested to improve overall availability of the data collection system. The lab testing showed that over 98% of the time the system achieved availability rate greater than 98%. The system also achieved the maximum required rate of 10 local Caltrans 170E signal controllers sampling SPAT and ILD at every 200 millisecond (2.68kps) over the iDEN network for over 95% of the time.

## ACKNOWLEDGMENT

This work was sponsored by the State of California Business, Transportation and Housing Agency, Department of Transportation (Caltrans). The contents of this paper reflect the views of the authors, who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California.

## REFERENCES

Basu, A. P. (1964), Estimates of Reliability for Some Distributions Useful in Life Testing, Technometrics, Vol.6, 1964, pp 215-219

Caltrans (1995), Standard Communications Protocol for Traffic Signals in California, 1995, http://www.dot.ca.gov/hq/traffops/elecsys/reports/ab3418sp.pdf

Hoogendoorn, S. P., et al. (2003), Microscopic Traffic Data Collection by Remote Sensing, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1855, 2003, pages 121-128

Kell, J.H. and Fullerton, I.J (1982), Manual of Traffic Signal Design, Institute of Transportation engineers, Prentice Hall, 1982

Sinha, P. et al. (2002), WTCP: a reliable transport protocol for wireless wide-area networks, Wireless Networks, Vol.8, No.2, 2003, pp 301-316

Wolshon, B. and Taylor, W. C. (1999), Analysis of intersection delay under real-time adaptive signal control, Transportation Research Part C: Emerging Technologies, Volume 7, Issue 1, February 1999, Pages 53-72

# Reinforcement Learning Agents for Traffic Signal Control in Oversaturated Networks

Juan C. Medina[1], Rahim F. Benekohal[2]

[1]PhD Candidate, Department of Civil and Environmental Engineering. University of Illinois. Urbana, IL. 61801; PH (217) 333-5967; email: jcmedina@illinois.edu
[2]Professor, Department of Civil and Environmental Engineering. University of Illinois. Urbana, IL. 61801; PH (217) 244-6288; email: rbenekoh@illinois.edu

## ABSTRACT

An application for traffic signal control using reinforcement learning agents was implemented for small traffic networks with volumes close to saturation. The state observed by the agents and the rewards included information from the intersection being controlled and also from adjacent intersections. Communication between neighboring agents resulted in emergent coordination between agents and ultimately in better handling of traffic. Lower average, maximum and minimum delay values were found for the two tested networks compared to optimal pre-timed settings. Trends indicate that if the minor intersecting streets are one way, increased organization and coordination of traffic are expected as the network size increases. There is potential for reinforcement learning agents in traffic control applications as they can provide control in real time with flexible timing settings. Further research is being conducted with variable volume, bigger networks, and adjusting the agents' parameters.

## INTRODUCTION AND BACKGROUND

The increases in traffic volumes in the last decades and the limited available capacity in the roadway system have paved the way for improvements in traffic control. From traditional pre-timed isolated signals to actuated and coordinated corridors, traffic control has recently evolved into more complex "adaptive" signal control systems. New developments in information technologies have also resulted in applications that make use of machine learning techniques with great potential for real-time flexible traffic control. Such techniques include reinforcement learning (RL). Agents learning with RL techniques, different from actuated systems, can operate solely on the basis of traffic needs without restrictions such as maximum green times or cycle lengths. RL agents operate based on the perceived state of the system, making decisions and storing information about the "goodness" of the past actions. They can also communicate with each other to achieve a common goal (e.g. process vehicles with minimum delays).

An account of previous studies related to RL applications for traffic signal systems is difficult to summarize in a few paragraphs. Nevertheless, references to some of the work in this topic are provided next, as they may provide important sources for the reader to consult in more depth. One of the first algorithms using RL

and delayed rewards was introduced by Watkins (1992) when he presented the Q-learning algorithm. A few years later, a very complete and widely used publication was presented by Sutton and Barto (1998) with their book "Reinforcement Learning: An Introduction". Specifically for traffic control applications, the work by Thorpe (1997) used the RL algorithm SARSA for assigning signal timing to different traffic control scenarios. Wiering (2000) discussed a state representation based on road occupancy and mapping the individual position of vehicles over time, and Bakker (2005) later extended this representation using an additional bit of information from adjacent intersections. This allowed communication between agents, trying to improve the reward structure and ultimately the overall performance of the system. Using a different approach, Bingham (1998, 2001) defined fuzzy rules to determine the best allocation of green times based on the number of vehicles that would receive the green and red indication. He presented a neural network to store the membership functions of the fuzzy rules, reducing memory requirements. Also, a Cerebellar Model Articulation Controller (CMAC) has also been used in the past to store the learned information (Abdulhai, 2003), more specifically, approximate Q values.

To avoid the exponential explosion of the state space as the number of intersections increases, the scalability of the problem has been approached from different angles. Kuyer (2008) used coordination graphs and the max-plus algorithm to connect intersections close to each other. Networks having up to 15 intersections were tested, finding improved results compared to Wiering and Bakker. Additional work has been conducted by Oliveira and Bazzan (2004, 2005) in terms of coordination and cooperation of agents in a traffic network, using novel approaches such as swarm intelligence and dynamic formation. Also, Xie (2007) and Zhang (2007), explored the use of an actor-critic temporal difference agent for controlling a single intersection, and expanded the setup to an arterial using the same scheme previously applied to the isolated intersection.

In this paper, the focus is given to the application of a Q-learner that used a special reward structure to learn from the environment. A mixed state representation was also defined to combine information from the intersection controlled by the agent and from adjacent intersections. In addition, experiments were conducted in small networks that received traffic volumes close to the saturation level.

Thus, the main objective of the paper is to present a different approach to operate the traffic signals of a small network, by using a set of RL agents with improved state representation and the reward structures in cases at or close to oversaturation.

The remaining of the paper is organized as follows. The next section explains the basics of reinforcement learning, followed by a description of the state and the reward structures. Then, the procedure to conduct the experiments and the conditions simulated using the RL implementation are explained. The results and discussions are presented next, and the paper finalizes with conclusions and plans for future research.

## BASICS OF REINFORCEMENT LEARNING

Assuming that a system can be represented as a Markovian system, and that the values of the states are known (based on discounted rewards), the well-known

recursive Bellman equation shows that the value of a given state can be expressed solely on the value of the potential states following the immediate action:

$$V^{\pi}(s) = \sum_a \pi(s,a) \sum_{s'} P^a_{ss'} \left( R^a_{ss'} + \gamma V^{\pi}(s') \right) ,$$

where $P^a_{ss'}$ is the probability of transitioning to state *s'* given the current state *s* and that the action taken is *a*, $R^a_{ss'}$ is the reward of such transition due to action *a*, and $\gamma$ is a discount factor for the value of the next state. Thus, the problem can be thought as finding the policy that guarantees maximum expected rewards:

$$V^*(s) = max_{\pi} V^{\pi}(s), for\ all\ s \in S ,$$

or simply put in terms of state-action pairs, the problem is one of finding the policy with action-value functions ($Q^{\pi}(s,a)$) leading to maximum expected total rewards:

$$Q^*(s,a) = max_{\pi} Q^{\pi}(s,a)$$

However, since the true discounted value of the states or actions are not known (otherwise finding optimal policies would not be a problem), some algorithms, including temporal difference (TD) learning methods, have been used for this purpose. TD learning methods, compared to other methods to solve reinforcement learning problems (e.g. dynamic programming and Monte-Carlo) are well suited for traffic control applications since: *a*) the agent can learn without knowing the dynamics of the environment, and *b*) the agent updates the best action estimates based on previous data (bootstrapping), so there is a solution for every state at every point in time (i.e. any-time algorithms).

Out of a handful of TD algorithms, Q-learning has shown to produce good performance for a variety of problems under stationary conditions, even though the convergence of Q-values has only been proven if the states are visited an infinite number of times (Watkins, 1989, 1992). Practical decision making, however, do not require full convergence of Q values as long as they are "sufficiently" different for the agent to commit to the best choice. On this regard, precise boundaries of the algorithm for decision-making purposes only are not well defined and require further research.

For the specific case of traffic signal control, the maximum expected reward for each state can be obtained (by means of an appropriate TD algorithm) if the decision of giving the right of way to a particular traffic movement can be properly expressed as a function of variables that are memory-less and that also account for the main components that affect the efficient movement of vehicles.

A second challenge in formulating a good RL application for traffic purposes is to define a meaningful function that estimate the discounted reward for taking a decision. These two components of a typical Q-learning algorithm (state and reward) have been often overlooked. The results are functions that over-simplify the traffic problem to the point that it may not longer represent the dynamics of the system. This is the main reason why a more elaborated state representation and reward structures are considered in this paper.

## STATE REPRESENTATION AND REWARD STRUCTURE

A mixed state representation that combines information from the approaches at the controlled intersection and also from neighboring approaches was used. The general form of the equation that defines the state perceived by the agents is as follows:

$$S = \{s_{ew}, s_{ns},\ b_{ew}^{up}, b_{ns}^{up}, b_{ew}^{down}, b_{ns}^{down}, g, d\}$$

Where S is the state vector; s represents both the number of vehicles and the time they have spent in the link in the east-west (ew) and in the north-south (ns) directions; $b^{up}$ and $b^{down}$ describe the occupancy of the links in upstream and downstream intersections, respectively; *g* indicates the approach receiving the green indication; and *d* stores the duration of the current phase.

Note that one of the components of the state vector are the occupancies from upstream and downstream intersections. Given that downstream links can be oversaturated, producing intersection blockages due to queue spills, monitoring the remaining capacity of these links can be critical. In addition, coordination and anticipation of increased arriving volumes can be achieved by keeping track of the occupancy of upstream intersections, allowing a faster discharge of the current queues prior to the arrival of new vehicles.

Similar to the state representation, the reward structure consists of a combination of several factors, as follows:

$$R = \beta_1 x_g + \beta_2 x_r + \beta_3 p_t + \beta_4 p_d + \beta_5 i_u$$

Were *R* is the total reward; $x_g$ and $x_r$ indicate the number of vehicles receiving green and red, respectively; $p_t$ is a penalty for transitioning from one phase to the next (when the phase is terminated) and accounts for the lost time; $p_d$ is a penalty for giving green to downstream links close to capacity (discouraging queue spills); and $i_u$ is an incentive for giving green when the upstream link is close to capacity (favoring coordination). The vector of coefficients β provides weights to give more importance to specific components of the rewards.

## SIMULATION EXPERIMENTS

Two scenarios were created in a microscopic simulation environment provided by Vissim: a) a network of size 2x3; and b) a network of size 3x3. The default Vissim values for the different drivers' characteristics and car-following parameters were used in the simulations presented in this paper.

The Vissim COM interface was used in order to make decisions and operate the signals while the simulation was running. This interface has an option to define an external controller that can be programmed in C++. Vissim provides the structure of the C++ code, allowing the user to interact with the simulation in real time through a series of predefined functions. These functions provide direct access to values such as simulation time, the status of the signals, and detector information.

Thus, the complete reinforcement learning application, including the q-learning update rule, the reward structure, and the state representation were programmed in C++. Then, a dynamic linked library (.dll) was created so that Vissim could read the controller instructions at running time.

The purpose of the simulations was to determine if there was potential for the reinforcement learning scheme defined above to adequately manage the traffic signals located in close proximity to each other. The complexity was increased as the scenarios moved from *a*, to *b*, mainly because of the increased number of participating intersections.

All approaches had two lanes and no turning movements were allowed. The main roadways had two-way traffic and the minor streets had only one-way traffic. A schematic representation of the 2x3 network is shown in Figure 1 for illustration purposes.

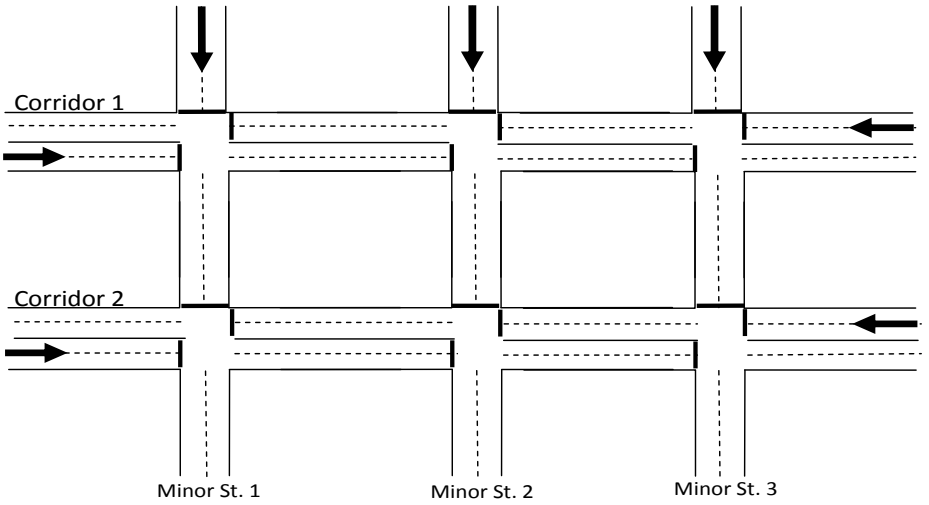



**Figure 1. Schematic representation of 2x3 network**

The two scenarios were tested under two different volume conditions: a) high volume, undersaturated conditions, and b) high volume, close-to-saturation conditions. For the undersaturated conditions, 1700 vph were input at each of the entry points. For the close-to-saturated conditions, 300 vph were added to the arterials in each direction, so that the total volume to be processed at a given intersection was the sum of 1700 vph from the minor road, 2000 vph in the EB and 2000 vph in the WB of the arterial.

The reinforcement learning agents were trained during 60000 simulation seconds for the 2x3 and 3x3 networks. After the trained Q-values were obtained, they were stored and served as the starting point for an additional run based on the learned policies. The results presented in this paper are based on the best actions from all agents after their trained period.

It is noted that the actual time required to train the agents was about $1/30^{th}$ of the simulation time, using a PC with two processors running at 2.4 GHz and 4 GB in RAM. Thus, in practice less than one hour was required for such stage.

## RESULTS

Results from the reinforcement learning agents were compared to the best pre-timed signal timing settings under the specified volumes. The values provided in this section are based on the last 2 hours of simulation time from a run where the agents used the policies created in the training period. A value was obtained from Vissim every 5 minutes of simulation time, thus a total of 24 values were obtained from the last 2 hours of simulation.

Given that traffic volumes per hour did not change in the simulation, pre-timed signal operation is expected to provide adequate results. In addition, the pre-timed signals were set such that one direction of traffic was coordinated along the two-way arterials, and also along the one-way roadways (minor streets).

Results from the pre-timed and the reinforcement learning agents were compared in terms of delay. A group of data collection points were defined in the networks to obtain the measures of performance. These points covered roadway sections starting at about 200 ft upstream from the initial signal, and ending about 200 ft after the last signal. For example, in the 2x3 network, the delay along one of the arterials started upstream from the first traffic signal and ended downstream of the third (and last) traffic signal.

The first scenario (2x3 network) is analyzed next. The results in terms of delays are shown in Tables 1 and 2 for the undersaturated and the close-to-saturation conditions, respectively. The average, maximum, and minimum delays per vehicle for each of the links in the networks are included. In the pre-timed case, it is obvious that the coordinated traffic in the main corridor was eastbound (EB) direction.

From Table 1, the first observation is the lower total average delay obtained with the reinforcement learning compared to the pre-timed signals. Second, there is more balance between the delays in the EB and the WB direction when using the agents than in the pre-timed. This is expected since one direction is typically coordinated with pre-timed signals, increasing the expected delays for the traffic in the other direction. Third, the maximum and minimum delays experienced by a vehicle were lower using the agents compared to the pre-timed signals. Overall, it could be said that for this specific condition, the reinforcement learning agents showed better performance that the pre-timed signals.

**Table 1. Delay for the 2x3 Network with High Volume and Undersaturated Conditions**

| Link | Direction of Traffic | Pretimed | | | | Reinforcement Learning | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | # of vehicles | Delay (veh/s) | | | # of vehicles | Delay (veh/s) | | |
| | | | Average | Max | Min | | Average | Max | Min |
| Corridor 1 | EB | 3216 | 21.7 | 36.6 | 16.7 | 3349 | 30.0 | 50.0 | 16.9 |
| | WB | 3329 | 69.3 | 85.2 | 57.4 | 3356 | 23.1 | 37.0 | 15.2 |
| Corridor 2 | EB | 3269 | 20.1 | 26.2 | 15.0 | 3370 | 28.3 | 43.2 | 18.7 |
| | WB | 3392 | 70.0 | 84.5 | 56.8 | 3370 | 20.1 | 34.2 | 10.0 |
| Minor St. 1 | | 3188 | 41.2 | 54.0 | 35.6 | 3289 | 45.7 | 67.3 | 28.3 |
| Minor St. 2 | | 3328 | 44.2 | 53.6 | 38.7 | 3318 | 54.7 | 70.3 | 36.0 |
| **Totals** | | **19722** | **44.7** | | | **20052** | **33.6** | | |

Regarding Table 2, the operation of the reinforcement learning agents in the close-to-saturation condition was, in general, superior to pre-timed settings. This is similar to the case described above (undersaturated condition). On average, delays on the two corridors were lower with the agents, with a tendency to prioritize the WB direction, generating higher delays for the EB traffic.

Results from the second scenario (3x3 network) are shown in Tables 3 and 4. The performance of the agents in these two cases was better than the pre-timed signal timings. For the undersaturated conditions, average delays were lower and more balanced with the agents, which also resulted in lower maximum and minimum values.

For the case close to saturation, also lower average, maximum, and minimum values were obtained (see Table 4). Similar to the results in the 2x3 network, the agents learned policies that generated improved and more balanced results. In particular, note that the ratio of delay from the EB and the WB for corridor #3 is close to 1 using the agents, but more than 3 for the pre-timed signals. Particular importance is given to corridor #3 because greater coordination effects are expected to be seen at the traffic progresses along the minor streets in the network, from corridor #1 through #3. Note how in Table 4 there is a decreasing trend in the delays from corridor #1 to #3, and how for the reinforcement learning

**Table 2. Delay for the 2x3 Network with High Volume and Close-to-saturated Conditions**

| Link | Direction of Traffic | Pretimed | | | | Reinforcement Learning | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | # of vehicles | Delay (veh/s) | | | # of vehicles | Delay (veh/s) | | |
| | | | Average | Max | Min | | Average | Max | Min |
| Corridor 1 | EB | 3526 | 93.4 | 152.1 | 61.7 | 3856 | 91.0 | 114.6 | 76.3 |
| | WB | 3971 | 65.8 | 89.0 | 53.5 | 3910 | 43.8 | 74.3 | 21.0 |
| Corridor 2 | EB | 3938 | 22.8 | 35.9 | 14.4 | 4051 | 57.7 | 100.9 | 25.2 |
| | WB | 3916 | 58.1 | 75.8 | 46.5 | 4091 | 40.2 | 64.9 | 22.7 |
| Minor St. 1 | | 2376 | 102.7 | 155.2 | 71.3 | 3703 | 97.6 | 143.4 | 75.1 |
| Minor St. 2 | | 3305 | 54.7 | 74.8 | 41.3 | 3962 | 45.3 | 113.7 | 29.0 |
| Minor St. 3 | | 4093 | 25.4 | 39.7 | 14.7 | 3983 | 58.1 | 101.2 | 22.6 |
| **Totals** | | **21032** | **68.3** | | | **27556** | **61.4** | | |

**Table 3. Delay for the 3x3 Network with High Volume and Undersaturated Conditions**

| Link | Direction of Traffic | Pretimed | | | | Reinforcement Learning | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | # of vehicles | Delay (veh/s) | | | # of vehicles | Delay (veh/s) | | |
| | | | Average | Max | Min | | Average | Max | Min |
| Corridor 1 | EB | 3350 | 55.8 | 159.7 | 15.4 | 3439 | 57.2 | 132.5 | 18.1 |
| | WB | 3456 | 71.7 | 82.5 | 62.5 | 3420 | 32.9 | 100.2 | 13.7 |
| Corridor 2 | EB | 3495 | 24.9 | 43.6 | 16.7 | 3495 | 44.7 | 75.8 | 23.2 |
| | WB | 3564 | 73.1 | 90.6 | 56.5 | 3564 | 24.5 | 40.7 | 13.0 |
| Corridor 3 | EB | 3403 | 18.4 | 31.4 | 13.2 | 3385 | 22.6 | 36.6 | 12.7 |
| | WB | 3507 | 72.1 | 83.1 | 65.2 | 3486 | 23.9 | 36.9 | 14.1 |
| Minor St. 1 | | 3463 | 56.8 | 135.0 | 18.4 | 3333 | 52.5 | 86.4 | 20.1 |
| Minor St. 2 | | 3371 | 71.4 | 110.1 | 61.0 | 3438 | 30.0 | 61.9 | 16.5 |
| Minor St. 3 | | 3462 | 24.1 | 48.0 | 15.0 | 3411 | 42.9 | 85.5 | 17.5 |
| **Totals** | | **27609** | **44.2** | | | **27560** | **30.1** | | |

**Table 4. Delay for the 3x3 Network with High Volume and Close-to-saturated Conditions**

| Link | Direction of Traffic | Pretimed | | | | Reinforcement Learning | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | # of vehicles | Delay (veh/s) | | | # of vehicles | Delay (veh/s) | | |
| | | | ` | Max | Min | | Average | Max | Min |
| Corridor 1 | EB | 3023 | 141.9 | 314.1 | 71.2 | 3749 | 104.8 | 136.3 | 82.3 |
| | WB | 4027 | 93.6 | 126.3 | 74.6 | 3985 | 51.4 | 87.3 | 25.6 |
| Corridor 2 | EB | 4076 | 25.3 | 45.5 | 15.6 | 4074 | 73.1 | 100.4 | 25.9 |
| | WB | 4070 | 64.4 | 91.6 | 47.4 | 4180 | 42.4 | 66.6 | 27.6 |
| Corridor 3 | EB | 3953 | 19.9 | 36.0 | 14.3 | 3996 | 30.0 | 50.1 | 18.3 |
| | WB | 4072 | 72.0 | 90.8 | 57.9 | 4130 | 31.1 | 54.7 | 15.4 |
| Minor St. 1 | | 2990 | 145.3 | 396.5 | 70.7 | 3714 | 102.4 | 127.6 | 80.9 |
| Minor St. 2 | | 3952 | 97.4 | 280.6 | 62.8 | 4092 | 49.5 | 90.3 | 26.7 |
| Minor St. 3 | | 4088 | 26.5 | 43.8 | 16.4 | 4132 | 71.3 | 95.1 | 34.6 |
| **Totals** | | **30163** | **66.0** | | | **31920** | **51.9** | | |

This trend is more clearly described in Figure 2, where the average numbers of stops per vehicle for each link in the 3x3 network are shown. The effect of using information from neighboring intersections works in two ways: 1) as vehicles move toward the center of the network in the east-west direction, and 2) as vehicles move south in the minor streets (since they have one-way traffic only). The dashed arrow in Figure 2 illustrates the trend, which is also supported by the reduction in the number of stops. Similar trends were found using the delay values, as it can be observed from Table 4.

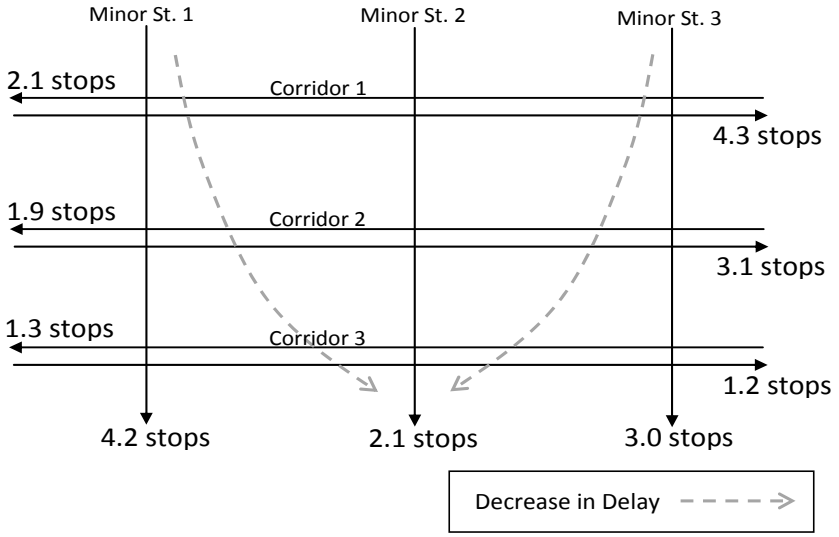



**Figure 2. Number of stops for the 3x3 network with volume close to saturation**

From the simulations, it was noted that at times the back of queue reached the advance detectors, limiting the ability of the agents to perceive the actual number of vehicles waiting for the green indication. This is critical mostly for the "outer"

intersections in the network (those located at the end of the corridors), as the agents do not have information from upstream neighbors regarding their current state. In these situations, the "inner" intersections would have information from neighbors, providing additional pressure for the green light when the queue spills may block the upstream signal.

## CONCLUSIONS AND FUTURE WORK

This paper presents an application of reinforcement learning agents to control the signals in traffic networks. The definitions of the state and the reward structures include information from neighboring intersections, making possible the eventual emergence of coordinated behaviors that may result in lower delays and more balanced signal timings for two-way roadways. This approach is different from previous works as the state and reward representations combine the use of penalties that discourage queue spills and incentives that encourage coordination when links are close to become saturated. In addition, since the information is stored in the form of Q-values, the knowledge of the agents can be accumulated over time and reused for future applications. The size of the state space does not grow exponentially with the number of intersections, making the procedure applicable to bigger and more realistic scenarios.

Results showed that reinforcement learning agents can cope with stable traffic conditions at high volumes and also for conditions close to the saturation point. For the 2x3 and 3x3 networks, improvements were obtained at all levels compared to pre-timed signals: lower average, lower maximum, and lower minimum delays. More balanced operation was also observed in two-way roadways compared to typical one-way coordination with pre-timed signals.

Using this reinforcement learning scheme, the level of coordination is expected to be higher as the vehicles move closer to the center of the network. This is particularly interesting for bigger networks since the level of organization may actually increase with the number of intersections. Further research testing different network sizes, volume variations, and adjusting the agents' parameters is underway.

## REFERENCES

Abdulhai, B., Pringle, R., and Karakoulas, G.J. (2003). Reinforcement Learning for True Adaptive Traffic Signal Control, ASCE Journal of Transportation Engineering, Volume 129, Number 3, pp. 278-285.

Bakker, B., Steingrover, M., Schouten, R., Nijhuis, E., Kester, L. (2005). Cooperative multi-agent reinforcement learning of traffic lights. In: Gama, J., Camacho, R.,Brazdil, P.B., Jorge, A.M., Torgo, L. (eds.) ECML 2005. LNCS (LNAI), vol. 3720.Springer, Heidelberg.

Bingham, E. (1998). Neurofuzzy traffic signal control. Master's thesis, Dept. of Engineering Physics and Mathematics, Helsinki Univ. of Technology, Helsinki, Finland

Bingham, E. (2001). Reinforcement learning in neurofuzzy traffic signal control, Eur. J. Operat. Res. 131 (2), pp. 232–241

Kuyer, L. Whiteson, S., Bakker, B. and Vlassis, N. (2008). Multiagent Reinforcement Learning for Urban Traffic Control using Coordination Graphs. In Proc. 19th European Conference on Machine Learning, Antwerp, Belgium

Oliveira, D., Bazzan, A. L. C., Lesser, V. (2005). Using Cooperative Mediation to Coordinate Traffic Lights: a Case Study. In: Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, Utrecht. New York : ACM, 2005. v. 1.

Oliveira, D., Ferreira JR., P. R., Bazzan, A. L. C., Kluegl, F. (2004) A Swarm-based Approach for Selection of Signal Plans in Urban Scenarios. In: IV International Workshop on Ant Colony Optimization and Swarm Intelligence Sutton, R.S., Barto, A.G. (1998). Reinforcement Learning: An Introduction. MIT Press (ANTS 2004), Brussels. Proc of the IV International Workshop on Ant Colony Optimization and Swarm Intelligence.

Thorpe, T. (1997). Vehicle Traffic Light Control Using SARSA, Masters Thesis, Department of Computer Science, Colorado State University

Watkins, C.J.C.H. (1989). Learning from delayed rewards. PhD Thesis, University of Cambridge, England

Watkins, C.I.C.H. and Dayan, P. (1992). Q-learning. Machine Learning, 8, 279-292

Wiering, M. (2000). Multi-Agent Reinforcement Learning for Traffic Light Control. In: Proc. 17th International Conf. on Machine Learning, pp. 1151–1158

Xie, Y. (2007) Development and evaluation of an arterial adaptive traffic signal control system using reinforcement learning, in Civil Engineering. Texas A&M University: College Station, TX.

Zhang, Y. Xie. Y., and Ye, Z. (2007). Development and Evaluation of a Multi-Agent Based Neuro-Fuzzy Arterial Traffic Signal Control System, Texas A&M University, 122 pp. (SWUTC #473700- 00092-1)

# Remote Data Collection for the IQA Delay Estimation Method

Y.M. Keita[1] and M. Saito[2]

[1]Department of Civil & Environmental Engineering, Brigham Young University, Provo, UT, USA; PH (801) 422-6326; FAX (801) 422-0159; e-mail: yayemalon@hotmail.com
[2]Department of Civil & Environmental Engineering, Brigham Young University, Provo, UT, USA; PH (801) 422-6326; FAX (801) 422-0159; e-mail: msaito@byu.edu

## ABSTRACT

One common way to measure the levels of service of signalized intersections is to estimate delays. The call for better methods for estimating delays is the source of much research that culminated in the Incremental Queue Accumulation (IQA) method. The IQA method consists of two procedures: IQA field and IQA model. Both procedures require intensive cycle-by-cycle data collection on approaching vehicles and traffic signal timing. In a study to evaluate the accuracy of the IQA method, the real-time remote data collection facility available at the Transportation Laboratory of Brigham Young University was used to create a digital video that simultaneously records the four approaches to a signalized intersection, together with signal changes for the four approaches. By using a video image analysis program that allowed frame-by-frame analysis, it was possible to collect both vehicle arrival/departure and signal timing data that were necessary for evaluating the two IQA delay analysis procedures.

## INTRODUCTION

Determining the level of service (LOS) of an intersection is an important task for traffic engineers. Estimations of intersection LOS are based on the values of the intersection delays. A long delay time indicates a low level of service. Several methods of estimating delays exist. Different organizations have developed methods for determining different kinds of delays. For example, the Transportation Research Board (TRB) defines control delays as the delay caused by a control device in its Highway Capacity Manual 2000 (HCM 2000) (TRB 2000). Also, the Institute of Transportation Engineers (ITE) defines stopped delays as the time a vehicle is stopped in queue while waiting to pass through the intersection in its Manual of Traffic Engineering Studies (ITE 1994). Furthermore, individual researchers in the traffic engineering field have also conducted their investigations and research to find better ways to estimate delay at intersections (Angel et al. 2002, Hereth et al. 2006, Kim 2006, Saito et al. 2008).

The standard equation for determining delay times at signalized intersections was the control delay estimation model introduced by TRB in 1994 (TRB 1994). The control delay estimation model in HCM 2000 uses three variables in its delay

equation: uniform delay, incremental delay, and delay caused by an initial queue. These variables are necessary for finding accurate delay times; however, the first variable (uniform delay) has the potential for error. The uniform delay variable is based on the assumption that the queue accumulation diagram always appears as a triangle. Although this assumption is correct for some basic cases, there are many cases where this variable appears differently. As of now, the cases that do not conform to the assumption that uniform delay is charted as a triangle are adjusted to fit into the assumption. This method requires researchers to ignore facts so that the equation can be used. Research to develop a method that addresses this problem has led to the development of the Incremental Queue Accumulation (IQA) method (Strong and Rouphail 2005; Strong et al. 2005). This method, because it is more precise, requires a large amount of vehicle and signal timing data. Traditional field data collection methods, which typically use a crew of two workers, are not adequate for using the IAQ method because signal timing data also need to be collected together with traffic data. Hence, tools that allow recording both traffic and signal timing simultaneously are needed.

In a study to evaluate the accuracy of the IQA method, the real-time remote data collection facility available at the Transportation Laboratory of Brigham Young University was used to collect all the necessary traffic and signal timing data for performing IQA delay analyses. This paper first presents the different types of data required to use the IQA method and then shows how traffic data are remotely collected and reduced to perform the IQA delay analyses.

## HCM 2000 MODEL VS. IQA METHOD

This section briefly summarizes delay modeling by HCM 2000 and IQA method. Details of other models and delay estimation procedures are discussed in Keita (2010).

### Highway Capacity Manual 2000 (HCM 2000) Delay Model

The HCM 2000 delay model is the latest version of HCM delay model. It is not considerably different from the 1997 version. The HCM 2000 delay model has three terms in the delay equations: uniform delay ($d_1$), incremental delay ($d_2$), and initial queue delay ($d_3$). As in the two preceding HCMs, the progression factor for the delay model is applied only to the uniform term; however, new to the 2000 model is the application of another progression factor ($PF_2$) that is used for determining back of queue to estimate initial queue delay (d3).

The limitations of HCM 1985-2000 delay models are related mainly to the first term of the model, which is the uniform delay term (TRB 1985, TRB 1994, TRB 1997, TRB 2000). This is because the calculation of uniform delay involves finding the area of the accumulation queue diagram, an area that researchers have assumed to be a triangle in HCM 1985-2000 delay models. For this to be true the researcher must assume three things: (1) there exists an unique triangle with only one red period and one green period, (2) the uniform arrival rate is represented by a single straight line on the leading edge of the queue accumulation diagram for the duration of the red

light interval, and (3) the difference between the uniform arrival rate and uniform saturation flow rate of departure is demonstrated by a single straight line on the falling edge of the queue accumulation diagram for the green light interval (Strong et al. 2005). Given that data do not always meet these assumptions, sub-models must be modified to make the data fit the equation. The queue accumulation diagram as a triangle is shown in Figure 1. For detailed discussions of complex cases, such as permitted left-turn and protected-permitted left-turn cases, refer to the HCM 2000 (TRB 2000).

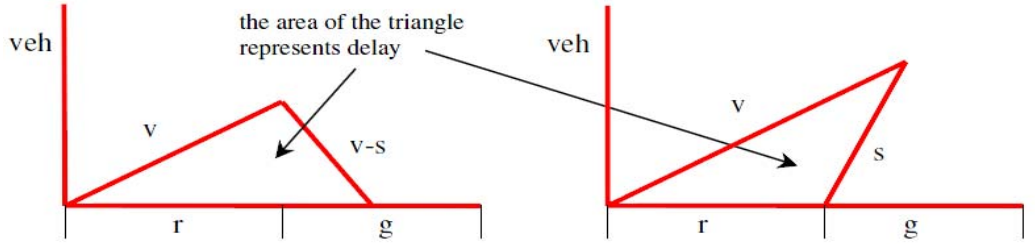



**Figure 1. Triangle shape of the queue accumulation and discharge function (TRB 2007).**

## IQA Delay Estimation Method

The concept of IQA is based on constant time increment; however, the IQA method does not need to have a fixed analysis interval. Variable analysis intervals can be used as long as they coincide with the change of inflow and outflow values. This is one example of where the flexibility of this new model is demonstrated. The queue accumulation diagram can be any shape because the non-constant saturation flow rate and arrival rate are both embedded in the method. To find the variable intervals, the analyst needs to determine intervals of the constant saturation flow rate and arrival rate. Each set of constant saturation flow rate and arrival rate data yields a trapezoid shape on the queue accumulation diagram, and the area of several trapezoids where inflow and outflow are constant are added in a cycle to find the total uniform delay of the cycle. Essentially, the IQA method for calculating uniform delay encompasses the current method for estimating delay in HCM 2000. That is, the current HCM method for estimating uniform delay can be derived from the IQA method.

**IQA Field Delay Measurement.** Estimating delay in the field using the IQA field procedure is different from determining delay using the IQA model (Kyte et al. 2009). The IQA field delay estimate appears to be more accurate than either the IQA model or the HCM 2000 delay model. One reason for that accuracy is that the IQA field procedure is able to correctly demonstrate the arrival patterns of vehicles better than the other two models can (Kyte et al. 2009). One may call the data obtained this procedure “ground truth” data because it requires the recording of arrival and departure times for all the vehicles that go through the study. The IQA field delay measurement is a counterpart of the field delay study mentioned in the ITE Manual of

Transportation Engineering Studies (ITE 1994). The data gathered here, however, are completely different from the ITE manual method. A sample data reduction table is shown later in Table 1.

**IQA Delay Model.** Delay diagrams of the IQA delay model are shown in Figure 2 (TRB 2007). To determine the area of each trapezoid, as shown in Figure 2, at first the queue at the beginning of the interval must be known and is in turn used to calculate the queue at the end of the interval. The queue at the start of the interval corresponds to the queue at the end of the previous one. If there is no previous interval, the queue at the start is zero. The formula to calculate the queue at the end of the interval is presented in Equation 1 (TRB 2007).

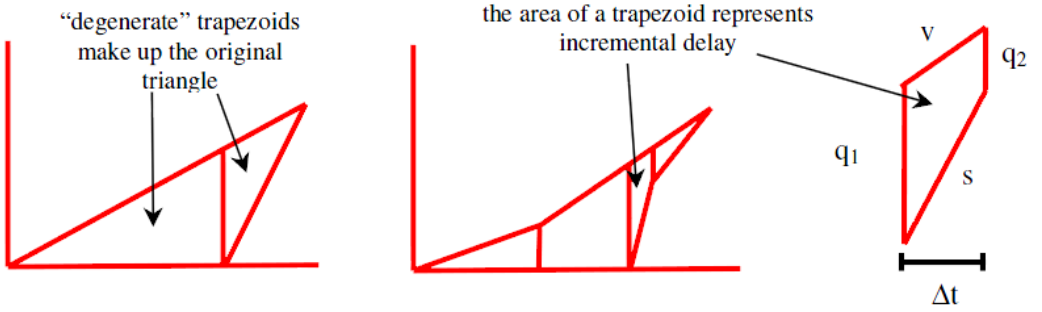



**Figure 2. Trapezoid shape of the queue accumulation and discharge function (TRB 2007).**

$$q_2 = q_1 + \frac{(v-s)}{3600} * \Delta t \geq 0 \quad \textbf{(1)}$$

where: $q_2$ = queue at the end of the interval (veh)
$q_1$ = queue at the start of the interval (veh)
$v$ = average arrival rate during the interval (veh/hr)
$s$ = average saturation flow rate during the interval (veh/hr)
$\Delta t$ = length of the interval (sec)

After finding all the information necessary to calculate $q_2$, the area of the trapezoid is calculated to find the uniform delay for that specific constant interval.

## DESCRIPTION OF REMOTE DATA COLLECTION

The BYU Transportation Lab has remote access to all traffic monitoring cameras monitoring highways and streets under the jurisdiction of the Utah Department of Transportation (UDOT). Some of these cameras are used for signal control. This feature, together with the digital recording capability of the system, can be used for collecting traffic data remotely for subsequent analysis. For this research, choosing a study site was an important task. Approaches to the selected study site needed to be clearly visible on the monitor in the BYU Transportation Lab in order to eliminate the need of extensive field data collection. The following sections outline

the problems and solutions encountered in selecting a study site. For more details on this topic, refer to Keita (2010).

## Site Selection Issues and Solutions

After examining several intersections that would possibly meet BYU Transportation Lab requirements for remote data collection, the intersection of University Parkway and Main Street in Orem, Utah, was selected. This intersection was chosen because a field video data collection trailer could be parked in a parking lot at a corner of the intersection, and its four approaches are simultaneously visible from a monitor in the Lab, where the images were digitally recorded for data reduction. Two video recordings—one recorded in the field by the data collection trailer and the other recorded in the BYU Transportation Lab—were used to supplement each other's data. The only significant problem with this site was that the signal timing information did not show the yellow interval in the monitor in the Transportation Lab; therefore, the yellow interval data had to be obtained separately.

To obtain the length of the yellow intervals for the chosen intersection, Orem traffic engineers were contacted. They said that they could not solve the issue because of the high cost involved in creating a simultaneous view of four approaches in one monitor, but they did provide the authors with a Synchro file containing the signal timing data of all signalized intersections on University Parkway within the City of Orem, which contained the lengths of the yellow intervals. The file indicated that the yellow intervals were 4.5 seconds for through movements and 3.5 seconds for turning movements for University Parkway approaches; and 4 seconds for turning movements and 3.5 seconds for turning movements for Main Street approaches. The yellow intervals were checked in the field and were found to be similar to the ones found in the Synchro file. To address the second challenge—the difficulty of coordinating the two video data sources—and in order to yield more accurate results, the approaches of University Parkway were excluded from the analysis and only the two approaches on Main Street were evaluated because only the ends of queues in the Main Street approach were visible in the monitor in the Transportation Lab.

## Data Collection

Because one goal of the field data collection was to serve as a supplement to the video data acquired in the Transportation Lab, different strategies were considered to avoid periods where queues would go beyond the view of the camera in the field data collection trailer (e.g., during peak hours, when it was certain that the end of the queue would not be visible in the lab).The field data collection was two hours long, and the day and hours chosen were Tuesday, July 28, 2009, from 11:00 AM to 1:00 PM.

## Video from the BYU Transportation Lab

Pictures of the lab monitors are shown in Figure 3 and Figure 4. The top right and bottom right quadrants of the main monitor show the Main Street approaches.

The main monitor was already connected to the digital recorder. The recording began at 9:00 AM because it was necessary to transport and set up the trailer before 11:00 AM. The video was recorded for six hours, from 9:00 AM to 3:00 PM. Very long queues were not visible during those six hours.

After recording, the video was divided into 15-minute segments so that the two hours from 11:00 AM to 1:00 PM were divided into eight 15-minute segments, each as a chapter in the recording. Following the division, the videos were transferred from the hard disk to the DVD. The DVD files were then converted to AVI files to be analyzed using the Delay Annotator software developed by Saito et al. (2008), which allows frame-by-frame analysis, as shown in Figure 5.

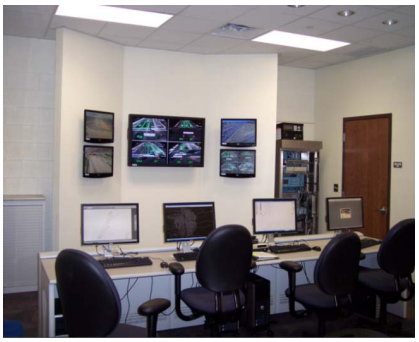


**Figure 3: BYU Transportation Lab monitors and computers.**

**Figure 4. Main monitor enlarged.**


**Figure 5. Screen shot of the image analysis software.**

## SAMPLE ANALYSIS RESULTS OF THE IQA METHOD

### IQA Field Delay Analysis

The IQA field delay analysis tracks individual vehicles passing through the intersection. In this study, four 15-minute videos were analyzed, with the analysis of each video encompassing two approaches, northbound and southbound. Unlike the HCM 2000 field delay analysis in which all the lanes of one approach are typically studied together, the IQA field analysis is always done lane by lane and cycle by

cycle; to find the delay for the entire approach, the weighted average of all the lanes in the approach is calculated.

The third cycle of the northbound second left turn lane from the middle of the road in Video 1 is used to demonstrate the procedure, as shown in Table 1. The same process is repeated for each cycle and each lane. A line perpendicular to the road under study, and parallel to conflicting traffic, and down the street from the stop bar was chosen as the lane beyond which a vehicle exited the intersection. The frame numbers are recorded at the time a vehicle stopped at the intersection (arrival), and then at the time the same vehicle crossed the reference line (departure). When the vehicle did not stop at the intersection, the frame numbers of when the vehicle got to the reference line and when the vehicle passed the reference line were recorded, indicating zero delay. When two or more vehicles came one after another and all stopped—that is, if multiple vehicles joined the existing queue—the frame numbers of each vehicle stopped were recorded; and when they crossed the reference line, the frame numbers of each vehicle crossing the reference line were recorded, as demonstrated in Table 1.

**Table 1. IQA Field Delay Computations (Cycle 3, 2nd Left Turn of NB, Video1).**

| Frame Numbers (1) | Clock Time, (sec) (2) | Arrival or Departure (3) | Time (Sec) (4) | # of Vehicle In (5) | # of Vehicle Out (6) | Incremental Queue (IQA) (7) | Incremental Delay (8)=(4)x(7) |
|---|---|---|---|---|---|---|---|
| 6497 | 216.6 | Arrival | 31.5 | 1 | | 1 | 31.5 |
| 7442 | 248.1 | Arrival | 6.7 | 1 | | 2 | 13.3 |
| 7642 | 254.7 | Arrival | 6.7 | 1 | | 3 | 20 |
| 7842 | 261.4 | Arrival | 42.2 | 1 | | 4 | 168.7 |
| 9107 | 303.6 | Departure | 6.2 | | 1 | 3 | 18.7 |
| 9294 | 309.8 | Departure | 4.1 | | 1 | 2 | 8.3 |
| 9418 | 313.9 | Departure | 1.0 | | 1 | 1 | 1 |
| 9448 | 314.9 | Departure | | | 1 | 0 | 0 |
| | | **Vehicle-seconds** | | | | | **261.5** |
| | | **Vehicles** | | | | | **4** |
| | | **Average delay (sec/veh)** | | | | | **65.4** |

All of the frame numbers were then converted to seconds according to a rate of 30 frames per second of film, and the arrival and departure times are specified in column 3 of Table 1. The time differences between adjacent arrivals and departures are shown in column 4. Column 5 indicates that one vehicle came in, and column 6 indicates that one vehicle went out. Column 7 shows the number of vehicles in the queue, and column 8 shows the incremental delay, which was calculated by multiplying the time difference between adjacent events and the incremental queue. The total delay was found by adding all of the incremental delays in column 8. The average delay for that cycle was calculated by dividing the total delay time by the total number of vehicles that entered the intersection or approached a lane during the cycle being studied.

## IQA Model Delay Analysis

The IQA model analysis is time consuming but not as time consuming as the IQA field analysis. This model analysis is data intensive and requires a lot of field data collection if data is not readily available. Table 2 shows the results of Cycle 1 of the IQA model analysis for the southbound right-turn lane of Video 1.

The same IQA model procedure was used for all the other cycles of the southbound right-turn lane of Video 1, after which, to get the delay time for the lane, the weighted average for all the cycles was calculated. To get the delay time for the southbound approach of Video 1, the weighted average of all the lanes included in that approach to the intersection was obtained. The same method was used for the northbound lanes and for all the other approaches in the other three videos. The input data were different for each cycle, each lane, each approach, and each video. As can be seen in Table 2, the input information was used to get the flow rate during the green and red intervals. The interval was then divided into green and red, and the duration of the green and red intervals was calculated. The saturation flow rates and the capacities during green and red intervals were also determined.

The second part of the analysis consists of breaking the green interval into blocked and unblocked intervals. The blocked interval indicates the portion of green interval when vehicles in the queue are clearing the intersection, the approach, and the lane. The unblocked portion is during the green interval but no vehicles are using the lane, approach, and intersection, because no queue of vehicles is being formed. For each interval— red, blocked, and unblocked—the duration of the interval ($\Delta t$), the queue at the start of the interval ($q_1$), the number of arrivals during the interval ($n_a$), the number of departures during the interval ($n_d$), the queue at the end of the interval ($q_2$), and delays ($d_i$) are calculated. Thereafter, the delay for the cycle is estimated as shown in Table 2.

## CONCLUSIONS

This paper described how a remote data collection facility was used to collect necessary data to perform the two delay analysis procedures of the IQA method: IQA field and IQA model. As described in the paper, the IQA delay analysis procedures can be more accurate than the HCM 2000 field method. For complex phases that are hard to model with the HCM 2000, such as permitted and protected-permitted phases, the IQA methods can help traffic engineers determine more accurate delays than the HCM 2000 methods. However, the IQA method requires a significant amount of data, as demonstrated in the sample IQA analyses. IQA analyses are performed lane by lane and cycle by cycle, a process that further complicates field data collection. Hence, it is practically impossible to collect the entire set of necessary data in one field visit by a couple of workers. To minimize field data collection efforts, it is recommended that traffic flow be video recorded for the extraction of all required data for the IQA delay analysis procedures. The remote data collection facility such as the one described in this paper can significantly help traffic engineers in performing IQA delay analyses.

**Table 2. Cycle 1 of IQA Model for Southbound and Right Turn Lane of Video 1.**

| | | | | |
|---|---|---|---|---|
| Yellow interval for movement, *y* (s) | 3.5 | | | |
| All red interval for movement , *ar* (s) | 1.5 | | | |
| Extension of effective green, *e* (s) | 4 | | | |
| Start up lost time, $l_1$ (s) | 2 | | | |
| Sum of yellow and all red, *Y*, (s) | 5 | | | |
| Clearance lost time, $l_2$ (s) | 1 | | | |
| Total lost time for movement (s) | 3 | | | |
| Actual green time, *G* (s) | 10.0 | | | |
| Effective green time, *g* (s) | 12.0 | | | |
| Effective red time, *r* (s) | 92.5 | | | |
| **Cycle1** | | | | |
| # of Vehicles in the cycle | 1 | | | |
| Volume, *V* (vph) | 34.4 | | | |
| Saturation flow rate, *S* (vph) | 1800 | | | |
| Cycle, *C* (sec) | 104.5 | | | |
| Effective green, *g* (sec) | 12.0 | | | |
| # of lanes, *n* | 1 | | | |
| Arrival Type, *AT* | 3 | | | |
| *Rp*= | 1 | | | |
| *P*= *Rp*x*g*/*C*= | 0.1 | | | |
| *Vg* | 34.4 | | | |
| *Vr* | 34.4 | | | |
| **Initial Interval Analysis:** | | | | |
| Interval # | 1 | 2 | | |
| Interval Description | red | green | | |
| $\Delta t$ (sec) | 92.5 | 12.0 | | |
| *v* (vph) | 34.4 | 34.4 | | |
| *s* (vph) | 0 | 1800 | | |
| *c* (vph) | 0 | 207.2 | X= | 0.2 |
| $v'$ (vph) | 34.4 | 34.4 | | |
| *v* (vpsec) | 0.01 | | | |
| *s*(vpsec) | 0.5 | | | |
| $t_c$ | 1.8 | | | |
| **IQA Computations:** | | | | |
| Interval # | 1 | 2 | | |
| Interval Description | red | Block | Unblocked | Total |
| $\Delta t'$ (sec) | 92.5 | 1.8 | 10.2 | 104.5 |
| $q_1$ (veh) | 0 | 0.9 | 0 | |
| $n_a$ (veh) | 0.9 | 0 | 0.1 | 1 |
| $n_d$ (veh) | 0 | 0.9 | 0.1 | 1 |
| $q_2$ (veh) | 0.9 | 0 | 0 | |
| $d_i$ (veh-sec) | 40.9 | 0.8 | 0 | 41.7 |
| | $d_1$= | 41.7 | sec/veh | |

## REFERENCES

Angel, A., Hickman, M.,Mirchandani, P. and Chandnani, D. (2002). Methods of Traffic Data Collection Using Aerial Video. *2002 IEEE 5th International Conference on Intelligent Transportation Systems*. IEEE, Singapore.

Hereth, W.R., Zundel, A. and Saito, M. (2006). Automated Estimation of Average Stopped Delay at Signalized Intersections Using Digitized Still-Image Analysis of Actual Traffic Flow. *ASCE Journal of Computing in Civil Engineering,* vol 20, Number 2, pp. 132-140.

Institute of Transportation Engineers (ITE). (1994). *Manual of Transportation Engineering Studies*, Prentice Hall, Englewood Cliffs, NJ. Chapter 5: Intersection and Driveway Studies, pp.69-87.

Kim, J. T. (2006). Estimation of Uniform Delay of Permitted Left Turns from Exclusive Turn Lanes. *ASCE Journal of Transportation Engineering,* vol 9, Number 708, pp.708-714.

Keita, Y. M. (2010). Evaluation of the Accuracy of the Incremental Queue Accumulation Delay Estimation Method. MS Thesis, Brigham Young University, Provo, UT.

Kyte, M., Abdel-Rahim, A., Dixon, M., Li, J.M. and Strong, D. (2009). Validating the Incremental Queue Accumulation Method for Left Turn Delay Estimation. *88th Annual Meeting Transportation Research Board 2009*, Transportation Research Board, Washington D.C.

Saito, M., Zundel, A.K., Taylor, C.N., Boyd, J. and Mendoza, M. (2008). Development of Automated Stopped Delay Estimation Software Using Video Image Processing Technology. *Report Submitted to Traficon USA LLC*.

Strong, D.W., Rouphail, N.M. and Courage, K. (2005). New Calculation Method for Existing and Extended HCM Delay Estimation Procedures. *85th Annual Meeting Transportation Research Board*, Transportation Research Board, Washington D.C.

Strong, D.W. and Rouphail, N.M. (2005). *Incorporating the Effects of Traffic Signal Progression* Into the Proposed Incremental Queue Accumulation (IQA) Method. *85th Annual Meeting Transportation Research Board*, Transportation Research Board, Washington D.C.

Transportation Research Board (TRB). (1985*). Highway Capacity Manual 1985* (HCM 1985). Transportation Research Board, National Academy of Science, Washington, D.C.

Transportation Research Board (TRB). (1994*). Highway Capacity Manual 1994* (HCM 1994). Transportation Research Board, National Academy of Science, Washington, D.C.

Transportation Research Board (TRB). (1997*). Highway Capacity Manual 1997* (HCM 1997). Transportation Research Board, National Academy of Science, Washington, D.C.

Transportation Research Board (TRB). (2000*). Capacity Manual 2000* (HCM 2000). Transportation Research Board, National Academy of Science, Washington, D.C.

Transportation Research Board (TRB). (2007*).Upcoming Highway Capacity Manual 2010 (HCM 2010)*. Restricted to Committee Members *Only*. Transportation Research Board, National Academy of Science, Washington, D.C.

# Evaluating Potential for Integrating Traffic Density with Guidance Information

Mostafa Reisi Gahrooei[1], Ryan Fries, Ph.D, P.E., M. ASCE[2], Huaguo Zhou, Ph.D., P.E.[3]

[1]Department of Civil Engineering, Southern Illinois University Edwardsville, Box 1800, Edwardsville, IL 62026;
[2]Department of Civil Engineering, Southern Illinois University Edwardsville, Box 1800, Edwardsville, IL 62026; PH (618) 650-5026; FAX (618) 650-5026; email: rfries@siue.edu
[3]Department of Civil Engineering, Southern Illinois University Edwardsville, Box 1800, Edwardsville, IL 62026; PH (618) 650-2815; FAX (618) 650-2815; email: hzhou@siue.edu

ABSTRACT

Although IntelliDrive has spurred significant research, little has been done to investigate the possibility of using traffic density information to time the release of in-vehicle guidance information. This study evaluated drivers' reactions to guidance information under varying traffic conditions to identify if such an integration of data is warranted. Researchers used a vehicle-driving simulator and participants between 18-25 years old, those with the highest US crash rate. The findings indicated that vehicle voice guidance software does not significantly distract drivers more than manually searching for road name signs. Results also suggest there is a direct relationship between the road segment level of service and the effectiveness of route guidance systems.

## INTRODUCTION AND BACKGROUND

Intelligent transportation systems (ITS) can improve the safety and efficiency of travel by using traffic engineering tools, computers, and software, hardware, and communication devices (Chowdhury, et al., 2003). The latest trend in ITS has been integrating roadside technologies with those inside vehicles, called IntelliDrive. Although the overarching purpose of ITS and IntelliDrive is to increase road safety, in-vehicle systems such as warning and guidance systems can lead to driver distraction and possibly crashes, particularly in urban areas. Recent research has focused on the operational abilities of such systems (Nowakowski, et al., 2008; Qi, et al., 2009; Qi, et al., 2009; Li, et al., 2009), suggesting that driver warning and guidance systems can effectively integrate with roadside infrastructure. Unfortunately, little is known about how much driver attention is required by these devices under different driving environments. Specifically, the amount of information and its frequency need to be considered in order to provide information to drivers without overloading them with information, thus preventing dangerous situations. This research study investigated the impact of in-vehicle route guidance systems on driver distraction to support a broader study on integrating these guidance systems with roadside infrastructure through the IntelliDrive concept.

## PREVIOUS WORK

Many studies have been conducted using driving simulators and can be categorized as either:

- Validating driving simulators
- Investigating drivers' behavior and
- Investigating innovative road design, pavement markings and signs.

### Validating Driving Simulators

Validation studies focus on either the physical representation of the driving environment or behavioral reactions of the drivers. The physical validation refers to the degree at which the visual and mechanical characteristics of the vehicle are realistically simulated. The behavioral validation refers to the similarity of driver's behavior and responses in simulated environment and real situation. Physical validation is a necessary condition for behavioral validation, thus it can be checked only if physical validation has been reached (Blana, 1996).

Studies of behavioral validity have examined both absolute and relative validity. An absolute validity study requires the same results between real-world and driving simulator experiments (Xuedong Yan, 2007). Tornros (1998) has declared that the absolute validation is not crucial for most behavioral research studies, yet relative validity is always necessary (Tornros, 1998).

Studies validating driving simulators commonly use speed and lateral position (within a travel lane) as measures of effectiveness. For example, Bella et al compared the speed of vehicles on existing two-lane rural roads near Rome, Italy, with the simulation of those roads. The results indicated that participants drove faster in driving simulator than on the actual road. Godly et al. used traverse rumble strips, horizontal curves and stop signs to validate speed. In this study, a group of drivers drove in six real courses where three of them contained rumble strips and others were control roads. Another group drove the same courses in a driving simulator. The comparison revealed that the relative validity can be obtained even if participants drive faster on one set of courses (Stuart T. Godley, 2002). To'Rnros et al. compared the speed and lateral position of motorists in a real tunnel and a simulated one, concluding that the driving speed was higher in simulated tunnel than real one (Tornros, 1998). Other validation studies have focused on simulating deceleration lanes (Bella, 2007), lane width impacts (Florence Rosey, 2009), route choice and speed (Nico A. Kaptein, 1996), work zone situations and devices (Bella, 2005; McAvoy, 2007) and traffic safety measures (Xuedong Yan, 2007).

### Investigating drivers' behavior

Several studies have used driving simulators to investigate the application of in-vehicle technologies on the driver's behavior. Particularly, several investigate the use of cell phones and its effects on drivers. Charlton et al. used a driving simulator to compare situations where a driver is conversing with a passenger versus using cell phone and a remote person who can see the driver. The findings suggested that the pace of the conversation is a key factor in avoiding hazardous effects of driver distraction. For example, passengers can pause the conversation while the driver makes a challenging maneuver while someone on the phone may not know to pause the conversation (Charlton, 2009).

**Investigating innovative road design, pavement markings and signs**

Research has also been conducted to investigate the effect of the road infrastructure on driver performance using driving simulators. For example, Anund et al. investigated the effect of different rumble strips on sleepy drivers (AnnaAnund, 2008) and Van der horst et al. investigated the effects of nearby lateral objects on drivers' speed choice and lateral position (Richard van der Horst, 2007).

In- vehicle technologies have been investigated by various researchers. Driel et al. researched the impact of a congestion assistant on driving behavior by using driving simulator. The congestion assistant provided information for drivers and controlled the pedals during congestion (C.J.G.van Driel, 2007). Researchers at Texas Southern University investigate the influence of different warning systems on driver behavior and traffic safety. They examined work zone warning, run off the road warning and lane change warning by manipulating driving simulator (Yi Qi, 2009).

Srinivasan et al. investigated in-vehicle navigation systems to identify the effect of guidance system attributes, traffic situtation, and driver charactrestic on driving performance. The study used four types of guidance systems: paper map, heads down electronic map, heads up display in combination with electronic map, and voice guidance in combination with electronic map (Raghavan Srinivasan, 1994). In other publications, these authors present methods of measuring the effect of in-vehicle guidance systems on driver workload and choice of vehicle speed (Raghavan Srinivasan, 1997).

METHODOLOGY

The primary research tools included a driving simulator and participant interviews. A driving simulator is a set of computer hardware and software that provides research subjects with a steering wheel, pedals, and 135-degree field of view while computer monitors display their driving progress through a predetermined test course. During the project, researchers first designed five courses, totaling approximately 30 minutes of driving time, including guidance information. The objective of the scenarios was to evaluate drivers ability to safely drive and comprehend directions under varying traffic and roadway geometric (i.e. curves) conditions.

Next, the researchers conducted a pilot study of younger drivers (18-25 years old) because this group has the highest traffic crash death rate (per mile driven) in the US (Insurance Institute for Highway Safety , 2006). Research subjects were solicited from the local student population and limited socioeconomic data was collected. During their participation, subjects were interviewed to identify their perceptions of the route guidance system after completing their driving test course.

Other questions helped identify basic demographic information of the drivers and the researchers also answered a few basic questions about the driver's performance to help explain any anomalies found during the simulated driving courses. All answers and driver performance data was associated with a number and the only record of participant names is on an informed consent sheet. Thus, there is no way to associate data to individual participants and the information remained anonymous.

After signing the consent letter and filling out the demographic questionnaire, drivers were asked to sit on the simulator seat and fix their location

to feel comfortable with the pedals and steering wheel. Next, the researcher introduced the driving simulator for the participant to make sure that they realized how to operate the controls. Then, divers were asked to drive five warm up scenarios each of about one minute long. These warm up scenarios were designed to help drivers become acclimated to the pedals, steering wheel, distances, and common traffic situations. During the warm up scenarios, no data was collected and the drivers were allowed to change their seat location, stop at any time, or ask questions.

After the warm up scenarios, drivers were asked to relax for approximately one minute; then the operator described the test and their task:

> *"The test is consisting of 5 different scenarios ranged from two to seven minutes. In each scenario you are supposed to turn at a specific road. In each scenario you might be provided with voice guidance that leads you to the destination road or you might not. In the latter case, you will be provided with the name of the street and the direction of the turn. Please note the speed limits and try to drive as you usually drive on real roads".*

To introduce randomness into the order of scenarios for each driver, they chose the name of the scenario by chance and a random number generator identified if the guidance would be voice or pre-disclosed destination.

During each run several measures of effectiveness were collected as shown in table 1. The most significant of these variables was the speed and lateral position, as suggested by the prevalence in previous driver simulator studies.

Table 1: Collected Measures of Effectiveness

| Time sec) | Distance (m) | Longitudinal and lateral acceleration($m/s^2$) | Longitudinal and lateral velocity (m/s) | Lateral position (m) | Road way and vehicle curvature (1/m) | Accident situation |
|---|---|---|---|---|---|---|

**Driving Simulator Scenarios**

The scenarios can be categorized based on their configurations or based on their traffic situation. Most of the scenarios contain road segments that have either two lanes or three lanes. These varying road segments help compare the effect of the number of lanes on driver's behavior. Additionally, the traffic situation varies throughout each scenario. The traffic situation was determined by the density of vehicles and measured via the Level of Services (LOS) criteria specified by the Highway Capacity Manual (from A to D). The significance of this metric was that drivers might be more distracted in denser traffic.

Overall, participants drove five road segments in this study. The findings from one representative scenario are presented herein.

**Street signs versus audio guidance system**

Each participant drove all types of scenarios and the method of providing directions and the direction of the turn was randomly selected, as previously discussed. For example, a participant might drive two of the scenarios equipped

by voice and others with only pre-disclosed directions. In each scenario drivers were asked to turn at one specific street (right or left).

When voice guidance was provided, drivers were not informed about the name of the street and the direction of the turn beforehand. Instead, the voice guides them at about 500 feet prior to the turn and their destination. On scenarios without voice guidance, drivers needed to find their destinations using the street signs, but were aware of the street name and direction of turn at the beginning of the scenario.

FINDINGS

The findings are broken into demographic data of the participant pool, analysis of average speeds, lateral lane positioning, speed variance, average acceleration, and number of lane changes.

**Participants demographic data**

Twenty-one people, mostly students, volunteered for participation in this study. At the beginning of the test, participants were asked to fill in a questionnaire to identify basic demographic information such as their age, driving experience and their driving background. Drivers' ages ranged between 18 to 25 with the weighted average of 23.7.

**Average Speed Analysis**

One of the most important variables evaluated was the change in speed due to guidance system distraction. During the study, driver speeds were collected every five feet for each scenario. Average speed can be considered as a good measure to show the trends in driving behavior for each group.

Figure 2 illustrates that the speeds were higher when the drivers were provided with voice guidance. The moving average line represents the average of the scatter plot for each scenario. In particular, participants are driving much faster when the level of service was B, indicating that light traffic and in-vehicle guidance could lead to faster speeds. Note that the horizontal axis indicates both the level of service (A-F) and the number of lanes of the road segment (two or three).

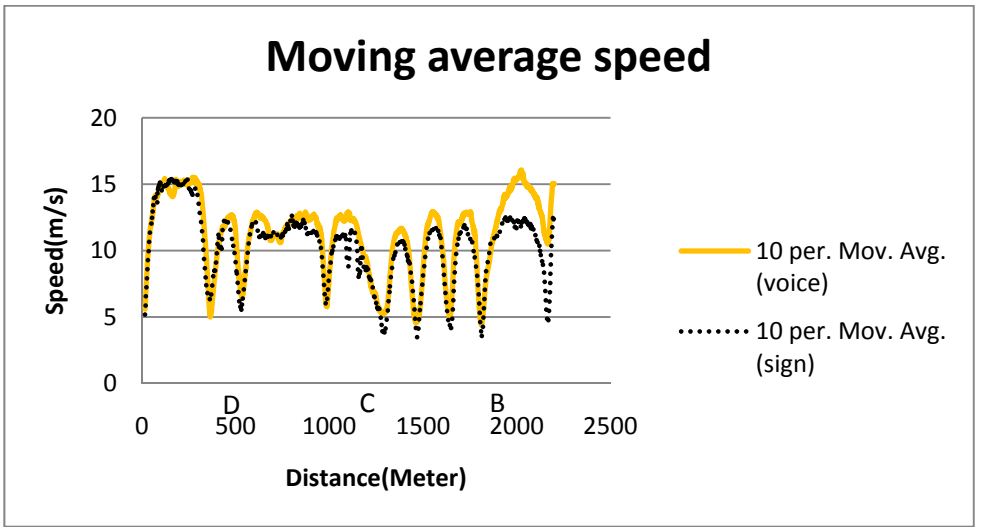



Figure 2: Average speeds between voice and sign guidance systems

**Lateral position analysis**

Based on collected data the lateral position of each driver has been found in each five-foot segment, similar to the speed analysis. Specifically, driver's path was compared to the centerline of the lane they are traveling in. In this analysis, it was assumed that a greater variance represents more driver distraction. Each scenario was analyzed for with and without voice guidance systems.

The variance illustrates divers' behavior similarity. High variance indicates a rapidly changing path and significant swerving. On the contrary, low lateral variance represents similar behavior throughout the scenario and indicates a focused driver. As figure 3 illustrates, the lane position variance in almost all scenarios was greater when drivers were not provided voice guidance, thus, these drivers were more distracted. Further, it is interesting to note that the participants using voice guidance drove very consistently throughout each scenario. Because the voice system provided information at the same locations, it encouraged a similar reaction from the participants. This consistency can provide uniform functionality and safer driving.

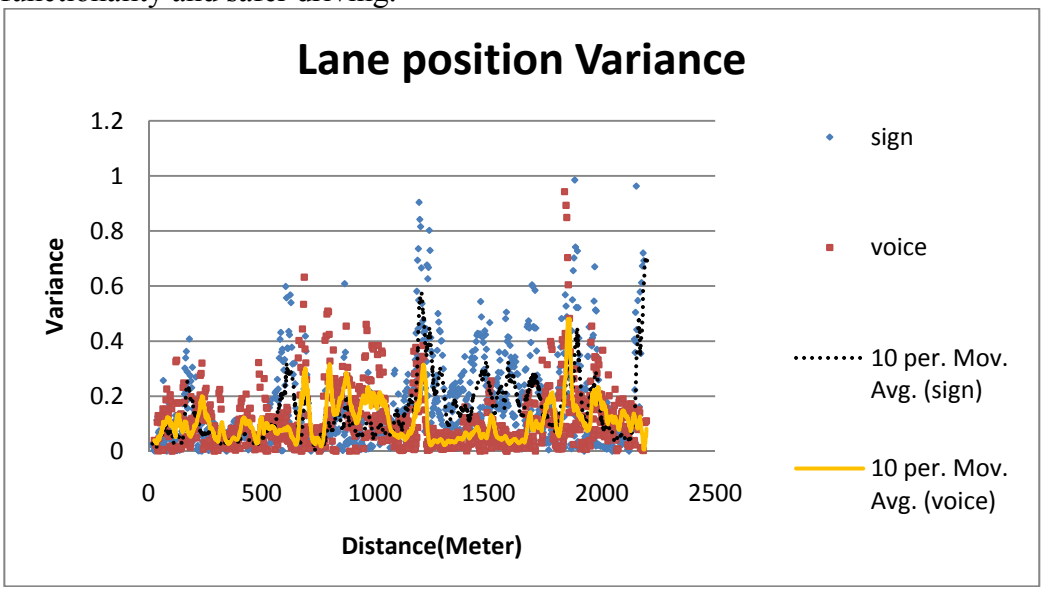



Figure 3: Lane position variance in scenario one

**Speed variances**

Researchers found speed variances in two different situations. First, for individual drivers every five feet (Figure 4) and then for every 100 foot segment (Figure 5). Figure 4 illustrates how differently the participating drivers behaved in each small portion of the road. The findings indicate that drivers behave more erratically while they are trying to read street signs to find their destination.

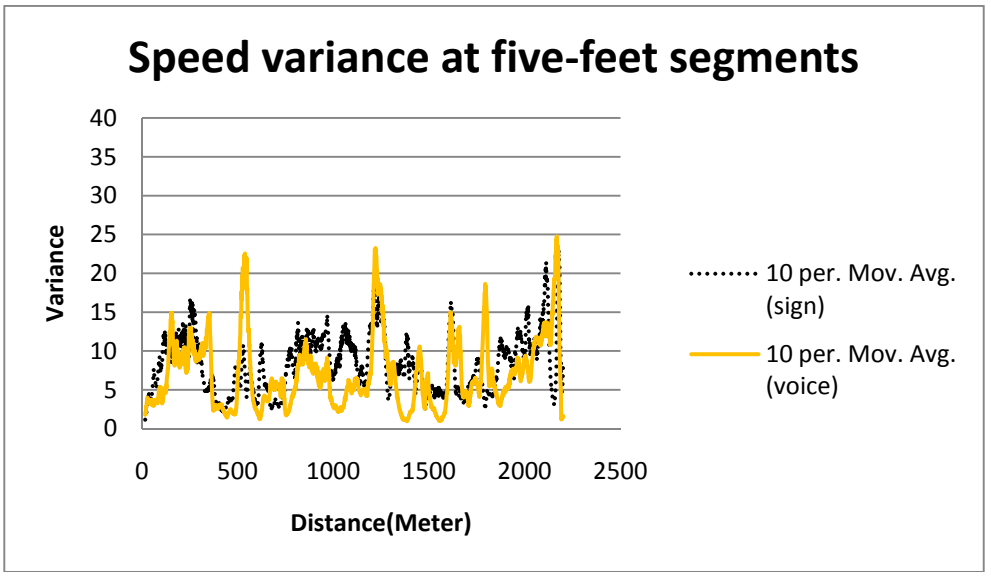



Figure 4: Speed variance in every five-foot segment for different drivers

Figure 5 depicts how individuals changed their speed throughout the scenario. For example each point indicates how much an individual driver changed their speed within a one-hundred-foot road segment. The figure displays more speed fluctuation for the drivers with no guidance system, indicating more stop and go driving while searching for street signs. Because the background traffic was controlled and the same for all participants, this fluctuation was due to driver breaking to read street name signs.

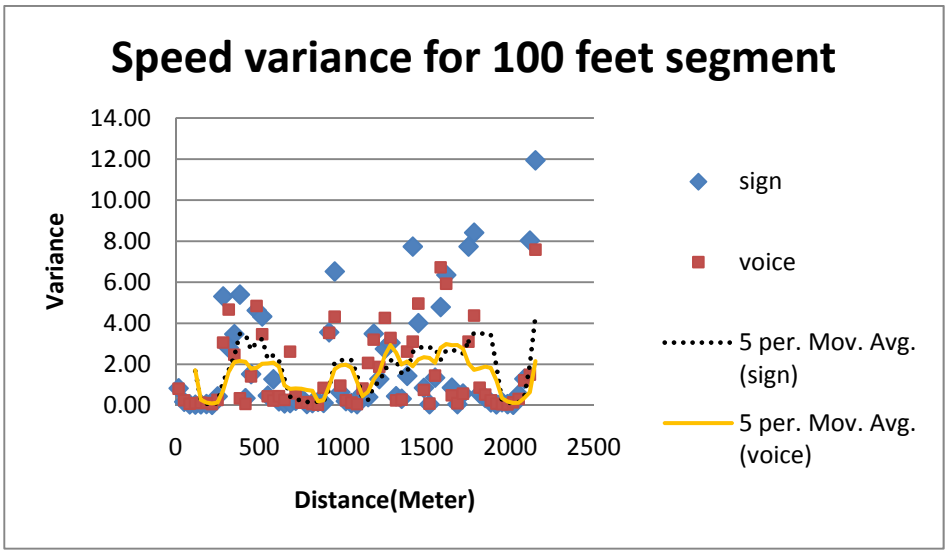



Figure 5: Average speed variance on every 100-foot segment

**Average Acceleration**

The acceleration parameter can show how aggressively drivers navigate the scenario. The average acceleration has been measured for both situations and compared in each 5-foot segment and is displayed in Figure 6. The results do not show any significant difference between the accelerating behaviors of drivers.

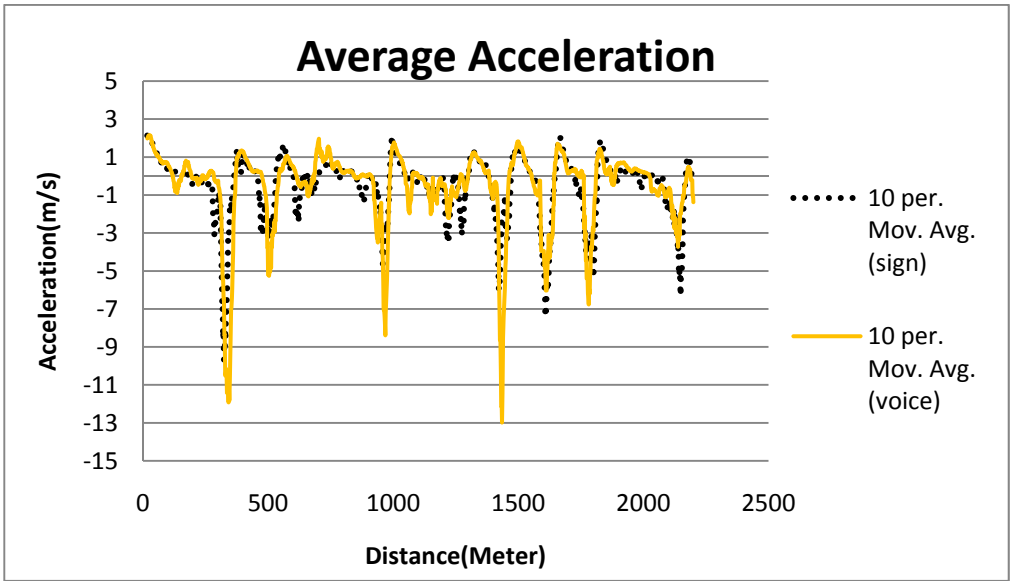



Figure 6: Average acceleration

**Number of the lane change**

Researchers believe that number of the lane change and car passing can probably show how confident a driver was during driving. Although the researcher observed that during the non-equipped scenarios drivers decide to stay on the left of right lane based on the direction of their turn during the scenario, an analysis of the average number of lane changes indicated that drivers changed lanes about five times for both situations.

## CONCLUSIONS

The findings of this pilot study indicate that in-vehicle voice guidance systems do not significantly distract drivers more than roadside guide signs. Experiment results from the driving simulator revealed that those drivers using voice navigation drove faster and with less variance, indicating they were more confident in their ability to find their destinations. Further, participants without any in-vehicle navigation system did not drive as straight within their lanes, as indicated by a higher lateral variance, compared to those using guidance software.

These findings indicate that there is significant potential for advancing in-vehicle guidance technology, particularly those integrating with the vehicle infrastructure integration concept. Results also suggest there is a direct relationship between the road segment level of service and the effectiveness of route guidance systems.

## ACKNOWLEDGEMENTS

This study was supported by the Institute for Urban Research.

## REFERENCES

AnnaAnund Goran Kechlund, Anna Vadeby, Magnus hjalmdahl, Torbjorn Akerstedt "The alerting effect of hitting a rumble strip- A simulator study with sleepy drivers". - Stockholm : ELSEVIER, 2008. - Vol. 40.

Bella Francesco Validation of a driving simulator for work zone design [Journal] // Transportation Research Record. - Rome : *Transportation research board*, 2005. - Vol. 1937.

Bella Francesco "Driving simulator for deceleration lane design" // Transportation research board. - Francesco Bella : *Transportation reseach board*, 2007.

Blana E "Driving simulator validation studies: A literature review". - Leeds : [s.n.], 1996.

C.J.G.van Driel M. Hoedemaeker, B.Van Arem "Impact of a congestion assistant on driving behaviour and acceptance using a driving simulator". - [s.l.] : *Tranportation research part F*, 2007. - 2 : Vol. 10.

Charlton Samuel G Driving while conversing: "Cell phones that distract and passengers who react" // *Accident Analysis & preventation*. - 2009. - pp. 160-173.

Chowdhury M A and Sadek A "Fundamentals of Intelligent Transportaiton Systems Planning" [*Book*]. - Norwood, MA : Artech House, 2003.

Florence Rosey Jean-Michel Auberlet, Olivier Moisan and Guy Dupre "Impact of narrower lane width comparision between fixed-base simulator and real data" // *Transportation research board*. - 2009. - pp. 112-119.

Garrett Weinberg Bret Harsham "Developing a Low-Cost Driving Simulator for the Evaluation of In-Vehicle Technologies" . - 2009.

Hakan Alm Lena Nilsson "The effects of a mobile telephone task on driver behavior in a car following situation" [Journal]. - Linkoping : *Pergamon*, 1995.

Insurance Institute for Highway Safety Fatality facts, older people [*Report*] : digital. - Arlington (VA) : "Insurance Institute for Highway Safety" , 2006.

Kristen E Beede Steven J. Kass "Engrossed in conversation: the impact of cell phones on simulate diriving performance" [Journal] // *Accident Analysis & prevention*. - 2006. - pp. 415-421.

Li K [et al.] "The Development of a Map-Based Driver Assistant under the Vehicle Infrastructure Integration Framework" [Conference] // *Transportation Research Board Annual Meeting* CD-ROM. - Washington, D.C. : National Academy of Science, 2009. - p. 14.

Linda Ng Boyle Gred mannering "Impact of traveler advisory systems in driver speed: some new evidance" [Journal]. - [s.l.] : *Transportation research part C,* 2004. - Vol. 12.

LLANERAS NEIL D. LERNER AND ROBERT E. "NCHRP report 488" [*Report*]. - Washington DC : Transportation Research Board, 2003.

McAvoy Schattler, Datta "driving simulator validation for nighttime construction work zone devices". - Washington DC : *Transportation research board*, 2007. - Vol. 2015.

Nico A. Kaptein Jan theeuwes, Richard Van Der Horst "Driving simulator validity: Some consideration". - Washington DC : *Transportation research board*, 1996. - Vol. 1550.

Nowakowski C, Cody D and O'Connel J "A Comparison of Infrastructure Versus In-Vehicle Driver Interfaces for Left-Turn Warnings" [Conference] // *Transportation Research Board Annual Meeting* CD-ROM. - Washington, D.C. : National Academy of Science, 2008. - p. 14.

Qi Y [et al.] 19 [Conference] // *Transportation Research Board Annual Meeting* CD-ROM. - Washginton, D.C. : National Academy of Science, 2009. - p.

"Vehicle Infrastructure Integration (VII) Based Highway Lane Change Warning System for Collision Prevention".

Qi Y [et al.]" Vehicle Infrastructure (VII)-Based Driver Warning System for Run-Off-Road Crash Prevention "[Conference] // *Transportation Research Board Annual Meeting* CD-ROM. - Washington, D.C. : National Academy of Science, 2009. - p. 16.

Raghavan Srinivasan Chun Zin Yang, Paul P Jovanis, Ryuichi Kitamura, Mohammad Anwar "Simulation study of driving performance with selected route guidance systems". - [s.l.] : *Transportation research part C*, 1994. - 2 : Vol. 2.

Raghavan Srinivasan Paul P. Jovanis "Effect of in-vehicle route guidance systems on driver workload and choice of vehicle speed: finding from a driving simulator experiment" [Book Section] // *Engonomics and safety of intelligent driver interfaces* / book auth. Noy Y. Ian.

Richard van der Horst Selma de Ridder "Influence of roadside infrastructure on driving behavior: Driving simulator study" . - Washington Dc : *Transportation research board*, 2007. - Vol. 2018.

Shiaw-Tsyr Uang Sheue-Ling Hwang "Effects on driving behavior of congestion information and of scale of in-vehicle navigation systems". - [s.l.] : *Transportation research part C*: Emerging Technologies, 2003. - 6 : Vol. 11.

Stuart T. Godley Thomas J. Triggs, Brian N. Fildes "Driving simulator validation for speed research" // *Accident Analysis and Prevention*. - 2002.

Tornros Jan Driving behaviour in a real and simulated road tunnel- A validation study [Journal]. - 1998.

Vincet Cantin Martin Lavalliere, Martin Simoneau, Normand Teasdale "Mental workload when driving in a simualtor:Effects of age and driving complexity" [Journal] // *Accident analysis and prevention*. - 2009.

Xuedong Yan Mohamed Abdel-Aty "Validating a driving simulator using surrogate safety measures" // *Accident Analysis & Preventation*. - 2007.

Xuedong Yan Mohamed abdel-Aty, Essam Radwan,Xuesong Wang, Praveen Chilakapati "Validating a driving sumulator using surrogate safety measures" [Journal] // *Accident Analysis & preventation*. - 2007.

Yi Qi Xin Chen, Lane Yang, Bin Wang and Lei Yu "Vehicle Infrastructure Integration (VII) Based Road-Condition Warning System for Highway Collision Prevention" [*Report*]. - Texas : [s.n.], 2009.

# Identifying Spatial Gaps in Transit Accessibility to Maximize Market Potential

**Srinivas S. Pulugurtha[1], Venkata Ramana Duddu[2] and Rakesh Mora[3]**

## ABSTRACT

The focus of this paper is (1) to explore features available in commercial Geographic Information Systems (GIS) software and estimate accessibility index as a function of potential captive riders for each transit stop, and, (2) identify spatial gaps based on accessibility index to provide improved public transportation systems that maximizes market potential. Potential captive riders are identified within a pre-defined accessible walking distance or walking time (0.25 miles or 5 minutes) from each transit stop in the City of Charlotte, North Carolina. These riders comprise unemployed, household population with 0 or 1 automobiles, population by age group, low income population (< $25,000), and ethnicity. Results from statistical analysis conducted considering transit ridership and estimated accessibility index show that market potential can be maximized by selecting transit stop locations with high accessibility index or in locations with high accessibility index. Therefore, the methodology and estimated accessibility index to identify spatial gaps helps expand transit area coverage, identify new routes and transit stops.

## INTRODUCTION

In recent years, proportional growth in traffic congestion has been a by-product of rapid growth in population. Several solutions were documented in the past literature to reduce huge economic and environmental losses associated with traffic congestion. Reducing congestion by providing an effective and efficient public transportation with maximum market potential is one such solution. The successful deployment of

[1] Associate Professor of Civil & Environmental Engineering, Assistant Director of Center for Transportation Policy Studies, The University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001, USA; Phone: (704) 687-6660, Fax: (704) 687-6953, E-mail: sspulugurtha@uncc.edu

[2] Graduate Student of Civil & Environmental Engineering, The University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001, USA; Phone: (704) 687-2913, Fax: (704) 687-6953, E-mail: vduddu@uncc.edu

[3] Graduate Student of Civil & Environmental Engineering, The University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001, USA; Phone: (704) 687-2913, Fax: (704) 687-6953, E-mail: rmora@uncc.edu

public transportation systems are often bound by factors such as service frequency, total travel time, fare, accessibility, comfort and convenience of travel.

Geographic Information System (GIS) software is widely being used by public transportation system planners in making decisions to improve transit system performance and achieve their objectives. The use of GIS in transportation planning has increased over the last two decades due to its exclusive features which simplify the process of conducting spatial analysis. It allows users to store, manage, analyze and display large databases in an effective and efficient manner.

This paper focuses on identifying spatial gaps in transit accessibility to enhance its market potential. Transit accessibility is defined as the ability for a rider to access public transportation or transit service in order to reach a destination within a pre-established or expected walking time, walking distance or comfort level. In general, a walking distance of 1,320 feet (0.25 miles) or walking time equal to 5 minutes (using 4 feet per second as an average walking speed) from an origin to a transit stop or transit stop to a destination is often considered to define accessible areas. In a near perfect situation, the minimum spacing between transit stops would be 0.5 miles in this case.

Factors that are used to measure accessibility are based on potential captive riders. These riders typically include unemployed population, population with 0 or 1 automobile (auto-ownership), population by age group, low income group (< $25,000), and population by ethnicity.

This paper illustrates the use of features available in GIS software to estimate accessibility for each transit stop. It is expressed as an index value and defined as a function of potential captive riders. The estimated accessibility index can be used to identify spatial gaps in transit accessibility to select ideal locations for transit stops along a route, identify new transit routes, and extend an existing route. This helps maximize transit market potential and ridership.

## LITERATURE REVIEW

Sanchez (1998) identified the connection between public transit and employment using GIS and two stage least squares regression method. The study also showed that characteristics such as gender and education can also have an effect on accessibility index. Pulugurtha et al. (1999) defined a measure for accessibility to transit service facility so as to maximize transit market potential. A GIS based methodology was used to identify locations for transit service facilities along various routes. An index for transit potential (measure for accessibility) was defined based on potential captive riders belonging to demographic variables such as age group, household income, ethnicity, household auto ownership, unemployment and physically handicapped persons.

Research by Ammons (2001) on transit stop spacing standards recommended that the range of stop spacing in urban areas is 656 feet to 1,968 feet. Saka (2001) developed an optimization model for determining the optimum spacing between the transit stops in urban areas.

Bhat et al. (2002) developed an urban accessibility index based on accessibility measures (such as spatial separation, cumulative opportunities, gravity measure, maximum utility and time space measures) and found varying results. Beimborn et al.

(2003) used accessibility and connectivity to identify potential captive riders. Murray (2003) developed a coverage model to improve public transit system accessibility by minimizing the number of stops and maximizing the proportion of population covered by the transit stops.

Lee (2004) recommended a parcel level measure of public transit accessibility to destinations using GIS. This method considered walk time, wait time and travel time using transit service and by walk. Kuby et al. (2004) developed a raster based algorithm for determining off-network routes to identify trip origins and destinations. Kimpel et al. (2007) used GIS to measure the effect of overlapping service areas on passenger boarding at transit stops. A distance decay function was used to calculate walking accessibility from dwelling units to bus stops.

### Limitations of Past Research

Past research focused on accessibility for either a single transit stop or stops along a transit route. Literature does not document studies to examine the relationship between ridership and estimated indices. This paper illustrates the working of a GIS based methodology (1) to estimate accessibility index for all the transit stops in the study area, and, (2) to examine the statistical relationship between ridership and estimated accessibility index. If a relationship exists, the estimates from the methodology can be used in decisions to identify new transit stops, new routes or the need for extension of existing route. These decisions help increase market potential by minimizing spatial gaps and inaccessible areas.

## METHODOLOGY

The GIS based methodology to estimate accessibility index comprises the following steps.

1. Selection of variables and weights
2. Spatial analysis
3. Data processing
4. Estimate accessibility index for each transit stop

Each of the above steps is discussed next in detail.

### Selection of Variables and Weights

Unemployed, population in households with 0 or 1 automobiles, population by age group, low income population (< $25,000) and population belonging to certain ethnic origins are the groups that typically are considered as potential captive riders. These groups of population are considered as variables to estimate accessibility index in this paper.

Different weights could be used in estimating accessibility index from these variables. Unemployed and low income group populations typically contribute to high ridership. Hence, these variables are given 30 percent weight each. These groups are followed by population with 0 or 1 per household. This variable is given a weight of 20 percent based on their expected contribution to ridership. Typical transit system

riders are in age group 15 years to 74 years. The numbers also depend on ethnicity. Hence, both these variables are given a weight of 10 percent each.

The 10 percent weight for age group is further divided into three parts - 20 percent, 70 percent and 10 percent for age groups 10-14, 15-64 and 65-74, respectively. Similarly, population by ethnicity is further divided into 50 percent, 30 percent, 5 percent, 5 percent and 10 percent for Africa-Americans, Hispanics, Caucasians, Asians and others, respectively.

## Spatial Analysis

The accessibility index in this paper is defined in terms of walking distance or time for a rider to/from a transit stop from/to an origin/destination. Representing this walking distance (0.25 miles) or walking time (5 minutes), buffers are generated around each transit stop in the transit system. As an example, Figure 1 shows two 0.25-mile buffers around two transit stops. Census data layer with 2008 estimates is then overlaid on the generated buffers to capture variables that help identify potential captive riders. The data are processed next to estimate accessibility for each transit stop.

## Data Processing

Spatial overlay of census data on generated buffers does not automatically adjust census data of the census block that falls within a buffer around a transit stop or a route. Data are further processed to estimate values pertaining to each selected variable. As an example, the total number of low income group population in a buffer around a transit stop is estimated using the following equation.

$$I_i = \sum_j \frac{A_{j,i}}{A_j} * I_j \qquad \text{.....Equation (1)}$$

where,

$I_i$ = Number of low income group population in the buffer "i" around transit stop "s",

$I_j$ = Number of low income group population in census block "j",

$A_{j,i}$ = Area of census block "j" in the buffer "i" around transit stop "s", and,

$A_j$ = Area of census block "j".

Similarly, equations are developed to extract population in households with 0 or 1 automobile, population by age group, unemployed population, and population by ethnicity for each transit stop.

## Estimate Accessibility Index for Each Transit Stop

As discussed before, 30 percent, 20 percent, 10 percent, 30 percent and 10 percent are used as weights for unemployment, auto-ownership, age group, low income and

ethnicity, respectively in this paper. The accessibility index for each transit stop "s" based on data extracted for each variable is calculated using the following equation.

$$A_s = 0.3 * U_s + 0.2 * AO_s + 0.1* AG_s + 0.3* I_s +0.1* E_s \quad \text{……. Equation (3)}$$

where,
$A_s$ = Accessibility index for transit stop "s",
$U_s$ = Total unemployed in buffer for transit stop "s",
$AO_s$ = Population with auto-ownership 0 or 1 in buffer for transit stop "s",
$AG_s$ = Age group between 15 to 74 in buffer for transit stop "s",
$I_s$ = Low income population (< $25,000) in buffer for transit stop "s", and,
$E_s$ = Ethnicity in buffer for transit stop "s".

## ANALYSIS & RESULTS

Data for the City of Charlotte, North Carolina transit network are used to illustrate the working of the methodology. The transit system in the City of Charlotte, North Carolina is operated by Charlotte Area Transit System (CATS). There are currently 3,662 transit stops and 80 routes in the study area. The recently opened light rail corridor (called as Lynx Blue Line) was not considered as the emphasis is more on bus operated transit system.

Accessibility index for all the 3,662 transit stops in the study area was calculated using the above discussed methodology.

### Accessibility Index for Transit Stops

Figure 2 shows spatial distribution of accessibility index by transit stop in the study area. The size of the point representing each transit stop is defined as a function of its accessibility index. One can clearly identify transit stops with high accessibility index from the figure.

The estimated average accessibility index value was 105 whereas the estimated maximum accessibility index value was 514. Table 1 summarizes estimated accessibility indices by range of all the transit stops. The accessibility ranges are divided based on average and standard deviation values.

It can be seen from Table 1 that 56.7 percent of transit stops have accessibility index values less than the average value accessibility index (~105). On the other hand, 4.8 percent of transit stops have accessibility index values greater than average plus 2 standard deviation of accessibility index (~236).

### Statistical Analysis

Ridership data for year 2008 was obtained for about 94 percent of the transit stops and all the routes the City of Charlotte Department of Transportation (CDOT). Statistical analysis was done to examine the relationship between accessibility index and ridership for transit stops. The total ridership was considered as the dependent variable while the estimated accessibility index was considered as the independent variable.

Table 2 summarizes results from statistical analysis. Statistical parameters (Table 2) obtained indicated that a statistically strong relationship exists between transit ridership and accessibility index for transit stops. The T-Statistic and P-Values are greater than 2 and less than 0.001 (>99 percent level of significance or confidence level), respectively for transit stops. The F-Statistic was high for the model i.e., greater than 4, showing that there exists a strong relationship between ridership and accessibility index estimated and used for analysis in this paper.

The coefficient obtained for accessibility index was positive indicating that ridership increases as accessibility index increases. In other words, one can say that having transit stops in locations with high accessibility index increases market potential and ridership.

## CONCLUSIONS

This paper presents a GIS based methodology to estimate accessibility index for transit stops. The working of the methodology was illustrated using the City of Charlotte, North Carolina transit network that comprises 3,622 transit stops. The average accessibility index for the transit stops in the study area was 105. Statistical analysis indicated a strong relationship between ridership and accessibility index estimated for analysis in this paper. This shows that having transit stops at locations with high accessibility index or in locations with high accessibility index increases ridership and market potential.

The GIS based methodology can be used to identify spatial gaps in transit accessibility so as to select new transit stop locations or relocate existing transit stops along a route. Spatial overlay of accessibility index for transit stops on area-wide demographic data will help extend an existing route or add new routes. Overall, the methodology developed and results obtained from this study could be used to enhance market potential and ridership.

Only demographic and socio-economic data were considered to estimate accessibility index in this paper. These data are typical indicators of weekday trip productions during morning peak hours and weekday trip attractions during evening peak hours. Developing accessibility index using land use characteristics will help account for weekday trip attractions during morning peak hours and weekday trip productions during evening peak hours. The estimation of accessibility index based on land use characteristics needs an investigation.

## REFERENCES

1. Ammons, D. N. (2001). “Municipal Benchmarks: Assessing Local Performance and Establishing Community Standards.” 2nd edition, Thousand Oaks, Sage Publications, California.
2. Beimborn, E. A., Greenwald, M. J., and Xia, J. (2003). “Accessibility, Connectivity and Captivity Impacts on Transit Choice.” Center for Urban Transportation Studies, University of Wisconsin, Milwaukee, Wisconsin.
3. Bhat, C., Handy, S., Kockelman, K., Mahmassani, H., Weston, L., Gopal, A., and Srour, I. (2002). “Development of Urban Accessibility Index: A

Summary." Center for Transportation Research, The University of Texas at Austin, Texas.
4. Kimpel, T. J., Ducker, K. J., and El- Geneidy, A. M. (2007). "Using GIS to Measure the Effect of Overlapping Service Areas On Passenger Boarding at Bus Stops." *Urban and Regional Information Systems Association Journal*, Vol. 19, No.1.
5. Kuby, M., Upchurch. C., Zoldak. M., and Barranda, A. (2004). "Using GIS To Generate Mutually Exclusive Service Areas Linking Travel On and Off a Network." *Journal of Transport Geography*, Vol. 12, 23-33.
6. Lee, B. H. Y. (2004). "Parcel Level Measure of Public Transit Accessibility." TransNow Student Conference, Portland State University, Portland, Oregon.
7. Murray, A. T. (2003). "A Coverage Model for Improving Public Transit System Accessibility and Expanding Access." *Annals of Operations Research 123*, Netherlands, 143–156.
8. Pulugurtha, S. S., Nambisan, S. S., and Srinivasan, N. (1999). "Evaluating Market Transit Potential and Selecting Locations of Transit Service Facilities Using GIS." *Journal of Public Transportation*, Vol. 2, No. 4, 75-94.
9. Saka, A. A. (2001). "Model for Determining Optimum Bus-Stop Spacing in Urban Areas." *Journal of Transportation Engineering*, Vol. 127, 195–199.
10. Sanchez, T. W. (1998). "The Connection between Public Transit and Employment." Center for Urban Studies, Portland State University, Portland, Oregon.

## ACKNOWLEDGMENTS

The authors acknowledge the staff of the City of Charlotte Department of Transportation (CDOT) and Charlotte Area Transit Systems (CATS) for their help with transit network and ridership data.

## DISCLAIMER

The contents of this paper reflect the views of the author(s) and not necessarily the views of the University, CDOT or CATS. The author(s) are responsible for the facts and the accuracy of the data presented herein.

**TABLE 1 Summary of Accessibility Index by Transit Stops**

| Accessibility Range | No. of Transit Stops | Percent of Transit Stops |
|---|---|---|
| 0 | 17 | 0.5 |
| 1 to 105.2 | 2,076 | 56.7 |
| 105.2 to 236.4 | 1,393 | 38.0 |
| > 236.4 | 176 | 4.8 |

**TABLE 2 Statistical Analysis – Ridership vs. Accessibility**

| Independent Variable | Coefficient | T- Stat | P - Value | F - Stat |
|---|---|---|---|---|
| Accessbility | 0.16 | 22.74 | < 0.001 | 516.96 |

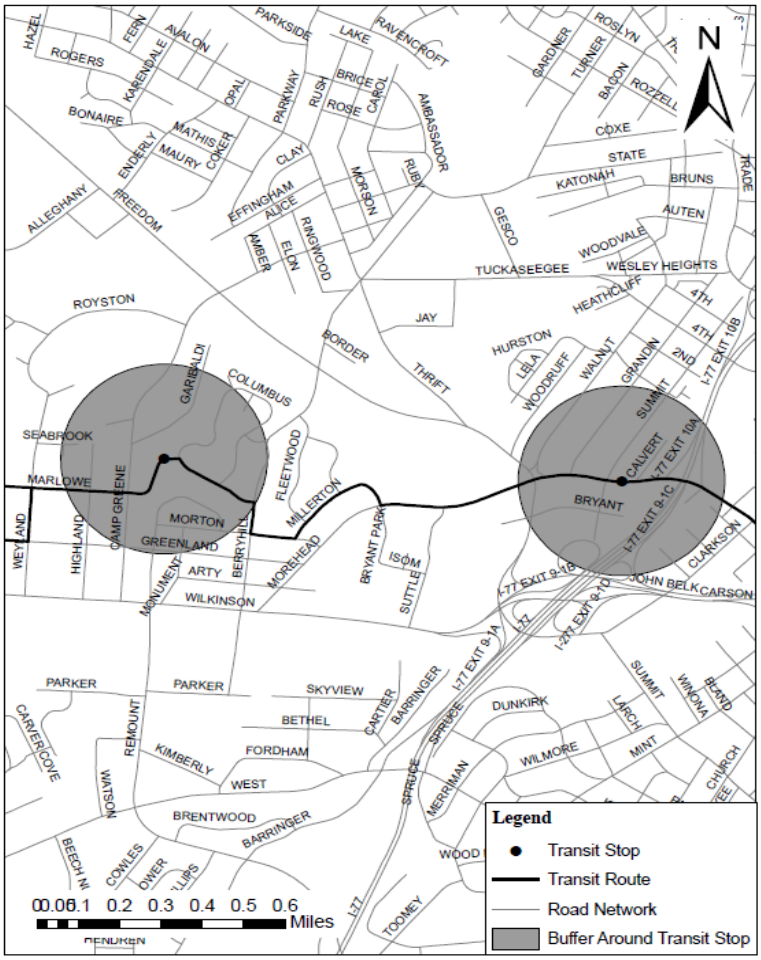



**FIGURE 1 Buffers Around Transit Stops.**

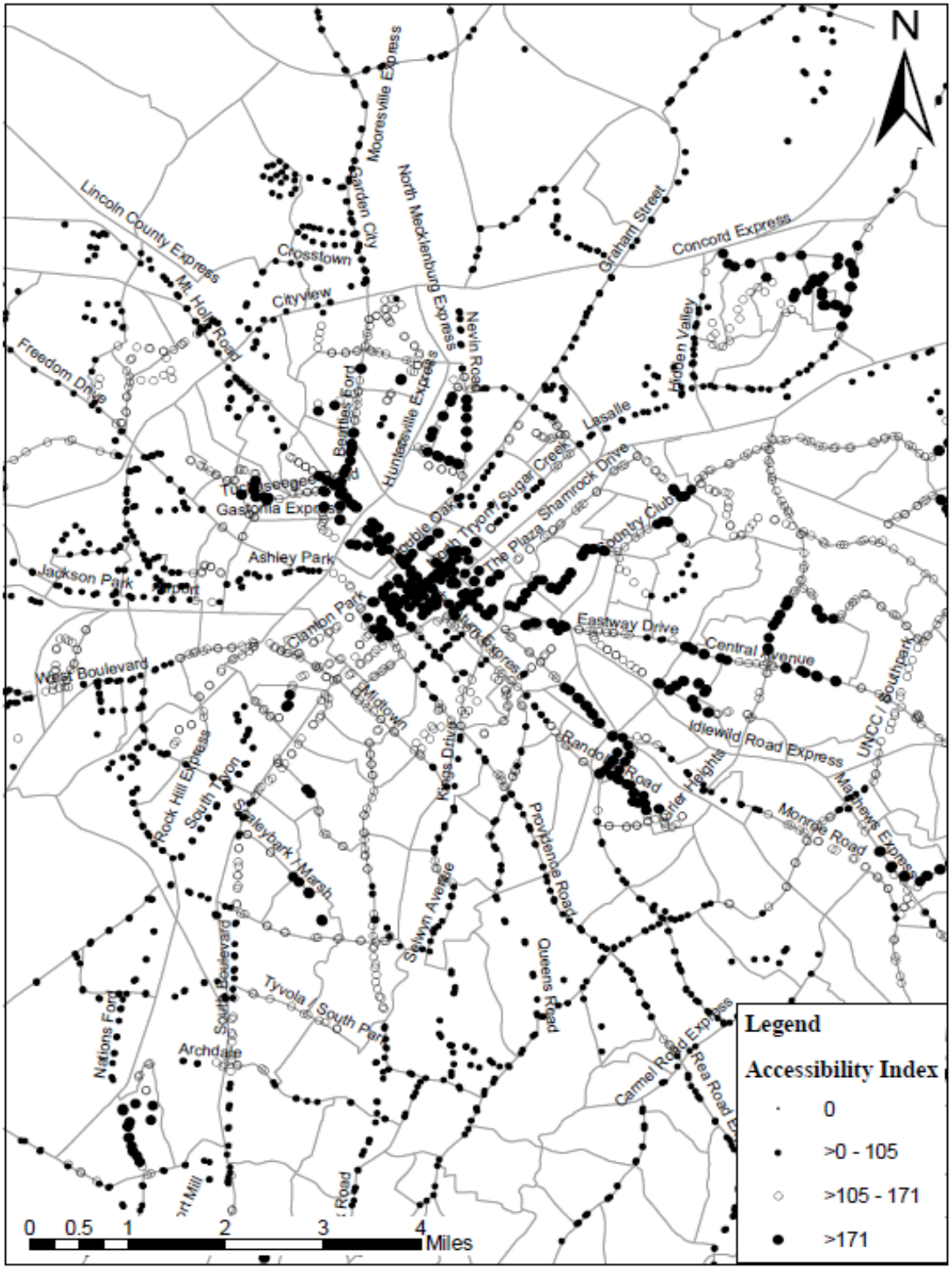



**FIGURE 2 Spatial Distribution of Accessibility Index for Transit Stops.**

# INVESTIGATE THE EFFECTIVENESS OF GRAPHIC ROUTE INFORMATION PANELS THROUGH A SURVEY STUDY

Ricardo Jesus Aitken[1], Guohui Zhang[2*], C. Michael Walton[3], and Alison Conway[4]

[1]. Keith and Schnars, P.A. 901 Northpoint Parkway, Suite 103, West Palm Beach, Florida 33407, Tel: (561) 682 -1095, Email: ricardo_aitken@yahoo.com

[2*]. Corresponding Author, Center for Transportation Research, Department of Civil, Architectural and Environmental Engineering, University of Texas at Austin, 3208 Red River, Austin, TX 78705, Tel: (512) 785-2296, Email: guohui@mail.utexas.edu

[3]. Department of Civil, Architectural and Environmental Engineering, University of Texas at Austin, Austin, TX 78712, Tel: (512)-471-1414, Email: cmwalton@mail.utexas.edu

[4]. Department of Civil Engineering, City College of New York, NY, Email: alicon@mail.utexas.edu

## ABSTRACT

Advanced Traveler Information Systems (ATIS) provide a vital platform to assist travelers in planning more informed trips on route selection and diversion for congestion avoidance and safety enhancement. As one major information dissemination mode, Dynamic Message Sign (DMS) can provide real-time en-route information including roadway congestion situations, construction activities, incident events, and diversion needs. Compared with traditional text-only messages, graphic-aided messages can be better recognized and understood. Although text-based DMSs have been in operation for many years, research on graphic-based DMSs is still in its early stages. To enhance our understanding of critical issues on designing suitable graphic-based DMSs, this study concentrates on design of an innovative Graphic Route Information Panel (GRIP), and proposes a methodology for understanding its potential usefulness to drivers. An online survey is employed to examine sign rates of understanding by and usefulness to drivers. In addition to basic statistical analysis of the survey results, a Logit model is employed to investigate the significance of various drivers' characteristics and design attributes impacting the effectiveness of GRIPs.

Key words: Advanced traveler information system, graphic variable message sign, legibility and recognition, and traffic survey.

## INTRODUCTION

In the last two decades, dramatically increasing travel demands and insufficient traffic facility supplies have induced severe traffic congestion. Traffic congestion costs billions of dollars every year, including lost time, increased risk of traffic accidents, wasted energy, excess air pollution, and lost productivity. From 1980 to 2005, annual vehicle miles traveled increased by 96%, while roadway lane

miles increased by only about 4% nationally *(1)*. According to the 2007 Annual Urban Mobility Report *(2)*, in 2005, the annual average delay per capita was 38 hours for 437 surveyed urban areas, a 171% increase from 1982. The total travel delay of 4.2 billion vehicle hours generated 2.9 billion gallons of excess fuel consumption and $78 billion in excess cost due to congestion. It is of practical importance to manage traffic system operations more efficiently using advanced technologies and management strategies, particularly in metropolitan areas. Advanced Traveler Information Systems (ATIS), as one important component of Intelligent Transportation Systems (ITS) provide travelers with dynamic traffic information to assist their trip preparation, and alternative routing diversion and guidance for congestion avoidance and safety enhancement. One major ATIS information dissemination mode, the Dynamic Message Sign (DMS) has been widely deployed to provide travelers with real-time and reliable en-route information. DMSs enable travelers to make better route decisions leading to better utilization of the road network. Normally, the message signs are mounted on overhead sign bridges and display extensive information including roadway congestion situations, construction activities, incident events, and diversion needs. Although the message signs may contain different contents in various formats, DMSs are roughly classified into two categories: text-only message signs, and graphic-aided message signs. Traditional text-only DMSs have been in operation for many years since 1961 when a section on signs for freeways and expressways was added to the *Manual on Uniform Traffic Control Devices* (MUTCD) *(3)*. However, these text-only message signs are limited in display capacities. Physically, a DMS is usually restricted to 3 lines of up to 20 alphanumeric characters. Taking into consideration that regulations do not allow more than two displays within a message cycle and that the entire message cycle should be readable at least twice by drivers traveling at the posted speed *(4)*, these constraints do not always allow all relevant information regarding a traffic situation to be effectively presented. As a result, posting a short but effective message becomes a challenge *(5)*. Additionally, text-only DMSs are less effective in sign legibility and recognition, and cannot satisfy customized routing information requests *(6)*.

Graphic-aided DMSs can enhance the effectiveness of traditional text-only DMSs in delivering information. The use of graphics over text in presenting information allows the sign to convey more information in a limited space and to be processed faster than written language. This is particularly important since under normal driving conditions, drivers only have 4-6 seconds to interpret sign information *(5)*. Although the MUTCD provides guidance on the design and use of diagrammatic signs *(3, 7)*, only a few studies were conducted to investigate the understandability and legibility of graphic-aided DMSs. Furthermore, the complexity of the roadway environments and the traffic patterns augments the difficulty in designing and deploying suitable message signs with high-quality readability. Inappropriate design of message signs in content, format, and degree of complexity can adversely affect drivers' comprehension *(8)*. It is important to ensure that graphic-aided message signs can be recognized and understood by travelers and that the corresponding responses can be made in a reasonably short time period before these signs are permanently mounted on traffic facilities. Therefore, this research concentrates on designing an innovative Graphic Route Information Panel (GRIP) and investigating its attributes

and drivers' preferences under various test scenarios through traffic survey studies. This paper is organized as follows. The next section briefly describes the state of the art regarding DMS design and implementation. It is followed by the detailed description of design principles for the proposed prototype GRIPs. Then the experimental tests for GRIPs' performance evaluations and travelers' comprehension examination are presented, including test scenario design, traffic survey study, statistical analysis, discrete choice model development and result discussions. The final section concludes this research effort and proposes further research topics.

## STATE OF THE ART

Many studies were conducted to investigate the effectiveness of traditional text-only DMSs. In Benson's study *(9)*, a survey was conducted to investigate the impacts of DMSs on drivers. It found that about 50% of the participants often responded to DMSs while 38% occasionally or sometimes responded to them. Chatterjee *(10)* employed surveys to investigate drivers' response to DMSs in London. Most of the participants indicated that DMS information could be very useful to them. In a similar study conducted by Peng et.al. *(11)*, 62% of the drivers indicated that they responded to DMS messages more than once per week and 66% of them changed their route at least once per month. An empirical survey analysis conducted by Emmerink et. al. *(12)* indicated that over 70% of the drivers were sometimes influenced by DMS information. Based on video-based driving simulation experiments on DMSs, Wang and Cao *(13)* found that discretely displayed messages demanded less response time than sequentially displayed messages and that single line messages were better than multiple-line messages. Dudek and Ullman *(14, 15)* studied message displays on DMSs and suggested that one-frame DMS messages should not be flashed and a line on a two-frame DMS message should not be alternated while other lines are kept the same. The MUTCD *(3)* has established a standard for DMSs that limits the number of frames per message to two and recommended that the message be in capital letters with a clearly legible letter size. Wang et al. *(16)* found that the best settings in regard to drivers' preference and response time are to display messages in amber or amber–green color combination. Wardman, Bonsall, and Shires *(17)* and Peeta *(18)* suggested that the impact of VMS information on travelers' route choices depended upon the message content, local circumstances, and motorists' characteristics. Although DMSs are capable of presenting information in various formats, only limited display formats are recommended by researchers.

Because the capacity of human visual receptivity considerably exceeds that of verbal receptivity, graphic-aided DMS signs have many advantages over text-only DMS signs. A graphic-aided message could be identified more easily, more quickly, and from a further distance; seen better under adverse viewing conditions; and understood better by people who cannot understand the language in the text *(19)*. In many European countries such as Britain, the Netherlands, Italy, Spain and France, graphical traffic information has been used on DMSs to influence drivers' route choices *(20)*. Driving simulation experiments were used to study graphical information displays, and no potential safety issues were found. The Conference of European Directors of Roads (CEDR) report also recommended that graphics and symbols should be used as much as possible to avoid the problem of disseminating

information to drivers who speak and use different languages *(21,22)*. Previous studies found that graphically presented information allowed faster responses *(23, 24)*. Bruce, et al. (*25*) found that driver recognition time for text messages was appreciably longer than symbolic messages. Alkim et al. *(5)* tested and compared drivers' comprehension of both regular text-based DMSs and GRIPs and found that drivers exhibited better route choice behavior with GRIPs than with regular DMSs. Wang et. al. *(8)* conducted a study on adding graphics to DMSs and found that most drivers preferred graphics over text. They concluded that graphics might help enhance drivers' understanding of and responses to those messages.

## GRIP DESIGN

These existing studies on graphic-aided message signs provide valuable insight for the proposed GRIP design and performance investigations. Based on the Texas MUTCD standards *(4)* as well as the lessons learned from the previous projects, a prototype GRIP was designed. The proposed GRIP is a message sign that provides travelers with real time, color coded traffic condition information in a graphical format. The basic panel consists of two components: a static component and a dynamic component. The static component usually consists of a reflective sheet containing the road network outline, landmark or road names, and a directional arrow. The dynamic component consists of a two or three color light emitting diode (LED) array which lights up to display the congestion levels in the roadway sections. Dynamic traffic information, such as average speed, travel time, incident location and lane closure could be displayed. As emphasized before, although GRIPs can convey more complex information to facilitate customized routing decisions, they should be carefully designed and tested to avoid information overload and driver distraction, especially for the freeway corridors with high speed limits and high demands. Since there are no specific MUTCD guidelines available to satisfy the GRIPs' unique combination of static and dynamic components, based on the operational purpose and classification, the prototype GRIP is designed by following the guidelines provided in TMUTCD 2A.05, 2E.19, and 2E.32. Each GRIP component follows the guidelines best suited for it, as long as they support the main function of GRIPs, which is to provide travelers with clear and concise information for safe and efficient trips. The design principles for different GRIP components are detailed in the following section.

### *Static Component*

As previously mentioned, the GRIP static component consists of a reflective sheet background containing a directional arrow, road names, route numbers, and a road network outline. The specific design considerations for each element are detailed below:

**Background Color.** A variety of different colors have been used for GRIP backgrounds. The colors have ranged from yellow and blue in the Netherlands, to white in Germany, to green in Japan, to green and blue in China. In the U.S. however, according to the MUTCD, there are 13 colors that have been identified as being appropriate for use in conveying traffic control information. Since GRIPs are to be considered guide signs, based on TMUTCD 2E.04 the background color for GRIPs should be green.

**Lettering.** According to the TMUTCD standards, guide signs shall have white lettering. All destinations on guide signs have the option of being composed of lowercase letters with initial uppercase letters or being all uppercase. The font, word and lettering spacing should follow the TMUTCD 2E.13 standards.

**Sign Orientation and Panel Dimension.** Regarding the roadway network orientation, there are two options: 1) display the network with a true north orientation, and 2) re-orientate the network in the direction the driver is traveling. Many studies concentrated on this issue *(20)*. The proposed GRIP adopts the first option, which is consistent with daily experience when using a map. The panel dimension is designed by following the Standard Highway Sign Designs Manual recommended by the TMUCD. Other factors including operational and maintenance costs, specific environments and sight distance also should be taken into account.

**Road Network Outline.** On a GRIP, the road network may be portrayed as realistically as possible or in a more stylized manner. If realistically portrayed, the network might be more recognizable but have limited display capacities. The design principle used in this study is to portray a stylized network that is as simple as possible to enable fast comprehension, but realistic enough to be recognizable. Also, because the adequate sight distance is crucial to the success of GRIPs, a suitable resolution for the network outline is another consideration factor.

Additionally, other GRIP attributes, such as retroreflection and illumination are considered and designed by following the TMUTCD standards.

*Dynamic Component*

The dynamic component consists of three color LEDs on the GRIPs. Black, orange, and red are used to symbolize different traffic flow conditions: free flow, crowded, and congested. The roadway network is divided into segments which are the smallest units used for representing the entire network. The smaller the number of colors used, the greater the reduction of complexity and increase in efficiency of information processing *(20)*. Dynamic traffic information, such as travel time indication, can be displayed. More descriptions of drivers' preferences are detailed in the next section.

Following the proposed design principles, a diagrammatic sketch for the proposed prototype GRIP is illustrated in Figure 1.

## EXPERIMENTAL TESTS

To examine the effectiveness of the proposed GRIPs and quantify the significance of diverse attributes and drivers' characteristics, experimental tests are conducted. There are two major approaches employed for assessing sign comprehension, a questionnaire survey and a driving simulation experiment. Although driving simulation has been widely used to gauge drivers' behavior to reduce the high risks and limitations involved in actual driving, the findings from simulation experiments cannot be validated sufficiently by actual driving and their reliability is questionable. In this study, the questionnaire survey method is applied, and the test scenarios are designed to gain insight into drivers' comprehension and preference for a GRIP and to assess the design and display attributes and characteristics.


**FIGURE 1 Diagrammatic Sketch for the Proposed Prototype GRIP**

*Test Scenario Design*

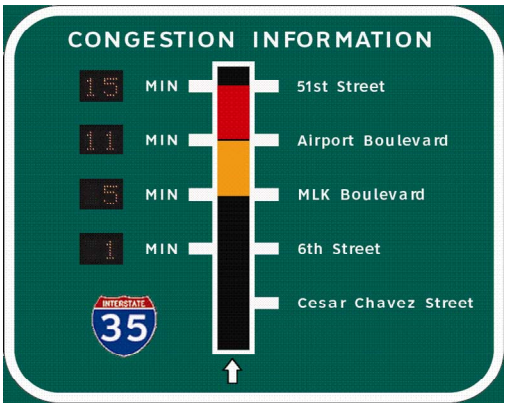
Four test scenarios were designed to examine the appropriateness and effectiveness of the proposed GRIPs from different perspectives. Figure 2 illustrates the schematic sketch of the traditional text-only DMS and the proposed GRIPs for different test scenarios. Test Scenario A and B were designed to present similar travel information by the traditional Text-only DMS and by the proposed GRIP along Interstate 35 in Austin, TX. Normal travel time is given as 10 minutes to travel along this corridor section. Drivers were asked to answer a series of questions regarding the effectiveness of each message sign and to indicate their comprehension and preference. For comparison purposes, the exact same questions were used for these two scenarios. Similarly, Test Scenarios C and D were designed to further examine the drivers' comprehension on travel conditions displayed in the proposed GRIPs. Scenario D provides additional travel time information to verify drivers' preference on traffic time measurements.

*Survey*

The survey was performed through the online survey service, www.SurveyMonkey.com between April 13, 2008 and April 24, 2008. Survey information was sent out to local email listserves. At the beginning of the survey, the basic traffic conditions and network geometric characteristics are briefly introduced. Each questionnaire contained a total of 30 questions and multiple choices in each question surveying drivers' preferences concerning the effectiveness, legibility and readability of the text-only DMS and GRIPs for four test scenarios. Many previous studies have found that drivers' responses to message signs are impacted by their demographics. Emmerink et al. *(12)* found that female drivers were less likely to be influenced by traffic information displayed. Wardman et al. *(17)* and Wang and Cao *(13)* found that young people are less inclined to comply with message sign advice. A total of 61 participants took the survey. Most were local Austin residents, including

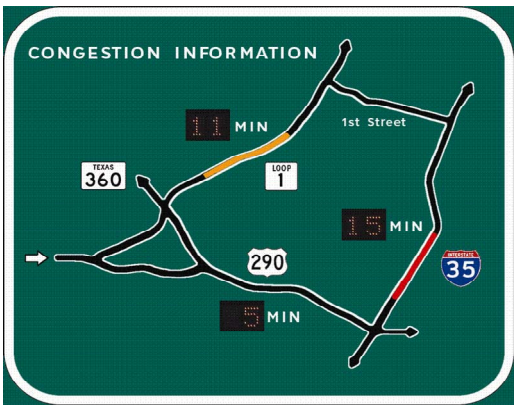
travelers from different age, gender, and travel pattern groups. To avoid any potential confusion, only one scenario was examined at a time, and the participants were required to complete all questions before moving to the next scenario. Demographic information was collected at the end of the survey. Survey results were automatically stored online to facilitate the following result analysis. The survey questions and corresponding results are presented in the next section.


Test Scenario A




Test Scenario B


Test Scenario C




Test Scenario D

**FIGURE 2 Schematic Sketch of the DMS and GRIPs for Test Scenarios A to D**

*Result Analysis and Discussion*

Results collected from the 61 participants were analyzed. Tables 1 and 2 show the survey questions and results for Test Scenario A to D. Table 3 shows the participants' demographic and travel pattern data. As can be seen in Table 1, the effectiveness of the GRIPs is verified by the fact that 41% of participants believe GRIPs are highly effective in their routing decisions, while 9.8% of participants think text-only message signs are highly effective under the same travel conditions. Based on the text-only message signs, only 63.9% of participants can identify the congestion levels, which is significantly improved by using the GRIPs, when about 95.1% of participants can identify the congestion levels. Significance tests were conducted for the effectiveness comparisons. T-statistic test was calculated based on the survey data. The *t-stat* is equal to 3.62, which corresponds to the p-value of 0.0003. Therefore, the

results indicate the GRIP can significantly improve the drivers' comprehension on travel time and routing selection at the level of $p<0.01$.

Similar analysis was conducted for Test Scenarios C and D to verify the significance of travel time indication in Table 2. In addition to the color coded level of congestion, 90.2% of participants prefer having travel time indication to present the congestion levels. 42.6% of participants believe the GRIPs are highly effective in choosing a route with estimated travel time indication. When travel time estimation is eliminated from the panel, only 4.8% of participants think the GRIPs are highly effective. The analogous results are consistent across all demographic categories of gender and age as illustrated in Table 4. Both the majority of females (65.4%) and males (71.4%) judge Scenario C as effective. For Scenario D, the majority of females (53.8%) classify the sign as effective; a higher share of males (45.7%) classify it as highly effective. Overall, the GRIP with travel time estimation received more highly effective ratings from both males (45.7%) and females (38.5%). For Scenario C, within the highly effective rating, the percentage of each bracket varies from 4.5% for 15-25 year olds, to 0% for those 26-35, 0% for those 36-45, and 22.2% for those in the 46+ age bracket. Scenario D indicates greater highly effective ratings across all age groups. The percentage increased to 54.5% for 15-25 year olds, 34.8% for those 26-35, 28.6% for those 36-45, and 44.4% for those in the 46+ age bracket. These results demonstrate that drivers prefer the travel time indication on the GRIPs. These also indicate that the older the driver, the more effective the information conveyance of the GRIPs.

**TABLE 1 Survey Questions and Results for Test Scenarios A and B**

| Survey Questions | Responses | Scenario A | | Scenario B | |
|---|---|---|---|---|---|
| | | Frequency | Percent | Frequency | Percent |
| Does the sign above clearly convey congestion levels? | Yes | 39 | 63.9 | 58 | 95.1 |
| | No | 22 | 36.1 | 3 | 4.9 |
| In the sign above, how long might it take to get to 51st Street? | 5 minutes | 0 | 0.0 | 0 | 0.0 |
| | 11 minutes | 0 | 0.0 | 0 | 0.0 |
| | 15 minutes | 61 | 100.0 | 49 | 80.3 |
| | 20 minutes | 0 | 0.0 | 12 | 19.7 |
| Would the sign above influence you to change your planned route? | Yes, I would take local roads instead | 18 | 29.5 | 32 | 52.5 |
| | No, I would remain on my original route | 43 | 70.5 | 29 | 47.5 |
| Does the sign above allow you to judge travel time effectively? | Yes | 56 | 91.8 | 57 | 93.4 |
| | No | 5 | 8.2 | 4 | 6.6 |
| How effective is the sign above in choosing a route? | Highly Ineffective | 1 | 1.6 | 2 | 3.3 |
| | Ineffective | 10 | 16.4 | 1 | 1.6 |
| | Neither | 2 | 3.3 | 6 | 9.8 |
| | Effective | 42 | 68.9 | 27 | 44.3 |
| | Highly Effective | 6 | 9.8 | 25 | 41.0 |

**TABLE 2 Survey Questions and Results for Test Scenarios C and D**

| Survey Questions | Responses | Scenario C | | Scenario D | |
|---|---|---|---|---|---|
| | | Frequency | Percent | Frequency | Percent |
| Does the sign above clearly convey congestion levels? | Yes | 45 | 72.6 | 55 | 90.2 |
| | No | 17 | 27.4 | 6 | 9.8 |
| In the sign above, which route shows heavier congestion: Loop 1 or I-35? | Loop 1 | 2 | 3.2 | 3 | 4.9 |
| | I-35 | 60 | 96.8 | 58 | 95.1 |
| Would the sign above influence you to change your planned route? | Yes, I would travel on Loop 1 instead | 41 | 66.1 | 41 | 67.2 |
| | Yes, I would travel on local roads instead | 14 | 22.6 | 10 | 16.4 |
| | No, I would remain on my original route | 7 | 11.3 | 10 | 16.4 |
| Does the sign allow you to judge travel time effectively? | Yes | 13 | 21.0 | 54 | 88.5 |
| | No | 49 | 79.0 | 7 | 11.5 |
| How effective is the sign above in choosing a route? | Highly Ineffective | 0 | 0.0 | 2 | 3.3 |
| | Ineffective | 9 | 14.5 | 1 | 1.6 |
| | Neither | 8 | 12.9 | 4 | 6.6 |
| | Effective | 42 | 67.7 | 28 | 45.9 |
| | Highly Effective | 3 | 4.8 | 26 | 42.6 |

To further analyze various factors associated with drivers' comprehension of the GRIPs and find the statistical significance behind these associations indicated by the cross-classifications of the results, a discrete choice model, a Logit regression model *(26)* was developed to Test Scenarios B and D. The regression estimates the impacts of the related characteristics on the probability of correctly understanding the travel condition information on the panels

$$P_{HE} = \frac{e^{U_{HE}}}{e^{U_{HI}} + e^{U_{IN}} + e^{U_{Nei}} + e^{U_{EF}} + e^{U_{HE}}} = \frac{e^{U_{HE}}}{\sum e^{U_{Eff}}} \quad \text{(Eq. 1)}$$

where, $P_{HE}$ is the probability that travelers can comprehend the information on the panels completely and believe the panel is highly effective. U is the utility function, and is formulated as:

$$U = \beta_1 * \text{Gender} + \beta_2 * \text{Age} + \beta_3 * \text{TraTime} + \beta_4 * \text{DrivingExp} + \cdots + \beta_{16} * \text{MajorHW} \quad \text{(Eq. 2)}$$

where, β is the coefficient that is calibrated using the survey data. The 16 dependent variables including the participants' demographic and travel pattern data are formulated in the utility function. These variables represent various attributes and characteristics that may influence drivers' comprehension of the GRIPs. Based on the survey data, the Logit model was calibrated and significant variables were identified. Measurements of goodness-of-fit were calculated: the Log likelihood function is -98.1757 and the R-squared is .4183, which indicate the Logit model can reasonably describe the data attributes and is capable of extracting their intrinsic associations.

The utility functions are computed as follows:

$$\begin{cases} U(\text{Highly Ineffective}) = \beta_{103} * TraTime \\ U(\text{Ineffective}) = \beta_{210} * DrivingExp \\ U(\text{Neither}) = \beta_{304} * TraTime + \beta_{311} * DrivingExp \\ U(\text{Effective}) = \beta_{411} * Mapskill \\ U(\text{High Effective}) = \beta_{502} * Age - \beta_{513} * Commute \end{cases} \quad \text{(Eq. 3)}$$

where, the variable explanations are provided in Table 5 and their coefficient significance is illustrated in Table 6.

**TABLE 3 Survey Participants' Demographic Data and Travel Patterns**

| Demographic | Response | Frequency | Percent |
|---|---|---|---|
| Gender | Male | 35 | 57.4 |
| | Female | 26 | 42.6 |
| Age | 18-25 | 22 | 36.1 |
| | 26-35 | 23 | 37.7 |
| | 36-45 | 7 | 11.5 |
| | 46+ | 9 | 14.8 |
| What is your level of education? | High School | 1 | 1.6 |
| | Some College, No Degree | 4 | 6.6 |
| | Associate's Degree | 0 | 0.0 |
| | Bachelor's | 22 | 36.1 |
| | Master's | 28 | 45.9 |
| | Doctorate | 6 | 9.8 |
| Are you familiar with the Austin, Texas road network? | Yes | 41 | 67.2 |
| | No | 20 | 32.8 |
| What is your driving experience? | 0-5 years | 5 | 8.2 |
| | 6 or more years | 56 | 91.8 |
| How do you rate your map reading skills? | Horrible | 0 | 0.0 |
| | Below Average | 5 | 8.2 |
| | Average | 9 | 14.8 |
| | Above Average | 20 | 32.8 |
| | Exceptional | 27 | 44.3 |
| Do you use a GPS navigation system? | Yes | 10 | 16.4 |
| | No | 51 | 83.6 |
| Do you commute to work? | Yes | 37 | 60.7 |
| | No | 24 | 39.3 |
| How many miles do you drive daily? | 0-5 miles | 9 | 37.5 |
| | 6-10 miles | 9 | 37.5 |
| | 11-15 miles | 1 | 4.2 |
| | 16-20 miles | 4 | 16.7 |
| | 21+ miles | 1 | 4.2 |
| How many miles is your daily commute (one way)? | 0-5 miles | 14 | 37.8 |
| | 6-10 miles | 5 | 13.5 |
| | 11-15 miles | 10 | 27.0 |
| | 16-20 miles | 4 | 10.8 |
| | 21+ miles | 4 | 10.8 |
| Does your commute involve travel on major highways? | Yes | 22 | 59.5 |
| | No | 15 | 40.5 |

**TABLE 4 GRIP Effectiveness Comparisons by Gender and Age for Scenarios C and D**

| Scenario | Effectiveness | Measure | Gender | | | Age | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Female | Male | Total | 15-25 | 26-35 | 36-45 | 46 | Total |
| C | Highly Effective | Frequency | 2 | 1 | 3 | 1 | 0 | 0 | 2 | 3 |
| | | Percent | 7.7 | 2.9 | 4.9 | 4.5 | 0.0 | 0.0 | 22.2 | 4.9 |
| | Effective | Frequency | 17 | 25 | 42 | 19 | 15 | 4 | 4 | 42 |
| | | Percent | 65.4 | 71.4 | 68.9 | 86.4 | 65.2 | 57.1 | 44.4 | 68.9 |
| | Neither | Frequency | 5 | 2 | 7 | 1 | 4 | 2 | 0 | 7 |
| | | Percent | 19.2 | 5.7 | 11.5 | 4.5 | 17.4 | 28.6 | 0.0 | 11.5 |
| | Ineffective | Frequency | 2 | 7 | 9 | 1 | 4 | 1 | 3 | 9 |
| | | Percent | 7.7 | 20.0 | 14.8 | 4.5 | 17.4 | 14.3 | 33.3 | 14.8 |
| | Total | Frequency | 26 | 35 | 61 | 22 | 23 | 7 | 9 | 61 |
| | | Percent | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| D | Highly Effective | Frequency | 10 | 16 | 26 | 12 | 8 | 2 | 4 | 26 |
| | | Percent | 38.5 | 45.7 | 42.6 | 54.5 | 34.8 | 28.6 | 44.4 | 42.6 |
| | Effective | Frequency | 14 | 14 | 28 | 7 | 12 | 4 | 5 | 28 |
| | | Percent | 53.8 | 40.0 | 45.9 | 31.8 | 52.2 | 57.1 | 55.6 | 45.9 |
| | Neither | Frequency | 2 | 2 | 4 | 3 | 1 | 0 | 0 | 4 |
| | | Percent | 7.7 | 5.7 | 6.6 | 13.6 | 4.3 | 0.0 | 0.0 | 6.6 |
| | Ineffective | Frequency | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| | | Percent | 0.0 | 2.9 | 1.6 | 0.0 | 4.3 | 0.0 | 0.0 | 1.6 |
| | Highly Ineffective | Frequency | 0 | 2 | 2 | 0 | 1 | 1 | 0 | 2 |
| | | Percent | 0.0 | 5.7 | 3.3 | 0.0 | 4.3 | 4.3 | 0.0 | 3.3 |
| | Total | Frequency | 26 | 35 | 61 | 22 | 23 | 7 | 9 | 61 |
| | | Percent | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

**TABLE 5 Variable Explanations for the Logit Model**

| Variables | Explanations |
|---|---|
| TraTime | Does the sign above allow you to judge travel time effectively? |
| | 0 (No), 1 (Yes) |
| DrivingExp | What is your driving experience? |
| | 0(0-5), 1 (6 or more) |
| Mapskill | How do you rate your map reading skills? |
| | 0 (Horrible ), 1 (Below Average ) , 2 (Average ), 3( Above Average ), |
| | 4 (Exceptional) |
| Age | 0 (15-25), 1 (26-35), 2 (36-45 ), 3 (46 or above) |
| Commute | Do you commute to work? |
| | 0 (No), 1 (Yes) |

As can be seen in Table 6 and Equation 3, the coefficient of the *Age* variable is positive (0.46) and that of the *Commute* variable is negative (-1.50) in the utility function for the option of highly effective. These results indicate that older drivers more prefer GRIPs. Using GRIPs, older drivers' message comprehension can be noticeably enhanced by the graphic information panel. This finding is consistent with the research findings from Wardman et al. *(19)* and Wang and Cao *(7)*. The results also indicate that commuters do not find GRIPs highly effective. This may be

partially because they trust their own judgment of local travel conditions more than information provided on the sign. More detailed traffic information, such as incident induced congestion, may satisfy their needs.

**TABLE 6 Variable Significance for the Logit Model**

| Variable | Coefficient | Standard Error | T-Statistic | P-Value |
|---|---|---|---|---|
| $\beta_{103}$ | -2.18 | 0.60 | -3.62 | 0.0003 |
| $\beta_{210}$ | -2.59 | 0.78 | -3.31 | 0.0009 |
| $\beta_{304}$ | -4.82 | 1.34 | -3.59 | 0.0003 |
| $\beta_{311}$ | 3.05 | 1.23 | 2.48 | 0.0132 |
| $\beta_{411}$ | -0.35 | 0.21 | -1.67 | 0.0956 |
| $\beta_{502}$ | 0.46 | 0.27 | 1.70 | 0.0894 |
| $\beta_{513}$ | -1.50 | 0.52 | -2.87 | 0.0041 |

In the utility function for the option of highly ineffective, the coefficient of the *TraTime* variable is negative (-2.18). This indicates that fewer drivers consider GRIPs ineffective when travel time information is provided. Similar analysis was conducted for the other utility functions. Two additional conclusions could be identified in the results: 1) map reading skills are not necessary for conveying traffic condition information by the GRIPs, and 2) more driving experience is beneficial for drivers to effectively use the information on the GRIPs. In general, these results reinforce the utility of statistical testing and utility function development in identifying the associations among the panel attributes, drivers' characteristics, and drivers' preference and comprehension of GRIPs. Once these associations are determined, corresponding improvements can be made to better address these significant attributes and drivers' characteristics to enhance the GRIPs' effectiveness in delivering traffic routing information.

Additional data collection would be required before application of analysis results for GRIPs implementation. This analysis was completed using an on-line survey, which biased the survey towards those with computer access. Some demographic biases were also identified. The vast majority of survey respondents were between the ages of 18 and 35; as a result, sample sizes for older age groups were very small. Additionally, more than 90 percent of respondents have completed at least a bachelor's degree. As a result, drivers with high school or associate's degree educations are not adequately represented in the sample. Before real design and implementation of a GRIP, new methods of data collection will need to be identified to capture underrepresented populations.

## CONCLUSIONS

This study concentrates on designing an innovative graphic-aided message sign, the GRIP, and determining a methodology for understanding its potential usefulness to drivers. GRIPs allow more complex information to be displayed to facilitate drivers' personalized routing needs. GRIPs can effectively overcome some limitations of traditional text-only message signs, such as limited display capacity. However, they

must be carefully designed to avoid traffic information overload and ensure message comprehension. In this study, based on the Texas MUTCD standards and previous studies, prototype GRIPs were designed to provide travelers with real time, color coded traffic condition information in a graphical format. Based on the proposed prototype GRIPs, four test scenarios were performed to examine their effectiveness and drivers' preferences, and to determine a method of analysis for identifying the related significant attributes. In addition to statistical analysis of the survey results, a Logit model was established to investigate the significance of various drivers' characteristics and design attributes impacting the effectiveness of GRIPs.

Results of the experimental tests indicate that for the given dataset, the proposed GRIPs can significantly enhance the legibility and recognition of message signs and improve travelers' comprehension. Relationships between age, gender, information types, travel patterns, and map use and GRIP effectiveness and ineffectiveness were examined. For the given dataset, driver age, the presence of travel time information, familiarity with travel on with the network (determined by commute type), and driver experience were found to significantly influence GRIP use. Overall, this analysis indicates that GRIPs can effectively provide improved traveler information to drivers. The statistical analysis performed in this study provides a framework for future examination of driver, sign, and information type variables to enable effective GRIP design and implementation with an improved dataset.

## REFERENCES


1. Bureau of Transportation Statistics. (2007). National Transportation Statistics 2005. U.S. Department of Transportation. Washington, D.C.
2. Shrunk, D. and T. Lomax. *The 2005 Urban Mobility Report*. Texas Transportation Institute. The Texas A&M University System (http://mobility.tamu.edu/, access on Jul. 14, 2009). 2005.
3. Federal Highway Administration, Manual on Uniform Traffic Control Devices (MUTCD), U.S. Department of Transportation, Washington, D.C., 2003.
4. Texas Department of Transportation, Texas Manual on Uniform Traffic Control Devices, 2006. (http://www.mctraffic.org/mutcd_2006_rev1_intro.pdf, access on Jul. 14, 2009).
5. Alkim, T. P., P. H. J. Van Der Mede, and W. H. Janssen. Graphical Route Information on Variable Message Signs. Proc., 10th International Conference on Road Transport Information and Control, London, April 4–6, 2000
6. Schönfeld, Grit, et al. Dynamic Driver Information Goes Graphical – A New Quality in Urban Traffic Information. 7th World Congress on Intelligent Transport Systems. Turin, Italy 2000.
7. Chrysler, Susan T ; Holick, Andrew James ; Williams, Alicia A ; Funkhouser, Dillon , Driver Comprehension of Diagrammatic Advanced Guide Signs and Their Alternatives. The 85th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.
8. Wang, J.H., S. Hesar, and C. Collyer. Adding Graphics to Dynamic Message Sign Messages. *Transportation Research Record*, Vol. 2018, 2007, p. 63-71.

9. Benson B.G. Motorist Attitudes about Content of Variable-Message Signs. In*Transportation Research Record: Journal of the Transportation Research Board,No. 1550*, 1996, 48-57.
10. Chatterjee, K., N.B. Hounsell, P.E. Firmin, and P.W. Bonsall. Driver Response to Variable Message Sign Information in London. *Transportation Research – C*, Vol. 10, No. 2, 2002, 149-169.
11. Peng Z. R., N. Guequierre, and J.C. Blakerman. Motorist Response to Arterial Variable Message Signs. In *Transportation Research Record: Journal of the Transportation Research Board, No.1899*, 2004, 55-63.
12. Emmerink, R., P. Nijkamp, P. Rietveld, and J.N.Van Ommeren. Variable Message Signs and Radio Traffic Information: An Integrated Empirical Analysis of Drivers' Route Choice Behaviour. *Transportation Research – A*, Vol. 30, No. 2, 1996, 135-153
13. Wang, J.-H., and Y. Cao. Assessing Message Display Formats of Portable Variable Message Signs. In Transportation Research Record: Journal of Transportation Research Board, No. 1937, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 113–119.
14. Dudek, C. L., and G. L. Ullman. Variable Message Sign Operations Manual. Publication FHWA-NJ-2001-10. Texas Transportation Institute, College Station, 2001.
15. Dudek, C. L., and G. L. Ullman. Flashing Messages, Flashing Lines, and Alternating One Line on Changeable Message Signs. In Transportation Research Record: Journal of the Transportation Research Board, No. 1803, 2002, pp. 94–101.
16. Wang, J.-H., C. M. Yang, and D. Waters. Effects of Message Display on Motorists Comprehension of and Response to Arterial Dynamic Message Signs. Presented at 84th Annual Meeting of the Transportation Research Board, Washington, D.C., 2006.
17. Wardman, M., Bonsall, P.W., and Shires, J. D., Motorist response to variable message signs: a stated preference investigation. Transportation Research – C, Vol. 5, No. 6, p.389-405, 1997.
18. Peeta, S., Ramos, J. L., and Pasupathy, R., Content of variable message signs and on-line driver behavior. Transportation Research Record: Journal of the Transportation Research Board, no. 1725, p.102-108, 2000.
19. Dewar, R., D. Kline, F. Scheiber, and A. Swanson. Symbol Signing Design for Older Drivers. Publication FHWA-RD-94-069. FHWA, McLean, Va., 1997.
20. Tsavachidis M., and H. Keller. Graphical Traffic Information on Dynamic Information Boards. Proceeding of European Transport Conference, Cambridge, England, 2000, pp. 201-215.
21. Lucas, A., A. Arbaiza, and M.T. Blanch. New Steps towards an International Symbolic Language: Some Results from Spain. Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems, 2005.
22. Conference of European Directors of Roads. Action FIVE-Framework for harmonized implementation of Variable message signs in Europe. Draft amended version 3.5.

2003.( http://www.esafetysupport.org/download/documents/Harmonisation_VMS_Systems.pdf, access on Jul. 14, 2009).
23. Staplin, L., K. Lococo, and J. Sim. *Traffic Control Design Elements forAccommodating Drivers with Diminished Capacity,* Volume II. Report No. FHWARD-90-055. Washington, D.C.: Federal Highway Administration, 1990.
24. Hanowski, R.J. and B. H. Kantowitz. Driver Memory Retention of In-Vehicle Information System Messages. *Transportation Research Record 1573*, 1997, pp. 8-18.
25. Bruce D., Boehm-Davis, D.A., and Mahach, K., In-Vehicle Auditory Display of Symbolic Information. Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society: Ergonomics for the New Millennium, San Diego, CA, p. 230-233, 2000.
26. McFadden, Daniel L. Conditional Logit Analysis of Qualitative Choice Behavior. In: Zarembka P (ed) Frontiers in Econometrics. New York, pp 105–142, 1974.

# Defining Benefits from Pavement Rehabilitation and Preservation

By

T. A. Dawson, G. Y. Baladi, C. M. Dean, S. W. Haider, and K. Chatti

3546 Engineering Building, Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI 48824-1226; PH (517) 355-5147; FAX (517) 432-1827; email: Dawsonty@msu.edu

## ABSTRACT

Pavement service life can be defined as the estimated number of years between pavement construction or rehabilitation and when the pavement section reaches a given condition (rutting, cracking, roughness, etc.) threshold value. At that time, the pavement section is typically subjected to rehabilitation or preservation actions. The two overarching methodologies for determination of pavement fix benefits are: 1) the extension in the pavement service life due to a given fix and 2) the calculation of the area between the pavement performance curves and a given threshold value. The life extension method places importance in prolonging the service life of the pavement. On the other hand, the area under the performance curve method places importance on the level of pavement distress over time. Selection of the pavement fix type and time with the lowest cost to benefit ratio could be considered "optimal".

Keywords: Remaining Service Life (RSL), pavement distress threshold, life cycle cost, optimal pavement fix type and time selection, and cost/benefit analyses

## INTRODUCTION

Highway agencies (HAs) across the country select pavement preservation and rehabilitation activities based on their benefits. The minimization of the costs to benefits ratio could be used as a tool for selecting the optimum pavement action and application time. Although no universal definition of benefits exists, the two overarching methodologies for determining the pavement fix benefits are: 1) the extension in the pavement service life and 2) the calculation of the area between the predicted pre-fix and post-fix pavement performance curves and a given threshold value. The cost of the fix is divided by its calculated benefit and the cost to benefit ratio is then compared to those of other fix types and times. Selection of the pavement fix type and time with the lowest cost to benefit ratio could be considered "optimal".

The above two methods for calculating the benefits from pavement actions are based on modeling the pre-fix pavement performance and estimating the post-fix performance. The pre-fix performance is modeled based on time-series pavement

distress data and the model is used to predict future conditions. The post-fix performance is a function of the pre-fix pavement conditions and rates of deterioration and must be estimated. The pavement pre- and post- fix performance models are used to calculate the estimated life extension and/or the area under the performance curves.

The life extension method places importance in prolonging the service life of the pavement. On the other hand, the area under the performance curve method places importance on the level of pavement distress over time. The selection of the optimum fix type and time could be based on providing the highest life extension or the maximum area under the curve for the given cost. It should be noted that both the area and the life extension methods could be based on the same or on different pavement condition thresholds. Additionally, the benefits of a pavement fix should be calculated for each pavement distress type and the minimum benefit is considered the overall benefit.

## PAVEMENT PERFORMANCE MODELING

The first step in defining the benefits of a pavement preservation or rehabilitation action is modeling the performance of the pre-fix pavement and estimating the post-fix performance. The performance is measured from the last resurfacing or reconstruction to a given distress threshold value. The modeling and prediction of pavement performance are discussed in the next few subsections.

### Pre-Fix Pavement Distress Modeling

For each distress type, the pre-fix time-series distress data and the proper mathematical function (see Table 1) are used to model the performance of a pavement section. The mathematical functions listed in Table 1 are based on known trends in the pavement deterioration. After determining the regression constants of the model, it can be used to predict future performance. It is important to note that a minimum of three data points (three data collection cycles) during which no pavement preservation or rehabilitation actions were taken are required to properly model the time-series pavement distress data.

Table 1 Pavement condition models

| Pavement distress type | Model form | Generic equation |
|---|---|---|
| Roughness (IRI) | Exponential | $IRI = a\exp^{bt}$ |
| Rutting | Power | $Rut = at^{b}$ |
| Cracking | Logistic (S-shaped) | $Crack = \frac{Max}{1+\exp^{(a+bt)}}$ |
| Where, a and b are regression parameters, t = time (year), and Max = the maximum value of cracking | | |

**Post-Fix Pavement Performance Estimation**

The pavement performance after a preservation or rehabilitation action must also be modeled to define the benefits of the action. The same mathematical functions listed in Table 1 are used for the post-fix performance models. However, the models cannot be fit to time-series distress data because none exist. Hence, the model parameters must be estimated. The factors which affect the post-fix performance are: the pre-fix conditions and rates of deterioration, the pre-fix repairs, the type of fix applied, and the pavement section. Experience and data from previous fixes applied to similar pavement types having similar conditions and rates of deterioration are needed to develop models to predict -post-fix pavement conditions and performance. In the absence of such data, historical national trends may be used to approximate the post-fix performance. However, the degree of confidence in the estimation will be minimal.

**Pavement Distress Thresholds**

The mathematical functions used to model pavement performance are typically based on the time dependent pavement condition data that were collected since the last pavement action (resurfacing, reconstruction, etc.). These models are also used to predict the pavement performance (both pre and post-fix) between the current conditions and the time when the pavement conditions reach specified threshold values. For each distress type, multiple threshold values could be specified as shown in Figure 1. One value triggers preventive maintenance actions, other rehabilitation actions, and a third reconstruction. Hence, the threshold values are based on the class or types of pavement fixes and/or the acceptable conditions to the user (such as serviceability or ride quality). Finally, the reconstruction threshold value indicates the end of the pavement service life and is typically used to measure the benefits (such as life extension and the area under the performance curve) of a pavement action.

**LIFE EXTENSION**

The life extension is a measure of the benefit from a pavement action in terms of the additional number of years the pavement section will perform at or better than the threshold value (see Figure 2). The threshold value could also be the conditions of the pavement section at the time of the fix (see Figure 3). In that scenario, the fix life is the benefit. The major difference between the two is the threshold value and therefore the length of time over which the pavement condition is predicted. The two methods are discussed in more detail below.

a) The life extension is calculated using Equation 1, where the difference between the $RSL_{Post\text{-}fix}$ and the $RSL_{Pre\text{-}fix}$ is the benefit (see Figure 2). The $RSL_{Post\text{-}fix}$ is the estimated number of years for the conditions of a pavement section subjected to a preservation or rehabilitation action to reach the threshold value. The $RSL_{Pre\text{-}fix}$, on the other hand, is the number of years remaining until the conditions of the

existing pavement section reach the threshold values if no actions are taken. Note that the pre and post-fix threshold values are the same for pre and post-fix RSL.

$$Life\ extension = RSL_{\text{Post-fix}} - RSL_{Pre-fix} \qquad \text{Equation 1}$$

The pre-fix mathematical functions and their parameters are statistically determined using the collected time series data. The post-fix pavement performance, on the other hand, is estimated based on a network model of previous fixes of similar pavement sections. The pavement condition threshold is a user input and could change depending on the scope, objectives, and policy of the HA. Therefore the pre and post-fix RSL values and hence the life extension could change depending on where the new threshold value is placed (assuming the pavement performance curves are not parallel). The threshold values to be used should be determined by the HA. It should be noted that the life extension based on each pavement distress type should be calculated, the minimum value is the true life extension of the pavement section.

b) The fix life is the estimated number of years for the conditions of a pavement section subjected to a preservation or rehabilitation action to reach the pre-fix conditions. The benefit is simply the fix life (see Figure 3). The main advantage of using the fix life as a measure of benefit relative to the life extension is that the length of time over which the prediction is made is reduced. The performance of the pre-fix pavement is simply based on the collected distress data, no prediction is required. The post-fix pavement performance, on the other hand, is predicted over shorter period into the future. This results in a higher confidence in the predicted time and in the calculated benefit.

The methods for calculating or estimating the benefits of pavement fixes based on life extension or fix life are practical and can be easily understood. The benefit (in years) can be expressed to engineers, management, the legislature, and the public. The method also enables a HA to optimize the selection of pavement fix type and timing by maximizing the longevity of a pavement section, thereby reducing the life-cycle costs. The most cost-effective pavement preservation strategy could also be determined using these methods.

## PAVEMENT PERFORMANCE AREA

In the pavement performance area procedure, the total benefit (TB) is calculated by the ratio of the areas expressing the benefit of the fix to the area of the pre-fix pavement with the do-nothing option as described by (Peshkin et. al. 2004) and Equation 2. These two areas are shown in Figure 4. The TB is a unit-less measure that should be estimated for each distress type and the smallest value controls. The performance area procedure could be divided into the two methods detailed below.

$$TOTAL \quad BENEFIT = \frac{AREA_{(BENEFIT)}}{AREA_{(Do-NOTHING)}} \qquad \text{Equation 2}$$

a) The benefit of a fix is calculated by the area bound by the predicted pre-fix and post-fix pavement performance curves and the threshold value as shown in Figure 4.
b) The benefit of a fix is calculated by the area bound by the predicted pre-fix and post-fix pavement performance curves and the pre-fix conditions as shown in Figure 5. The main advantage of estimating the benefits based on the pre-fix pavement conditions is the reduction in the prediction time and hence increased accuracy.

The pavement fix benefits measured in terms of TB are not as practical as the life extension or fix life measurements. It cannot be easily communicated or understood by non-technical users such as the public, the legislators, or even upper management. The life extension or fix life benefits expressed in term of years are better communicated and understood. As is the case for life extension and fix life, the TB method could be used to optimize the selection of pavement fix type and timing by maximizing the benefits for a given cost. In this case the benefit is the performance (condition versus time) as opposed to just time. It should be noted that the TB should also be calculated for each pavement distress type, and the minimum value should control.

## OPTIMIZATION

The purpose of calculating the benefits from a pavement fix is to assist the HA in the selection of optimum pavement fix type and time. An optimum pavement fix is one that provides the greatest benefit at the minimal cost. The cost calculation, objective function, and the effects of the benefit measurement are discussed below.

### Costs

The costs of any pavement preservation or rehabilitation action can be divided into two groups; agency and user costs. The Agency costs are the sum of the physical costs of all activities of the pavement project including design, materials, traffic control, and construction minus the residual value of the pavement section at the end of its life. These are often referred to as direct costs. User costs include vehicle operating costs, delay costs, and accident costs, which are much more difficult to estimate than agency costs as they are not based on specific monetary values (Morgado & Neves 2008). The life cycle cost is the accumulation of agency and user costs over the life of the pavement section, typically expressed in terms of present worth (Walls & Smith 1998). The life cycle cost is utilized to capture the cost of the entire pavement preservation or rehabilitation strategy.

### Objective Function

The objective function which will yield the optimum pavement fix type and timing is based on minimizing the costs to benefit ratio stated in Equation 3.

$$Optimum = Minimum\left(\frac{Cost}{Benefit}\right) \quad \text{Equation 3}$$

It should be noted that the method used to calculate the benefit has an impact on communicating the optimum fix. Use of the life extension or the fix life methods to determine the benefits of pavement fixes results in a cost to benefit ratio in dollar/year, which is an easily communicable term. The TB method, on the other hand, results in unit of measurement in dollar/area ratio, which is not easily communicable.

## IMPACT OF THE STATE OF THE PRACTICE ON BENEFITS

The calculation of the benefits from a pavement preservation or rehabilitation action directly depends on the pavement performance models. The performance models are affected by the methods and procedures for collecting, managing, and analyzing the pavement distress data (the state of the practice), which are not universal. The various factors affecting the pavement performance models and therefore the benefits are listed below.

- The frequency of pavement distress data collection
- The number of available distress data points (pavement distress history) and their accuracy.
- The data sampling procedure
- The availability of comprehensive maintenance records that include locations, actions and costs
- The availability of accurate and properly referenced cost data
- The historical record of all previous pavement rehabilitation actions and their associated costs

For a given pavement project, the pavement performance models and the calculated benefits of fix actions will be different depending on the state of the practice. To illustrate, consider the time series pavement distress data and the resulting pavement performance model shown in Figure 6, from SRID 020 in Washington State. The distress data are collected annually along the entire pavement network (no sampling), and comprehensive maintenance records are available although they are not shown in Figure 6. Then consider the following six alternative scenarios for the same pavement project under different state of the practice:

1) Scenario 1 – The full pavement distress data set and maintenance records are used to create the most accurate pavement performance model and benefits calculations.
2) Scenario 2 – The pavement distress data is collected every other year, as shown in Figure 7. This would decrease the number of available data points resulting in a different or less accurate performance model. Therefore, the calculated benefits will not be as accurate.
3) Scenario 3 – The first five years pavement distress data points are not available because the agency started collecting the data later. If the pavement surface age is not known (which is the case for many roads), a gap in the time-series distress will

ensue and the resulting performance model will not be accurate and so the benefits of a pavement action.

4) Scenario 4 – The pavement distress data are collected using some type of sampling technique, as shown in Figure 8. Hence, the data for most pavement miles are not measured but assumed based on the data from the sampled section. In a companion paper, the effects of sampling on the accuracy of the pavement management decision are addressed in detail (Dean et. al. 2010). It is shown that ten percent sampling causes about 50 percent errors in the pavement decisions and in the selection of project boundaries.
5) Scenario 5 – The time at which previous pavement actions are taken is unknown (incomplete pavement maintenance and rehabilitation records). This causes errors in the performance model and hence in the calculation of the benefits.
6) Scenario 6 – The inventory data are not complete or accurate, for example, the pavement type is unknown or incorrect. This scenario often takes place because the various districts or regions within a highway agency do not communicate changing pavement type (for example, rigid to composite) to the pavement management unit or to the database manager. This would precipitate the use of inappropriate mathematical function to model the data and hence errors in the pavement performance and benefits. Further, in most cases, the thicknesses of the pavement layers are not known. Said thicknesses are required to determine the type of fixes such as the thickness of the asphalt overlay. This would also generate errors in the calculation of the benefits.

For the first five scenarios, the pavement performance model was fitted to the nine available rut data points along SRID 020 in Washington and used to estimate the benefits of a given fix using all four methods. Note that the same post-fix performance model was assumed for each scenario. The nine data points were also used to simulate the impacts of data collection frequency (scenario 2), missing data (scenario 3), sampling (scenario 4), and so forth. The results, listed in Table 2, indicate that the benefits are impacted by the state of the practice. For example; the use of the TB methods requires pavement distress data covering the entire surface age of the pavement section to calculate the do-nothing area, missing, inaccurate, and incomplete data result in less accurate performance models and therefore less confidence in the calculated benefits, and incomplete maintenance records imply the type and time of the previous action are unknown. Therefore, the method of benefit calculation could be determined based on the comprehensiveness of the data to support the method. One other note, the results in Table 2 indicate that for some fixes and scenarios, the pavement life may not be extended (zero benefit). However, the fix life and the area methods yield some sort of benefits simply because the user is driving over a better pavement for a period of time.

Combination of all or some of the methods is also an option for determining benefits from pavement fixes. For example, the life extension could be considered the primary criteria and the TB to the threshold the secondary for fixes with similar cost to benefit ratios. This combined method would allow the HA to be cost effective and increase pavement performance.

Table 2 Factors affecting pavement performance models and benefit calculations

| Scenario | IRI model | Benefits | | | |
|---|---|---|---|---|---|
| | | LE | FL | TB (threshold) | TB (pre-fix condition) |
| 1 | Full data set | 3.85 | 7.52 | 0.53 | 0.80 |
| 2 | 2 year frequency | 2.41 | 7.05 | 0.46 | 0.81 |
| 3 | Missing data | 11.32 | 7.55 | 0.34 | 0.81 |
| 4 | 10% sampling | 0.32 | 7.02 | 0.36 | 0.82 |
| 5 | No maintenance record | 0 | 3.82 | 0.01 | 0.93 |

## CONCLUSIONS

The following conclusions were drawn with regard to defining the benefits of a pavement preservation or rehabilitation action.

- The benefits defined by the life extension and fix life methods place importance in prolonging the service life of the pavement.
- The benefits defined by the areas under the performance curves methods (TB) place importance on the level of pavement distress over time (performance).
- All methods of defining benefits of pavement actions can be used for cost/benefit analyses and optimization of the selection of pavement fix type and time.
- The state of the practice could determine the method used to calculate benefits of pavement fixes.
- Pavement fix selection using a combination of the four methods could be cost-effective.

## REFERENCES

1. Dean, C. M., G. Y. Baladi, T. A. Dawson, and S. W. Haider, "The Effect of Pavement Condition Data Sampling on Project Boundary Selection", submitted for publication at the 2011 T&DI Conference, Chicago, 2011
2. Morgado, J. and J. Neves, "Accounting for User Costs when Planning Pavement Maintenance and Rehabilitation Activities", Technical University of Lisbon, Portugal, 2008
3. Peshkin, D.G., T.E. Hoerner, and K.A. Zimmerman, "Optimal Timing of Pavement Preventive Maintenance Treatment Applications", National Cooperative Highway Research Program (NCHRP) Report 523, Transportation Research Board (TRB), Washington, D.C. 2004
4. Walls, J., & M. R. Smith, "Life-Cycle Cost Analysis in Pavement Design – Interim Technical Bulletin", No. FHWA-SA-98-079, Federal Highway Administration (FHWA), Washington, 1998

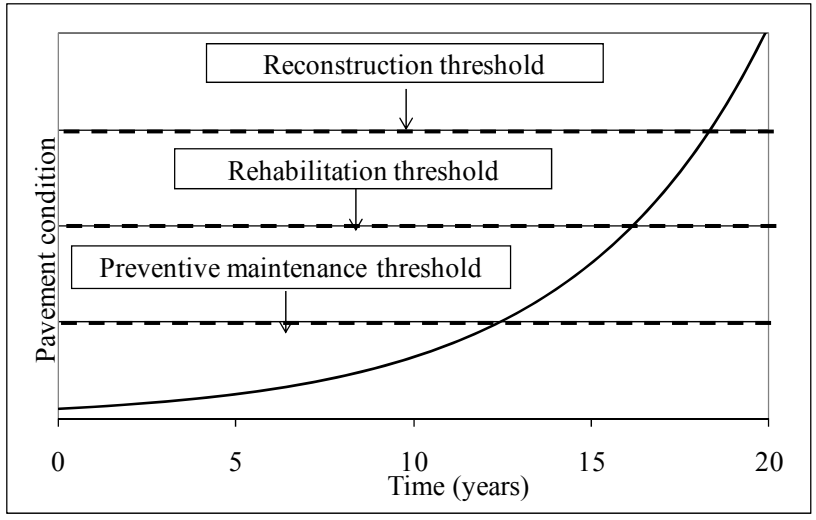



Figure 1 Various possible threshold levels

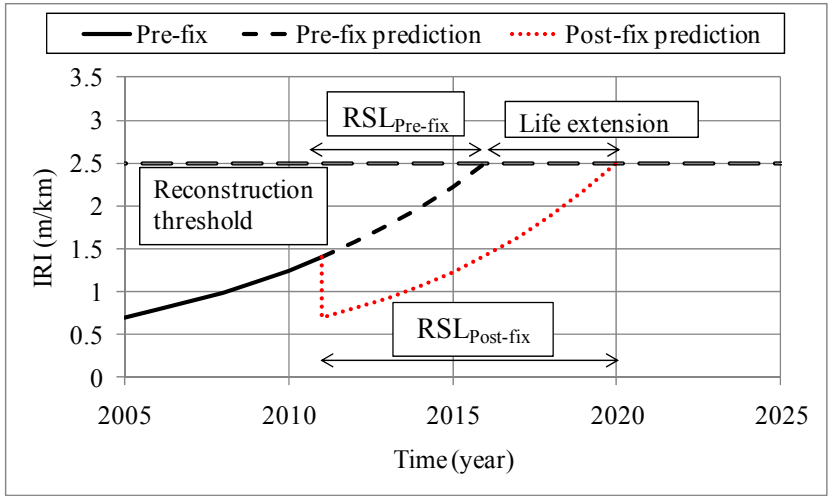



Figure 2 Life extension

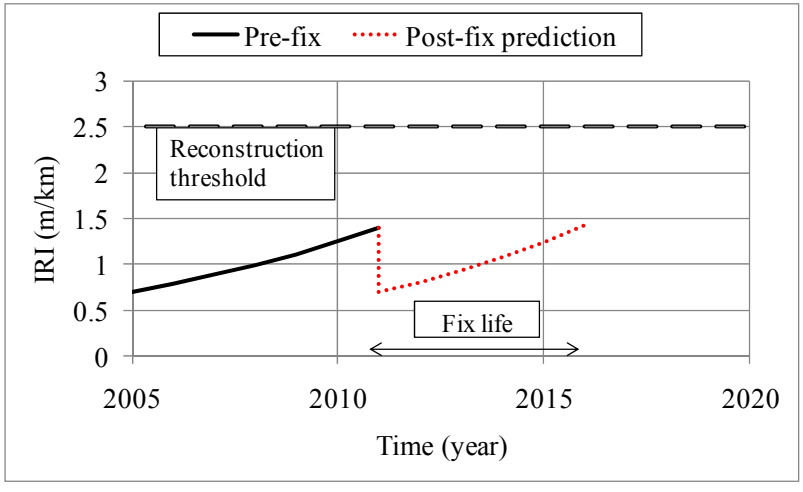



Figure 3 Fix life

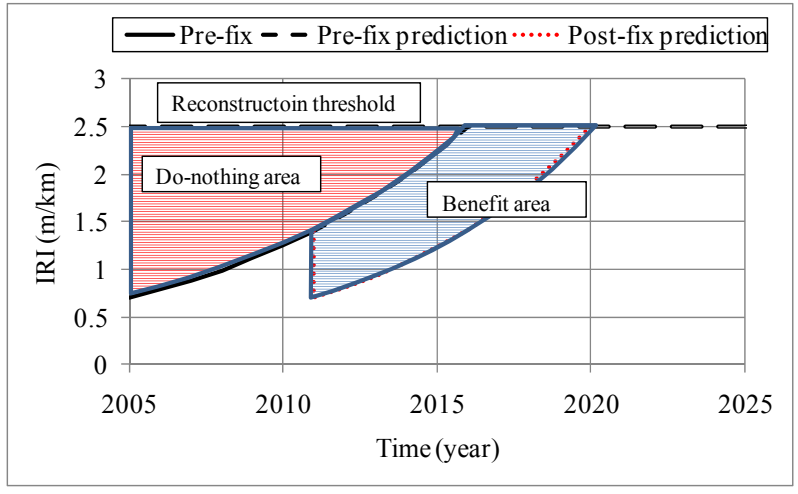



Figure 4 Total Benefit (threshold)

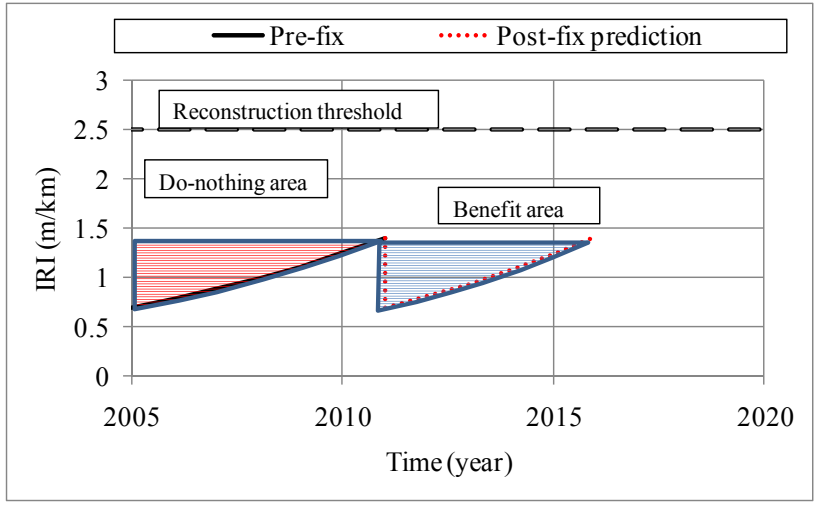



Figure 5 Total Benefit (pre-fix condition)

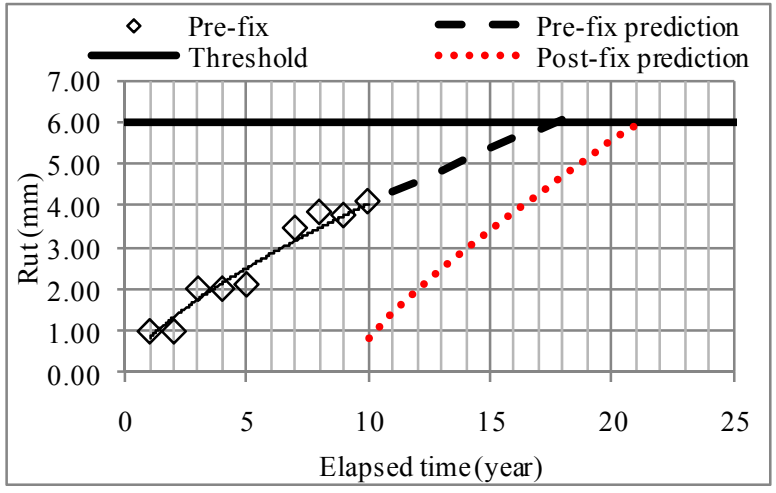



Figure 6 Scenario 1 pavement performance

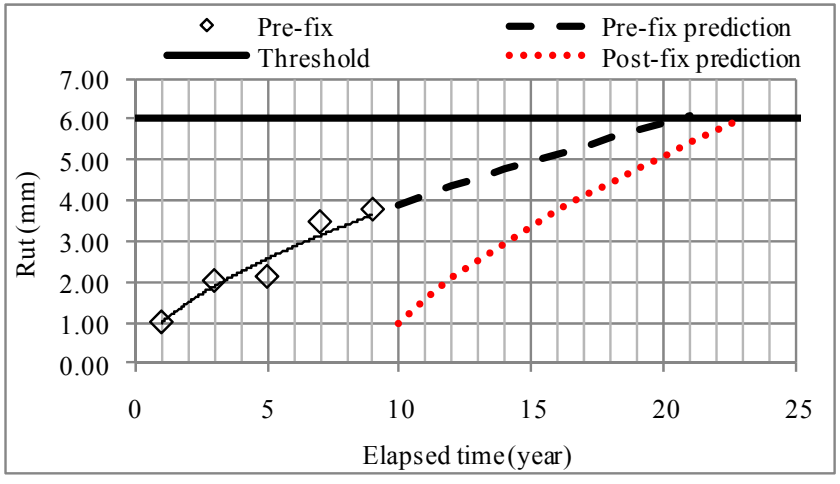



Figure 7 Scenario 2 pavement performance

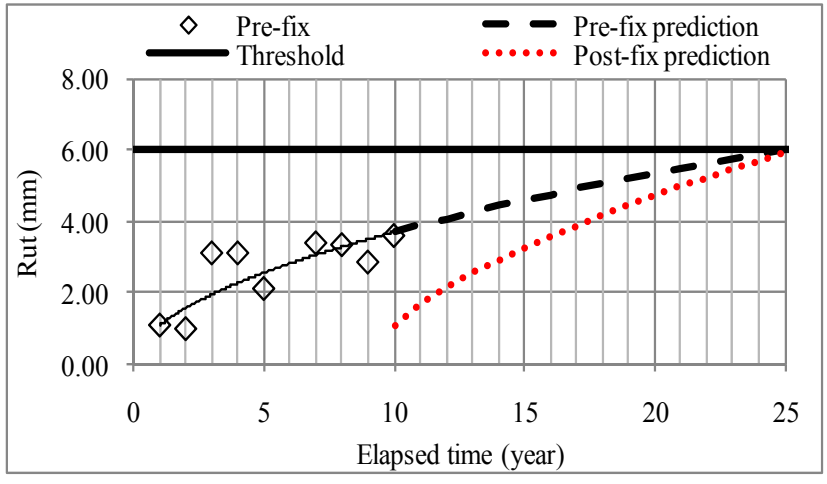



Figure 8 Scenario 4 pavement performance

# Freight Traffic Crossing the Eastern Continental Divide

By Robert T. Hintersteiner, P.E., F. ASCE
Transportation/Forensic Engineer

This paper addresses the problem of moving wholesale and retail goods around the New York City Metropolitan Regional Area. The existing roadways, bridges, and tunnels have already exceeded their capacity, and the author has assumed that freight traffic will continue to grow at the same rate as the population. To address this problem, a freight railroad tunnel under the Hudson River should be constructed to interconnect the west side of the Hudson River to the entire New England Regional Area and also to Long Island. This proposed railroad freight tunnel connection could remove over 30,000 tractor trailers per day, or about 8 million trucks per year, from the congested roadway network and the Hudson River Crossings.

In effect, the Hudson River in New York State has become a de facto "Eastern Continental Divide" that separates the entire New England Regional Area from the rest of the American continent. The Hudson River is 315 miles long and it extends from the Continental Shelf (New York Bay) north to Mount Marcy in the Adirondack Mountains at "Lake Tear of the Clouds" (4,293 ft elevation) south of the Canadian Border[1]. Interstate 87 extends from the Boroughs of Manhattan and the Bronx in New York City to the Canadian Border, a distance of about 334 miles[2].

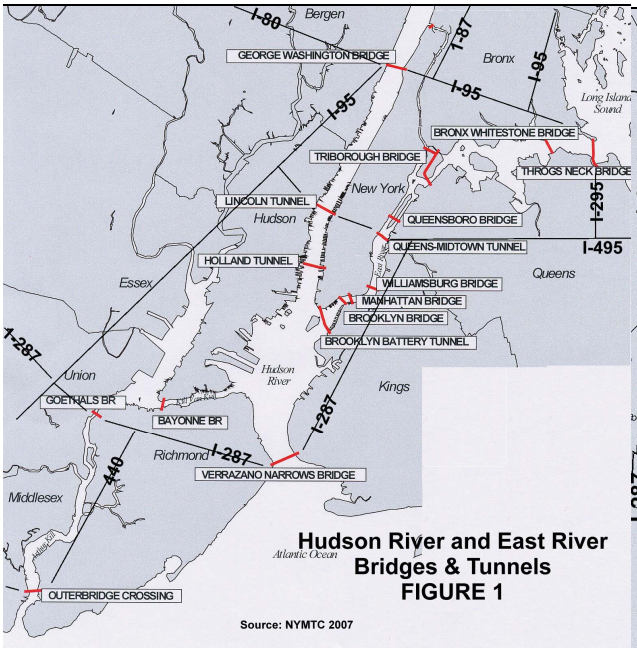
The Hudson River travels north to south. There are eight vehicle crossings of the Hudson River in the Lower Hudson Valley: The Outerbridge and the Goethals Bridge connect New Jersey with both Staten Island and Brooklyn via the Verrazano Narrows Bridge; the Holland and Lincoln Tunnels connect New Jersey with Manhattan, Queens and the Brooklyn local street network via the East River bridges and tunnels; the George Washington Bridge connects New Jersey with Manhattan and Long Island and New England via I-95 (see Figure 1); the Tappan Zee Bridge connects Rockland and Westchester Counties via I-87 and I-287 within New York State; the Bear Mountain Bridge connects Rockland and Westchester Counties via NYS Route 202; and the Beacon Bridge connects Ulster County with Putnam County and Connecticut via I-84 (see Figure 2).

There are three passenger railroads which currently connect New Jersey with Manhattan: the Path (Port Authority Trans-Hudson) Tubes, which terminate at the former World Trade Center; the Path Tubes which terminate at 6th Avenue at 34th Street (Journal Square) via Christopher Street and 6th Avenue; and the NJ Transit Tunnel which terminates at Penn Station, where they share the tunnel with Amtrak. Penn Station services the PATH trains, the Long Island Rail Road, and Amtrak which provides train service to Boston. There is a two track tunnel under the East River servicing only these three passenger railroads.

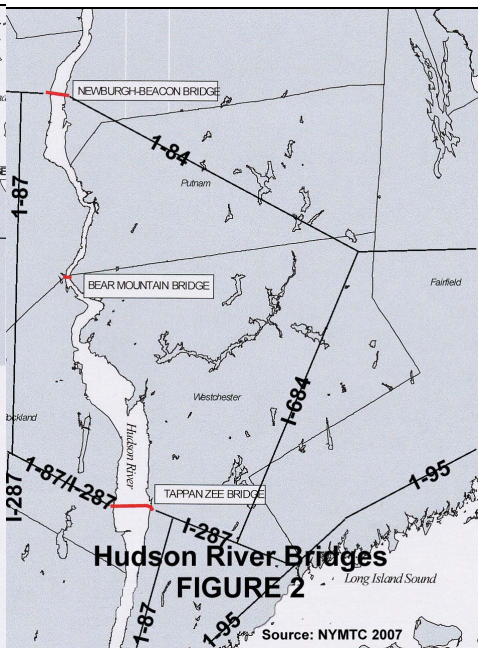
In the New York City Metropolitan Regional Area, the New England Regional Area, and the Long Island Regional Area, most commercial goods are transported primarily by trucks over an increasingly congested highway system across the Eastern

1 Wikipedia re Hudson River and Lake of Clouds
2 Wikipedia re NYS Thruway Route 87

Continental Divide. Building a railroad bridge across the Eastern Continental Divide in the New York City area would not be a viable solution, as the cost of acquiring land rights-of-way through residential and business areas would be prohibitive.




**Hudson River and East River Bridges & Tunnels**
**FIGURE 1**




**Hudson River Bridges**
**FIGURE 2**

Trucks crossing through the Borough of Manhattan, New York City must use the George Washington Bridge and the Lincoln Tunnel. This freight traffic has to transverse local streets in Manhattan to get to the East River Bridges to enter the Boroughs of Queens and Brooklyn, and to the Queens Midtown Tunnel to get to the Long Island Expressway (I-495). Routes I-95 and I-80 go to the George Washington Bridge and continue across Manhattan and the Bronx to the New England Regional Area via the Cross Bronx Expressway (I-95), and also to the Long Island Expressway via the Throgs Neck Bridge to Long Island. The New York State Thruway (I-87) and the I-287 cross the Tappan Zee Bridge and the Westchester Expressway (I-287) to the New England Thruway (I-95). I-95 is the most congested Interstate Route within the United States, and it extends from Florida to Maine (see Figures 1 & 2).

The Hudson River trench extends to the Continental Shelf, which is at least 100 miles into the Atlantic Ocean and over 200 miles up the Hudson River. The Palisades Mountains are located on both sides of the Hudson River, and they range from 400 to 550 feet above sea level from Manhattan Island northward, past West Point. The total depth of the Hudson River trench from the top of the Palisades to the bedrock is about 1,350 feet. The average width of the Hudson River is about one mile, but in Rockland and Westchester Counties the width is 3.2 miles, with two deep trenches. The eastern channel has a water depth of 40 feet, with about 300 feet of silt and clay to the bedrock on the Westchester County side. The western channel has a water depth of 7 feet, with about 800 feet of silt and clay to the bedrock on the Rockland County side.

Just as the Western Continental Divide goes over the Rocky Mountains and thereby restricts the number of viable transportation routes, the Eastern Continental

Divide is at the Hudson River and similarly restricts the number of viable transportation routes. Since the late 1800's bridges and tunnels have been built to ford the Hudson River.

Henry Hudson called this river a ZEE, which in the Dutch language means an ocean going river, which is different from the Dutch language word KILL, which means a fresh water stream or narrow river. Henry Hudson found the estuary of the Hudson River after travelling 150 miles north to Troy, New York. The Hudson River tidal flow averages about 7 feet, and ends where it meets fresh water. When the tide comes in the salt water is pushed up the Hudson River, and when the tide goes out it contains a mixture of fresh and salt water.

The traffic crossing the Eastern Continental Divide has to cross one of the seven bridges and tunnels. This has created bottlenecks along the major expressways and roadway arterials. Table 1 depicts the peak hour traffic, which has already exceeded the capacity of each of the crossings, resulting in traffic congestion lasting for hours. Accidents and vehicle breakdowns cause a ripple effect which has an adverse affect upon every Hudson River crossing, and the resulting traffic congestion can last for over half a day. Truck traffic comprises about 11 percent of the total traffic. The length of each truck is equivalent to 2 to 5 passenger automobiles. In addition, trucks move at slower speeds in heavy traffic and they take longer to accelerate when climbing grades, thus adding to traffic congestion.

In 2007, the New York Metropolitan Transportation Council studied toll collection data on all the bridges and tunnels within the New York City Metropolitan Regional Area for all the Hudson River crossings. Trucks pay a toll on the number of axles on each truck. Table 1 shows the number of tractor trailers in 2007 for each Hudson River crossing. The total truck count was doubled for one way tolls, which were located at each crossing for eastbound traffic only.

The New York and New Jersey Port Authority (NYNJPA) operates the Ports of Newark and Elizabeth, as well as other ports along the eastern shore of New Jersey. These port facilities transfer containers from ships to trucks and trains, and they also receive container shipments from trucks and trains. In the New York Metropolitan Regional Area almost all shipping is done by containers. Each ship delivers about 3,000 to 4,000 containers and receives about 3,000 to 4,000 containers per day. Currently, each container arrives by tractor tailor and leaves by tractor tailor. Only a very small percentage is delivered by train from the south or west. In 2007, about 53,000 containers per day were moved across the Hudson River through the New York City Metropolitan Regional Area, down from 87,000 containers in 1999[3].

Containerization of freight railroads has increased over the past 30 years. Freight from factories and warehouses are placed in containers, which provide security and also reduce handling cost. Railroad intermodal operations during the first 13 weeks of 2010 rose 12 percent to 2.2 million containers, which were shipped all around the country[4]. A number of these containers are delivered to the east coast from the south,

3 "The Tappan Zee Bridge Where Do We Go From Here?" by Robert Hintersteiner, P.E., July 2002
4 Progressive Railroading, May 2010

west and north. Containers that are not shipped overseas head to the New York, New England and Long Island Regional Areas.

**TABLE 1**

**TRACTOR TRAILER TRAFFIC CROSSING THE HUDSON RIVER**

| Hudson River Crossings | Average Annual 2007 Tractor Trailers Traffic [a] | Projected 2020 Average Annual Tractor Trailers Traffic [c] | Projected 2040 Average Annual Tractor Trailers Traffic [d] |
|---|---|---|---|
| Newburgh-Beacon Bridge | 2,446,556 | 2,764,600 | 3,254,000 |
| Bear Mountain Bridge | 47,404 | 53,600 | 63,100 |
| Tappan Zee Bridge | 1,911,630 | 2,160,200 | 2,542,500 |
| George Washington Bridge | 5,627,260 | 6,358,800 | 7,484,300 |
| Lincoln Tunnel | 377,350 | 426,400 | 501,900 |
| Holland Tunnel [b] | 0 | 0 | 0 |
| Bayonne Bridge | 373,406 | 422,000 | 496,700 |
| Goethals Bridge | 1,458,760 | 1,648,400 | 1,940,200 |
| Outerbridge Crossing | 1,045,692 | 1,181,600 | 1,390,800 |
| Total | 13,288,058 | 15,015,600 | 17,673,500 |
| Average Weekday [e] | 53,000 | 60,000 | 71,000 |
| Projected Tractor Trailers [f] | | 73,000 | 96,600 |

a. NYMTC report dated 2007 Truck Toll Volumes October 2008.
b. Since 2005 Truck Traffic were restricted in the Holland Tunnel.
c. Projected Average Annual Tractor Trailer with 1.5 % growth factor for 13 years.
d. Projected Average Annual Tractor Trailer with 1.2 % growth factor for 20 years due to Increase in Port Containers Ships.
e. Total Tractor Trailers/250 days.
f. With a proposed increase to 16 m containers per year (64,000 containers per day) and with about 40 % crossing the Eastern Continental Divide, that would be added by 2040 and half that in 2020.

On the eastern side of the Eastern Continental Divide, all freight traffic must be delivered by tractor trailers. There is only one operating freight railroad crossing in the New York City area, and it goes to Long Island by a freight railroad ferry service from Greenville, New Jersey to Bay Ridge Brooklyn, crossing New York Bay. The Hells Gate Railroad Bridge services Amtrak and the New Haven Railroad, but it has no connection with the MTA Hudson Division. Therefore, containers and freight service must be transported by truck across the Eastern Continental Divide via bridges and tunnels.

There is only one freight Railroad Bridge that crosses the upper Hudson River, called Smith Bridge. It is located south of Albany NY. The railroad tracks that service Smith Bridge are from Buffalo, NY and the Boston Regional Area. There is a freight service from Newark along the western side of the Hudson River, known as the River Line, and it services the northern areas of New York and Massachusetts. It takes about three days for a train to travel from Newark up the western side of the Hudson River, cross over the Smith Bridge, and travel down the eastern side of the Hudson River (Hudson Division) to the Hunts Point Market in the lower Bronx. It takes about two

hours for a tractor trailer to cross the George Washington Bridge during rush hour to go to the Bronx, and more time to go to Long Island and New England via I-95.

In 2006, the NYNJPA recorded 5 million TEUs (20-foot equivalent units) containers that were loaded and unloaded from ships at its port facilities. The NYNJPA has started to expand Port Newark/Elizabeth to accommodate up to 10,000 containers per ship by 2040. This would place an extreme burden upon the existing roadway network and all the bridges and tunnels crossing the Eastern Continental Divide. The current plan for the year 2040 is to expand the existing port facilities to handle at least 16 million containers per year, with 87 percent being transported by truck. Since the existing highway system has already exceeded its capacity, the projected growth in truck freight traffic would further overburden the highway system. The NYNJPA is currently spending $2 billion to lower the Elizabeth Channel to 55 feet from 45 feet.

During the past 10 years the number of truck drivers has fallen due to retirement and difficulty in recruiting new long haul drivers. Drivers employed by trucking companies are paid lower wages than independent drivers. However, independent drivers are paid by the load, and the cost of owning a tractor trailer has increased due to the cost of buying the tractor, maintenance, governmental regulations, fees, tolls, and the cost of fuel. Fuel costs increase when independent drivers are forced to waste time sitting in highly congested traffic.

After the current recession ebbs, the economy will start to improve, and that will increase the number of cars on the highways, as well as increase the need for consumer goods. This situation will increase the amount of freight traffic. The only viable solution to our traffic congestion dilemma is to construct a freight railroad to remove trucks from our regional highways.

A Cross-Harbor Rail Tunnel was proposed under New York Bay from Greenville, NJ to Bay Ridge, NY to replace the ferry service[5]. This would be a single track tunnel at least 12 miles in length using diesel engines. This would accommodate the low volume of freight trains servicing Brooklyn, Queens, and the rest of Long Island. It would also take about one hour to ventilate the tunnel each time a freight train passed through the tunnel.

Currently, the Tappan Zee Bridge Replacement Study has proposed a passenger commuter rail system within the I-87/I-287 Corridor across the proposed new Tappan Zee Bridge, in order to provide a one ride commute from Orange County, NY to Grand Central Station in New York City, via the Hudson Division Railroad. This system would by pass Westchester County, which is located immediately north of New York City. The Study was started in 1990, and after twenty years, is still ongoing.

The author has proposed building a freight rail tunnel paralleling I-287 in Rockland and Westchester Counties that would connect Fairfield, CT, and Nassau County, NY. The author refers to this proposal as the Lower Hudson Valley Metropolitan Transportation System (LHV Metro)[6], which would create a freight and passenger railroad system connecting the national freight railroad network west of the

5 NYC Economic Development Corporation proposed a Cross Harbor Rail Freight Tunnel, Wiki
6 "The Tappan Zee Bridge Where Do We Go From Here?" by Robert Hintersteiner, P.E., July 2002

Hudson River with the entire New England Regional Area east of the Hudson River. These suburban counties are located just north and east of New York City. These counties have undergone increased development over the past 60 years, and they are in need of a belt transportation system connecting Westchester County with Rockland, Nassau, and Fairfield Counties. The Metropolitan Transportation Authority (MTA) operates commuter service only in and out of the New York City Borough of Manhattan, without cross county services.

The basic concept of the "Lower Hudson Valley Metro Transportation System" (LHV Metro) would be to build a four-track subway system, 100 feet underground, from Suffern, NY to Port Chester, NY, connecting all the major residential and central business districts along NYS Routes 59 and 119 (see Figures 3 and 4). The Subway System would also include extensions to the Westchester County Airport; to Mineola in Nassau County; and pass through Stamford, CT (see Figure 5).

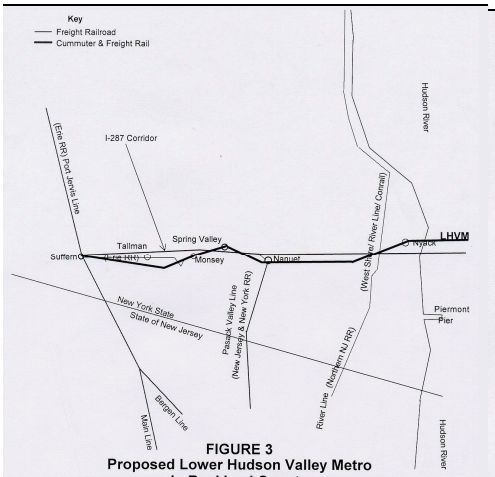



FIGURE 3
Proposed Lower Hudson Valley Metro
In Rockland County
Source: The Tappan Zee Bridge Where Do We Go From Here?

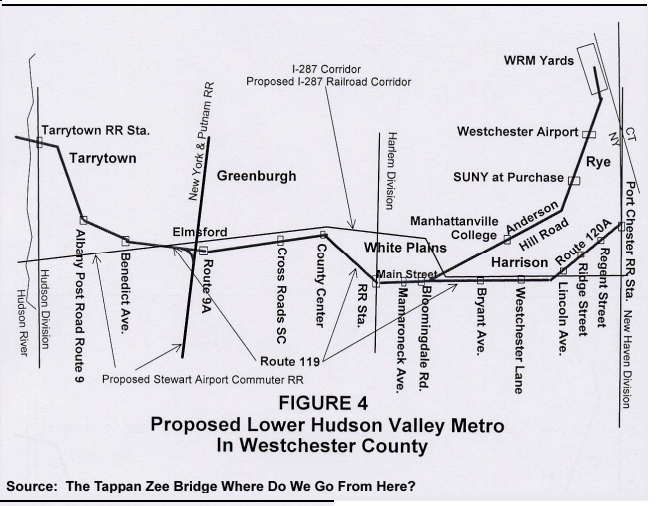



FIGURE 4
Proposed Lower Hudson Valley Metro
In Westchester County
Source: The Tappan Zee Bridge Where Do We Go From Here?

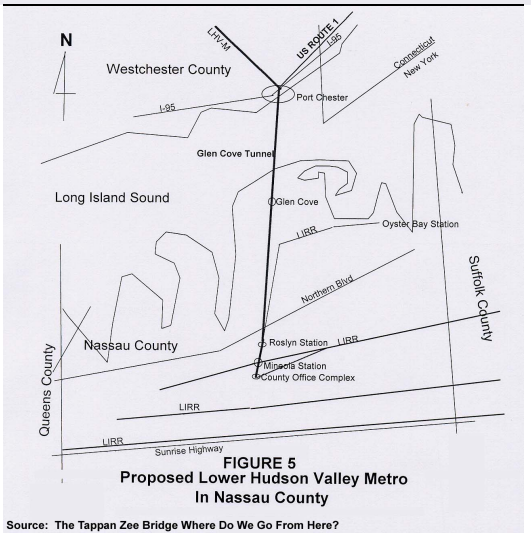



FIGURE 5
Proposed Lower Hudson Valley Metro
In Nassau County
Source: The Tappan Zee Bridge Where Do We Go From Here?

This proposal would reduce traffic congestion, improve transportation of people and goods, and improve the environment throughout the New York City Metropolitan Regional Area by creating a belt freight/passenger railroad system connecting the

suburban counties to the existing railroad networks servicing New York City. This proposed LHV Metro system would divert some 30,000 to 60,000 tractor trailers per day from passing through New York City local streets. With the NYNJPA improved Container Facilities, the traffic congestion and delays across the Eastern Continental Divide would make all the existing bridges, tunnels, and highways impassable by 2040, with continuous grid lock. An average capacity of the proposed tunnel would be about 28,800 containers per day one way, assuming that each train carried 120 containers, double stacked, on 60 rail cars with two engines, and the train length would be about 5,280 feet, traveling at 40 mph (58.8 fps), with a 5 minute headway between trains. However, this proposal would only bring traffic congestion on our highways back to the current 2010 levels.

The proposed LHV-Metro tunnel would have to be multi-model to be financially independent, and in order to accommodate different modes of transportation. Figure 6 shows a typical cross section of the proposed Lower Hudson Valley Metro System Tunnel. This proposed multi-modal system would be completely financially sound and not have to rely on federal and local funding sources. The following multi-modal items should be incorporation in its design and operations:

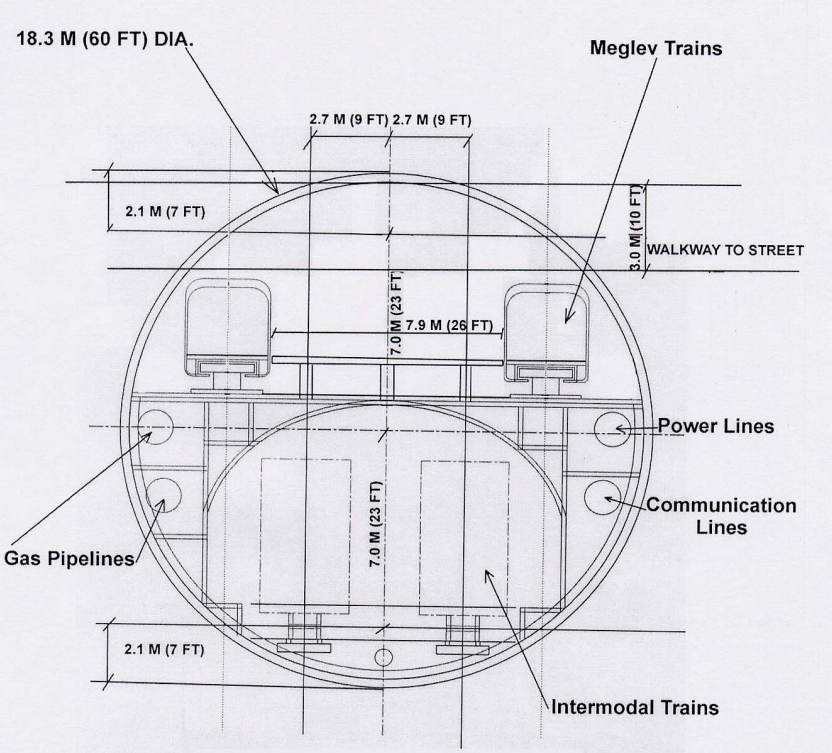



**FIGURE 6**
**PROPOSED LOWER HUDSON VALLEY METRO SYSTEM**
**CENTER PLATFORM STATION**

1. To recover the cost of design, construction, and maintenance, the lower two tracks of the subway tunnel would be leased (a freight toll road) to the freight railroad industry. The fees would be based on the current railroad ton per mile

fee schedule. It is estimated that it would take about 8 to 10 years to design and build using private capital and about 10 to 20 years to retire the bonds.

2. To reduce the cost of land acquisition the proposed tunnel would be built under public rights-of-way. The proposed tunnel would be at least 100 feet under NYS Route 59 in Rockland County, under the Hudson River, under NYS Route 119 in Westchester County, under US Route 1 in Fairfield County Connecticut, and under the Long Island Sound and local arterials to Mineola, NY. The location of the intermodal terminals will have to be determined.
3. The upper two tracks would be for passenger commuter service, with train stations along the tunnel route connecting residential, business and governmental centers along its route. Commuter service will only pay part of the operational cost and freight service will pay the remaining costs. A projected ridership of 25,000 riders/day [6.2 million riders/year] can be achieved during the first year of operations by diverting only 10% of the commuting traffic from the I-287 Corridor and the bridges to Long Island[7].
4. Providing Commuter Rail Service should inspire the political and local support required for this project.
5. A second means of paying for the subway tunnel and passenger operation services would be to lease space for a Natural Gas Pipe Line. The transportation of natural gas through the tunnel would be at a rate based upon millions of cubic feet per hour. In addition, other utilities could rent space for their electrical and communication lines crossing the Hudson River and Long Island Sound, to reducing the cost of providing their own underground facilities.
6. Both the passenger and freight railroad services would have to be electrified. However, since there is not enough power in our existing power grid system to meet the proposed LHV Metro System's operational requirements, it is proposed that the Lower Hudson Valley Metropolitan Transportation System provide its own power plants by using natural gas from the Natural Gas Pipe Line.
7. A third means of paying for the proposed LHV Metro System would be to build cogenerating facilities with excess power, which could then be sold on the Electrical Spot Market. This would provide a yearly return on the entire investment. Currently, there are Federal and New York State funding programs available to build cogenerating facilities.
8. A by product of the tunneling operations would be to dispose of over 30 million cubic yards of earth taken from the tunnel, which could be used to create new landfill areas along the banks of the Hudson River and Long Island Sound, for future development of new residential and commercial properties. An island would have to be built in Long Island Sound in order to ventilate the 12 mile tunnel and to serve as an emergency access and evacuation portal.
9. This proposed island could also serve as an emergency station for the Coast Guard, which would provide faster service to boaters within Long Island sound.

---

7 Ibid.

Another use of this proposed island would be to build a Liquid Natural Gas Ship Dock in the middle of Long Island Sound to bring natural gas into the New England Regional Area.

10. Compressed Natural Gas (CNG) service stations could be built along the LHV Metro System route to provide fuel to more environmentally friendly hybrid and CNG vehicles as an alternative fuel to diesel and gasoline.
11. A Transit Oriented Development (T.O.D) program could be created to address the following needs: retail; affordable housing for existing and future employment needs; median and luxury housing; industrial and commercial areas; recreational and entertainment centers; and multi-use facilities. The estimated 8,000 employees who would be required to operate the LHV Metro System will need addition housing and other facilities to meet their needs.
12. The LHV Metro System could sell Carbon credits to existing carbon producers to defray the cost of operation.

In conclusion, to reduce our future regional traffic congestion problems and the need to transport goods from the western United States and also from the increased PANYNJ port facilities to the entire New England Regional Area, an additional means of crossing the Eastern Continental Divide has to be developed. Otherwise we will have continuous regional grid lock. The Eastern Continental Divide is the choke point between the western states and the entire New England Regional area. During the 19th Century the New England States were the manufacturing centers of clothing and other goods and the railroads sent these goods to the port cities. Today, since our highways are extremely congested, the transportation of goods in a timely manner is becoming a hardship for many businesses. Providing a freight railroad across the Eastern Continental Divide would open the entire region to the development of new markets and the utilization of the abandoned existing railroad rights-of-way, resulting in reduced congestion along our highways.

Also, this proposal for a “Lower Hudson Valley Metro Transportation System” (LHV Metro) would reduce the number of trucks crossing the Eastern Continental Divide that transverse local street networks on both sides of the Hudson River, thus reducing air pollution, truck noise, traffic congestion, and economic loss to individuals and business. The implementation of a “Lower Hudson Valley Metro Transportation System” would efficiently improve our transportation services and our quality of life.

# Wilbur Wright, Hockey, and the History of Flight

William J. Sproule[1]

[1] Department of Civil and Environmental Engineering, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931; phone: 906-487-2568; e-mail: wsproule@mtu.edu

ABSTRACT

The Wright Brothers' invention changed the way we live and air travel has transformed our world. It is a fascinating story on how an injury in a hockey game changed Wilbur Wright's career plans. This paper relates the story and how his career evolved with his brother to build bicycles and then develop the airplane and an aircraft manufacturing company.

## THE EARLY YEARS

Wilbur Wright was born on April 16, 1867 in Millville, Indiana and was the third child of Milton and Susan Wright. His brother Orville was born in Dayton, Ohio, on August 19, 1871. They had two older brothers and a younger sister, Katharine.

Milton Wright settled the family in Dayton at 7 Hawthorn Street where he was editor of a newspaper published by the Church of the United Brethren of Christ. His various jobs as a minister led the family to move frequently, but they didn't sell their Dayton home and they kept returning. The Wright house provided an excellent setting for the children's intellectual and creative development. Their house had two libraries – one for the Bishop's theology collection and a second library with a large and diverse collection on a variety of topics. Orville wrote of his childhood – "we were lucky enough to grow up in an environment where there was always much encouragement to pursue intellectual interests and investigate whatever aroused our curiosity."

Milton often bought the children various souvenirs and trinkets that he found during his travels for the church. In 1878, one of the trinkets that be bought was a toy helicopter. It was based on an invention by French aeronautical pioneer Alphonse Penaud. Made of cork, bamboo, and paper, with a rubber band to twirl its twin blades, it was a little larger than an adult's hand. The brothers later said that this toy helicopter sparked their interest in flight.

Both boys attended high school but they did not receive diplomas. The family moved from Richmond, Indiana back to Dayton in June 1884, less than a month before Wilbur would have graduated from high school. Exhibit 1 is a photograph of Wilbur as a teenager. He was an excellent student and the next year he enrolled at Dayton Central High for additional studies to better prepare himself for Yale. Wilbur planned to be a teacher while his parents hoped that he would become a minister. He was remembered as an outstanding athlete and he played on the high school football team and was one of the swiftest runners in the school.

Exhibit 1 – Photograph of Wilbur Wright as a teenager, 1884
(*Wright State University Collection*)

## HOCKEY CHANGES WILBUR'S PLANS

Suddenly, all of that changed. In the winter of 1885-86, Wilbur was accidently stuck in the face by a stick while playing hockey with friends on a pond near Dayton. It resulted in the loss of his front teeth. A few weeks later he began to experience nervous palpitations of the heart. The details of the accident and its aftermath are so sparse that one cannot be certain when it occurred or to what extent of Wilbur's injuries actually were. Milton stated that the initial injury was minor but he did not see his son for some weeks after the accident. It was clear that the accident was less worrisome that the heart palpitations and the digestive complications that developed in its wake. It was a common chain of events to many nineteenth-century families - a sudden accident, an apparently simple illness, followed by the appearance of far more serious complications leading to lifelong debility.

The accident was clearly a turning point for Wilbur. It bought an end to any serious talk of going to college and marked the beginning of a period of withdrawal, depression, and self doubt. Friends and neighbors wondered at the sudden transformation of an active and athletic young man into a housebound person. The family insisted on a period of extended rest for Wilbur when the first signs of serious complications developed. But the accident, complications, and the long

recuperation left an indelible mark on Wilbur. In his own mind, he was now a potential invalid. Wilbur simply chose to ignore his own problems and devoted himself to nursing his mother. Susan Wright had contacted tuberculosis in 1883 and by 1886 she had become a helpless invalid. Milton travelled as part of his duties as bishop and his two older brothers had left home. Wilbur, whose own plans had been forestalled by ill health, stepped into the breach to look after his mother and younger brother and sister, Orville and Katharine. During that time he spent many hours reading and studying books in their family library.

Wilbur was clearly emerging from his shell by the spring of 1889, but on reflection the time had been well spent. He knew that he was now as read as any college graduate, and that he had the makings of a clear and confident writer and speaker. The long hours spent nursing his mother had made him feel needed and useful once again. Susan died on July 4, 1889.

## EARLY BUSINESS VENTURES

Orville dropped out of high school after his junior year to start a printing business in 1889. Later that year, Orville began publishing a weekly newspaper, *The West Side News*, and recruited Wilbur to write and be the editor for the newspaper. In 1890, they started a daily newspaper, *The Evening Item*, but the competition from a dozen other Dayton newspapers was intense and their newspaper lasted less than four months. The brothers returned to the far less risky business of operating a job printing plant. In 1892 they began to cast about for a business that they could run in addition to the print shop. By fall 1892, cycling had become a shared passion and they opened a bicycle repair shop and started to sell new bicycles in what would become known as the Wright Cycle Company. A few years later they started to manufacture bicycles. Their bicycle work lead them to another exciting area of science and invention of that era – aviation. Wilbur had always read extensively about the adventures of the early aviation pioneers and in the summer of 1896 when Orville became ill with typhoid fever and was unable to work, he began to read more and more about aviation and flight. That year the German aviator Otto Lillenthal died in a glider crash but the brothers were convinced that his pioneering work showed that manned flight was possible. The brothers interest in flight was ignited, and they set about to learn everything that they could about the subject, gathering, and reading whatever they could, and later designing their own experiments. It is reported that Wilbur wrote the Smithsonian Institution for publications on aeronautical subjects and developed a close friendship with Octave Chanute, a French civil engineer and aeronautical pioneer. In 1899, Wilbur wrote to Willis L. Moore, Chief of the U.S. Weather Bureau, asking about high wind conditions throughout the country to conduct flying experiments. The first place on the list that Moore sent back was Kitty Hawk, North Carolina. During the next three years, the brothers experimented at Kitty Hawk with kites and gliders.

## THE FIRST TO FLY

At the start of the 20th century, several efforts were underway to become the first to fly. Most inventors of the era were impulsive and undisciplined as they would build an aircraft one day and try and fly it the next day with either disastrous or unsuccessful results. The Wright brothers were much more scientific and methodical in their approach. As bicycle mechanics, they believed in testing before proceeding with further advancements. They built a wind tunnel to test aircraft wing design and they conducted almost a thousand flights in gliders before they felt ready to begin production of a motor-powered flyer.

After years of development, the brothers were ready to test their first powered plane, the Wright Flyer. On the morning of December 17, 1903, Wilbur and Orville stood on the beach at Kitty Hawk, North Carolina, and actually flipped a coin to see who would be first to pilot the Flyer. Orville won. The first flight lasted only 12 seconds and covered only 120 feet, but it was historic – it was the world's first powered, controlled flight. The next two flights covered 175 and 200 feet by Wilbur and Orville respectively. The fourth flight of the day with Wilbur on board traveled over 852 feet and lasted 59 seconds. The brothers were also interested in photograph and arranged for many of their flights to be captured on camera. Exhibit 2 shows the famous photograph of Orville's first flight.

After the successful flights at Kitty Hawk, the brothers returned to Dayton and continued their experiments at Huffman Prairie, a cow pasture 8 miles northeast of Dayton. They continued to refine their invention until it was considered a practical airplane. They made the first public demonstration of this machine to a group in Dayton in October 1905 and were awarded a patent for their flying machine in 1906 (Patent Number 821-393). Over the next few years, they sold airplanes to the U.S. Army and to a French syndicate, and demonstrated their invention throughout the United States. In 1908 Orville made the first flight of over one hour at Fort Myer, Virginia, in a demonstration for the U.S. Army which subsequently made the planes the world's first military airplanes. That same year, Wilbur made over 100 flights near Le Mans, France; the longest one on December 31st, a record flight of 2 hours, 19 minutes.

In 1908, the brothers built the Wright Model A - a two-seater with an improved control system and a more powerful engine. On September 17, 1908, Thomas Selfridge became the first person to be killed in an airplane when a propeller failure caused the crash of a Wright Model A plane that was being piloted by Orville during military tests at Fort Myer, Virginia. Orville was injured in the crash but he recovered.

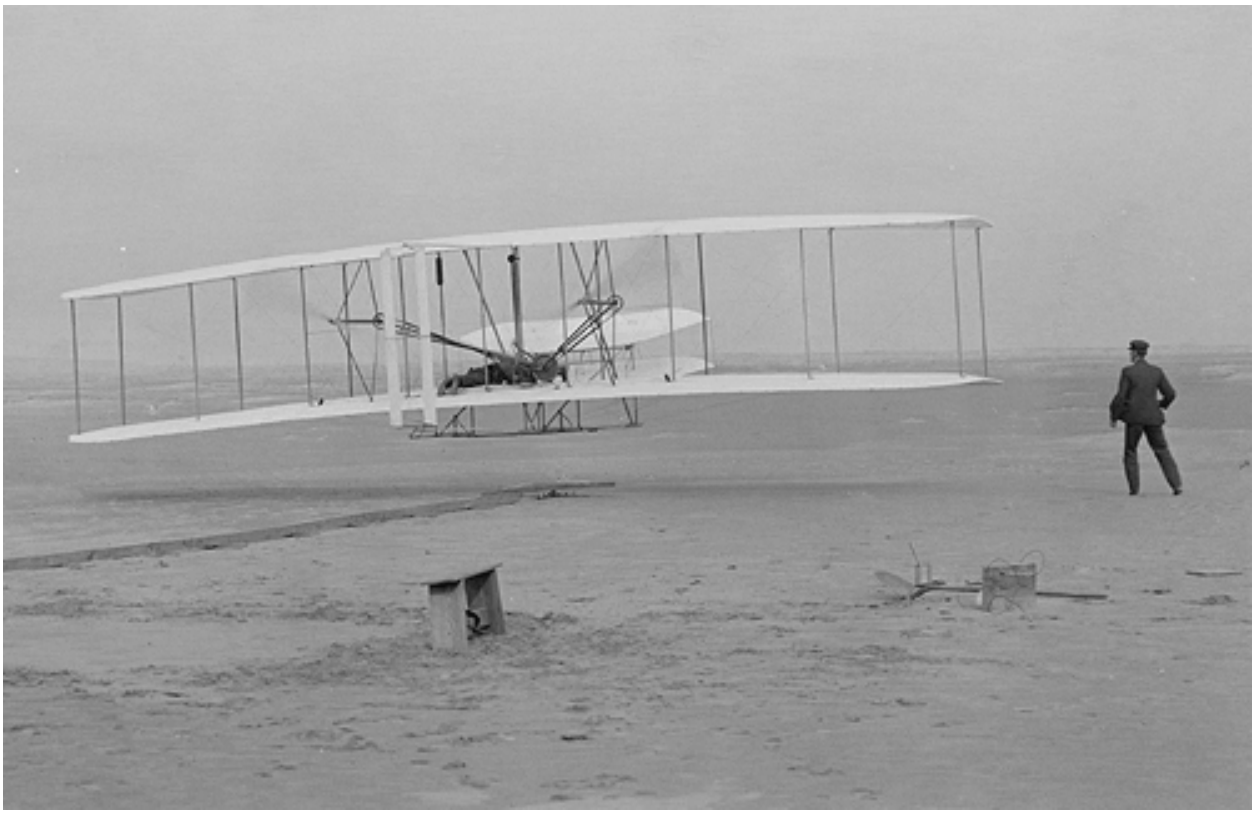


Exhibit 2 – Famous photo of the first flight at Kitty Hawk, North Carolina, December 17, 1903, Orville Wright was the pilot (*Library of Congress*)

## THE WRIGHT COMPANY

In 1909 the brothers organized a company to manufacture airplanes – the Wright Company. Wilbur was President and Orville was Vice President. They also began to file patent infringement suits against other manufacturers that were using their methods for aerodynamic control and Wilbur became the designated "expert witness" in these cases and traveled frequently to give testimony. The company headquarters was in New York City and they set up an airplane factory in Dayton and a flying/test flight field at Huffman Prairie (Huffman Field is now part of Wright-Patterson Air Force Base). In 1910 they founded the Wright Exhibition Company to demonstrate their planes and they started a flight school to train pilots. Orville was in charge of instruction and he trained numerous pilots. It is also believed that the Wright Company transported the first commercial air cargo on November 7, 1910 by flying two bolts of dress silk from Dayton to Columbus for the Moorehouse-Marten Department Store.

Exhibit 3 – Photo of Wilbur (left) and Orville Wright on their front porch, 7 Hawthorne Street, Dayton, Ohio, June 1909 (*Library of Congress*)

Neither brother married. Wilbur once quipped that he did not have time for a wife and an airplane. Wilbur became ill on a trip to Boston in April 1912 and after returning to Dayton he was diagnosed with typhoid fever and died on May 30 in the Wright home at the age of 45. Orville succeeded to the presidency of the Wright Company on Wilbur's death.

Orville sold the company in 1915 and it was merged with the Glenn L. Martin Company in 1916 to become the Wright-Martin Aircraft Company. Martin was an aviation pioneer that started out building military trainers in California a few years earlier. The merger did not go well and Martin left to form a second Glenn L. Martin Company in 1917. Over the years, the second Martin Company employed many of the founders and chief engineers of the American aircraft industry. Martin actually taught William Boeing how to fly and sold him his first aircraft. The second company went through several mergers and today exists as Lockheed Martin.

When Martin left Wright-Martin the company was renamed Wright Aeronautical and in 1929 it merged with Curtiss Aeroplane and Motor Company and other companies to become the Curtiss-Wright Corporation. By the end of World War II, Curtiss-Wright was the largest aircraft manufacturer and was the second largest company in the United States (behind only General Motors) but many believe that they failed to make the transition to the design and production of

jet aircraft. Today it is a shadow of its former self and it has become a component manufacturer specializing in actuators, controls, values, and metal treatment.

Orville made his last flight as a pilot in 1918 and retired from business to become an elder statesman of aviation serving on various boards and committees. He was a member of National Advisory Committee for Aeronautics which eventually became the National Aeronautics and Space Administration (NASA), and was elected as a member of the National Academy of Sciences in 1936. Orville died of a heart attack in Dayton in 1948.

Today, the Wright Flyer is on display at the Smithsonian National Air and Space Museum in Washington, DC. The Wright's childhood home, 7 Hawthorne Street, and their bicycle shop were purchased by Henry Ford in 1936 and were moved to Greenfield Village in Dearborn, Michigan, and placed side by side where they are available for tours.

## HOCKEY TEAM TRAVEL

Train travel was the common method of transportation for early hockey teams. The first hockey team to fly was the New York Rangers. They hired the Curtiss-Wright Corporation to fly them to Toronto for a game against the Maple Leafs on December 13, 1929 (The Rangers lost 7-6).

Today, charter buses are probably the most common transportation mode for hockey teams although they will often fly for longer trips. The National Hockey League (NHL) continued to travel by train until the late 1960s when the league was expanded to include west coast teams. NHL teams then started using commercial air service for longer trips and charter buses for shorter destinations. Today, most NHL teams travel using charter aircraft and there are a few teams that own or share ownership of an aircraft or lease an aircraft for the full season. An average NHL team will fly over 40,000 miles per year during the regular season.

## CONCLUSION

The Wright brothers always presented a unified image to the public, sharing equally in the credit for their invention, however biographers noted that Wilbur took the initiative. Author James Tobin asserts, "it is impossible to imagine Orville, bright as he was, supplying the driving force that started their work and kept it going from the back room of their shop in Dayton to conferences with capitalists, presidents, and kings. Wilbur was the leader from beginning to end. Although Wilbur's hockey injury did not appear especially severe, he became withdrawn and abandoned his plans to attend Yale and become a teacher. Had he enrolled, his career would have taken a very different path. Hockey did have a major impact on Wilbur Wright and the history of flight.

ADDITIONAL READINGS

Crouch, Tom D. (1989), *The Bishop's Boys – A Life of Wilbur and Orville Wright*, W.W. Norton: New York, NY.

Kelly, Fred C. (1943), *The Wright Brothers*, Fred C. Kelly, Harcourt, Brace and Company: New York, NY.

Tobin, James (2003), *First to Fly*, Free Press/John Murray Publishers: London, UK.

U.S. Department of the Interior (1976), *Wright Brothers*, National Park Service Historical Handbook Series No. 34, Washington, DC.

*Wilbur and Orville Wright – Dreams of Flying* (2001), Biography/A&E Television Networks, New York, NY – DVD/VHS

# The Collaborative Process for Developing Project Definition Documents for the Metropolitan Washington Airports Authority

By Geoffrey Baskir, C.M., AICP, LEED[1], M.ASCE and William Lebegern, P.E[2], M.ASCE

[1] Supervising Airport Planner, PB Americas, Inc., 465 Spring Park Place, Herndon, VA, 20170, PH (703)742-5872; e-mail: baskir@pbworld.com
[2] Manager of Planning, MA-32, Metropolitan Washington Airports Authority, Reagan National Airport, Washington, DC, 20001, PH (703)417-8161; e-mail: william.lebegern@mwaa.com

## Abstract

The conventional approach to preparing programming documents for new facilities is a process that tends to favor "stove-piping". Once a facility is planned and a program definition document has been prepared, often the project is turned over to design, and once designed, the project is turned over to construction. Little input is derived from the designers and constructors during the planning process, and during design, the input from the planners is thought to end. Likewise, when a project that enters construction, there is less input from the designers and the planners. The Metropolitan Washington Airports Authority (Airports Authority) has sought to invest all stakeholders in the success of a project by developing a collaborative planning, design and construction process that involves all three disciplines from the conceptualization of the project through its turnover to users. The intent of the process is to make sure that the completed project meets the goals and objectives first laid out in the planning phase, and to assure that all project costs are accounted for and that goals for schedules and budgets are identified early in the process and adhered to through completion. This collaborative effort has been used with success on some initial smaller projects at both Reagan National and Washington Dulles International Airports, and is now becoming the standard process for execution of larger projects.

## Need

The Airports Authority is in the final stages of an ambitious Capital Construction Program (CCP) that began in 1988 and has placed over $7 billion worth of improvements at both Reagan National and Dulles International Airports in Washington, DC. During the course of the CCP, lessons learned from the delivery of the more than 200 individual projects of the CCP have resulted in a change in the way

the Office of Engineering (MA-30) prepares the documents that define individual projects. The intent is to improve the process to avoid delays and additional costs associated with unforeseen conditions, disagreements about project requirements or inadequately defined project elements. Starting in 2006, MA-30 concluded that a process for preparing Project Definition Documents that brought together all stakeholders was needed, and that the collaborative framework established at the beginning of project definition needed to be carried through design to construction and project activation.

**Background**

The Airports Authority was created in 1987 to oversee the redevelopment and operation of Reagan National and Dulles International Airports. The Authority is structured with a Board of Directors and a Chief Executive Officer overseeing four major divisions: Reagan Washington National Airport (operations, maintenance, and administration), Washington Dulles International Airport (operations, maintenance, and administration), Public Safety (police and fire/emergency medical), and Consolidated Functions (communications, finance, engineering, marketing, legal, human resources, information technology).

Consolidated Functions' Office of Engineering (MA-30) promulgated a Capital Construction Program (CCP) and hired a program management team, the Parsons Management Consultants joint venture to oversee the CCP's completion as an extension to the Authority's Engineering staff. They supplement the staffs of the Planning (MA-32), Design (MA-34), Construction (MA-36) and Codes and Environmental (MA-38) Departments of the Office of Engineering. The Project Controls functions – scheduling, budgeting, cost control, safety and quality assurance – were provided by the program management team. The primary responsibility for preparation of Project Definition Documents falls to the Planning Department, with the assistance of the planners of the program management team.

**Project Definition - Elements**

The documents used to program facilities at Reagan National and Dulles International Airports have evolved over time. Three different documents have been used – project definition documents, program criteria documents and point papers. Point papers are used to address specific questions related to the planning, development or operation of new facilities and typically consist of an issue, an analysis process and a recommendation. They have been prepared for a diverse array of projects and were used most effectively to identify and lay out sites for facilities in the North Area of Dulles Airport, including a commissary, salt and sand storage facility, bus maintenance facility, police and fire station, service station and taxicab staging area.

Program Criteria Documents (PCDs) are comprehensive documents used to define complex projects and to outline the guidelines for their development. They typically provide information and direction covering all disciplines, including planning, schematic design, architectural, civil, utilities, interior MEP, special systems, scheduling, phasing and budgets. A PCD was prepared in 1993 to guide the design and construction of the Tier 1 Concourse at Dulles Airport, also known as Concourse

B. It was delivered to the selected Architect/Engineers and guided the development of each segment of the Concourse.

The Project Definition Document has also been an evolving document. PDD's produced in the 1990-2000 timeframe varied in content, with larger projects having more comprehensive PDDs than smaller projects. In 2001, the PDD format was standardized and included the following information:

- Purpose and statement of need;
- Project scope and requirements, including site identification, space planning, operational characteristics and requirements;
- Existing site conditions, including topography, drainage and utility availability;
- An environmental checklist;
- Codes and standards, including Federal guidelines;
- Budget (based on cost estimate);
- Schedule and phasing;
- Safety and security;
- Issues to be resolved during design.

**Case Study – The Legacy Process for PDD Preparation**

The original process for preparing PDDs usually began with a kickoff meeting consisting of the planners tasked with preparing the PDD and the expected end users of the facility. The PDD itself was produced with one-on-one input involving the planner tasked with preparing the document and those subject matter experts that the preparer included in the process. Upon completion, the document would be circulated for review, and if comments were received, then a review meeting would be scheduled to resolve the comments. The PDD would then be finalized, incorporating any revisions generated by the comments. When completed, a PDD was typically signed by the preparer, the Planning Department task lead, the Manager of Planning and the Vice President of Engineering prior to turnover to the Design Department for preparation of construction documents. Turnover of the PDD to Design typically signaled the end of active involvement of the Planning Department in the project (*Fig. 1*). Design and Construction followed in successive steps.

This process had certain drawbacks. In some instances, initiation of the PDD process did not begin until the project was approved for inclusion in the capital improvements budget, increasing the risk that the project scope would exceed the budget. Circulation of the PDD for comments often would generate little or no participation from the reviewing departments, increasing the likelihood of subsequent disagreements among the departments over the PDD scope and budget, with the result that changes would be made during design that would modify or abandon the original PDD objectives. These changes would often generate cost increases and cause delays in the development schedule.

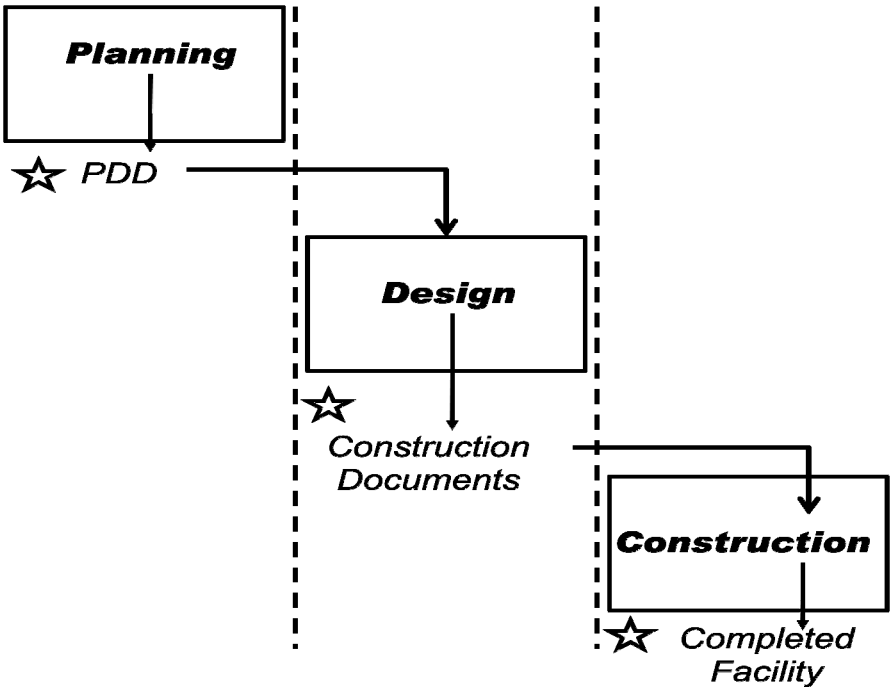



Figure 1. The Successive Step Process for Project Delivery

As an example, a PDD was begun for the extension of a sanitary sewer line from a security gate at the south boundary of Dulles Airport to an existing sewer trunk line discharging into a main flowing to a regional treatment plant off-airport. Conceived as a two-inch force main transitioning to an 8-inch gravity main, the sewer was intended to be designed and built in a year at a cost under $500,000. The project was entered as a $500,000 line item in the Authority's Capital Construction Program (CCP) budget in FY 2004. However, before the PDD could be circulated for signature, a scope increase was proposed that provided a connection to a complex of construction trailers along the path of the main. As a result, the cost of the project increased to over $800,000, and subsequent modifications to the scope increased the cost to over $1 million. Unforeseen design costs, the result of an extensive geotechnical investigation along the proposed route, drove the cost to $1.2 million. This additional cost required additional funding in the CCP budget. When completed, the sewer was three years behind the scheduled delivery date shown in the PDD, at a cost that was more than double the original estimate.

Another Project Definition Document, for construction of a shelter for multi-function snow equipment, was constrained by a capital budget that had been submitted prior to preparation of the PDD. In order to meet the budget, the resulting building was proposed as a three-sided lean-to without doors or climate control and accessed by a pathway made of crushed aggregate. Design and construction of the shelter has been placed on hold, awaiting procurement of the vehicles to be sheltered inside.

**Changes to the Process**

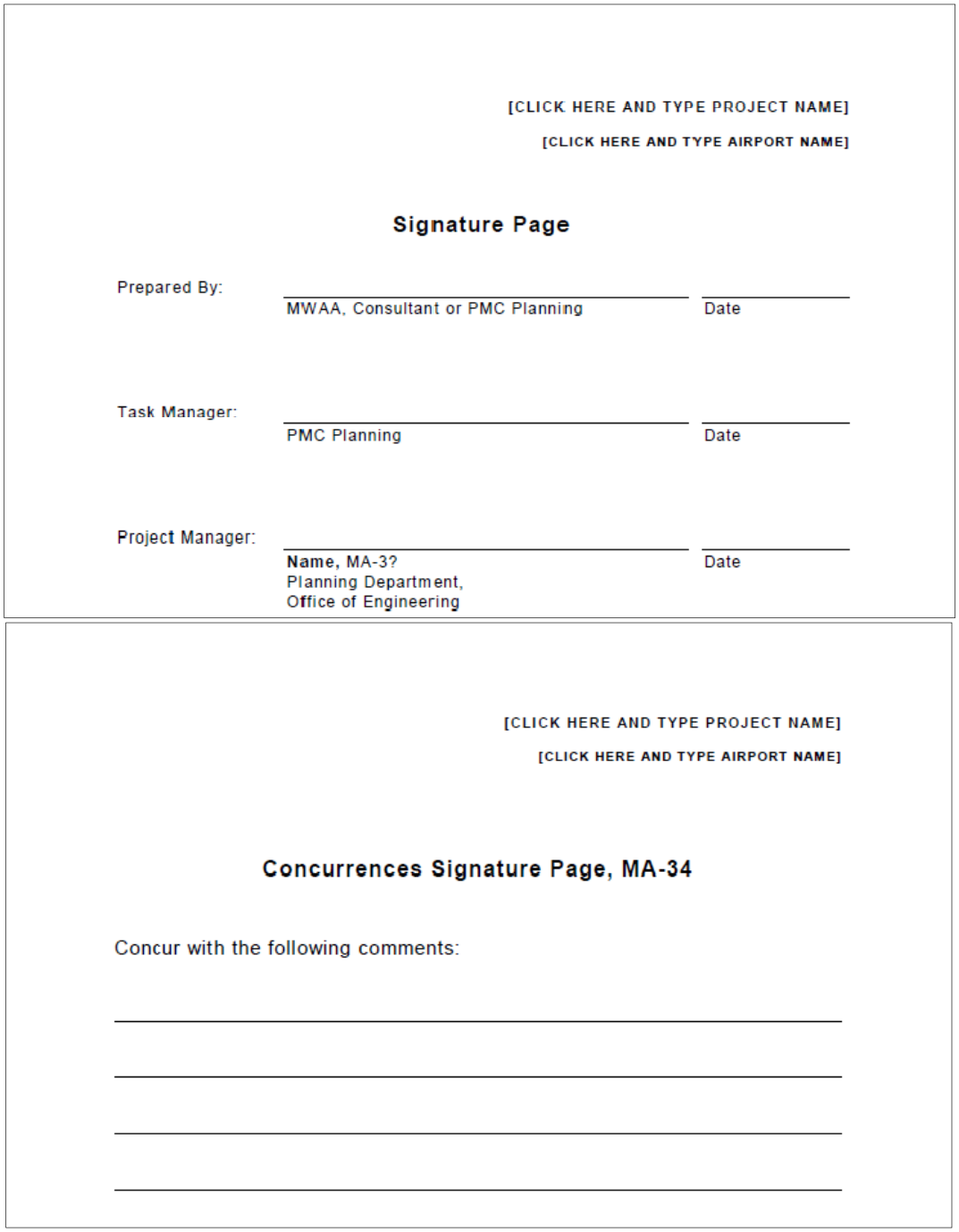
Improvements to the process for preparing Project Definition Documents were intended to assure that project scopes were defined in a way that met users' expectations and to minimize risks affecting budgets and schedules. The first change

was, to the extent practicable, to schedule the PDD preparation so that a defined scope and schedule for the proposed project would be confirmed in time for the Authority's capital budget process. Capital budgets are prepared by the Office of Engineering (MA-30) for submission to senior management in September of each calendar year, and the budget is approved by the President of the Authority and the Authority Board prior to the beginning of the fiscal year in January. Individual budget requests are therefore due in August, so that MA-30 can prepare a list of items that will be included in the overall capital budget. Striving to develop a draft of the PDD, including a cost estimate verified by MA-30, in advance of a budget request provides a greater assurance that the adopted budget will better reflect the expected cost of the project.

The list of stakeholders invited to participate in the PDD preparation process was also expanded. PDDs prepared prior to 2006 were typically prepared within the Planning Department (MA-32) with the input of the Design Department (MA-34), the staff at the airport where the project would be located and the user community as appropriate. With the implementation of the new process for PDD preparation, the list of stakeholders was increased to include the Construction Department (MA-36), the Codes and Environmental Department (MA-38), and cost estimators, schedulers, safety and quality assurance personnel provided by the program management team.

With the additional stakeholders came revisions to the PDD to acknowledge the participation of all stakeholders in its creation. First, a listing of all participants in the preparation process was appended to the PDD. Second, a list of all issues identified and resolved during the preparation process was included in the PDD. These two changes informed the approving departments that a consensus exists on the recommendations of the PDD. Finally, concurrence pages were added to the PDD that included signatures from the managers of the Design Department, Construction Department and Codes and Environmental Department of the Office of Engineering, as well as the Project Controls manager of the program management team (*Fig. 2*). Their signatures acknowledge the agreement of the implementing departments with the completeness of the PDD and concurrence with the recommended plan of action, schedule and budget. The Codes and Environmental Department is also vouching that all Code issues and environmental concerns that may be raised by outside agencies have been addressed and resolved. The overall intent is to assure that there is a consensus on the particulars of the PDD among all departments within the Office of Engineering that is documented in writing. That consensus will apply when the PDD is turned over to the Design and Construction Departments for implementation. The concurrence of the Construction Department assures that the project is constructible.

The Project Controls staff of the program management team has become more involved in the PDD preparation process. It is their responsibility to review and verify the accuracy of all cost estimates, including those prepared by the subconsultants tasked with preparing the PDD. They must also assure the PDD's conformance with Authority budgets established for the project and agree to the schedule proposed for implementing the project. Their role will revert back to the Airports Authority upon completion of the Capital Construction Program.



[CLICK HERE AND TYPE PROJECT NAME]

[CLICK HERE AND TYPE AIRPORT NAME]

**Signature Page**

Prepared By: _____ _____
MWAA, Consultant or PMC Planning — Date

Task Manager: _____ _____
PMC Planning — Date

Project Manager: _____ _____
**Name**, MA-3?
Planning Department,
Office of Engineering — Date

[CLICK HERE AND TYPE PROJECT NAME]

[CLICK HERE AND TYPE AIRPORT NAME]

**Concurrences Signature Page, MA-34**

Concur with the following comments:

_____

_____

_____

_____

Figure 2. PDD Signature and Concurrence Pages

**Benefits Expected**

The Authority's Office of Engineering expected that the changes implemented to the PDD process and the subsequent development of design and construction documents would result in the timely delivery of projects that more reliably met the Authority's expectations at a cost that was more predictable and within budget. Among the immediate benefits expected were:

- Fewer disagreements among the departments regarding the contents and recommendations of the PDD and greater "ownership" of the PDD recommendations by all stakeholders. Because all departments participate in the PDD development process and provide their signed concurrences to the recommendations, they are on record as agreeing to the implementation of the PDD's recommendations and are therefore more accountable as the project shifts from planning to design and then to procurement and construction;
- The collaborative approach to project implementation that begins with the PDD also assures continuity (*Fig. 3*). The handoff of a project from planning to design to construction does not end the participation of the planning department and then the design department in assuring the project's ultimate success. The planners will remain involved in reviewing the design documents to check for conformance with the PDD's recommendations and are also able to answer any questions the designers have about the PDD and its recommendations;

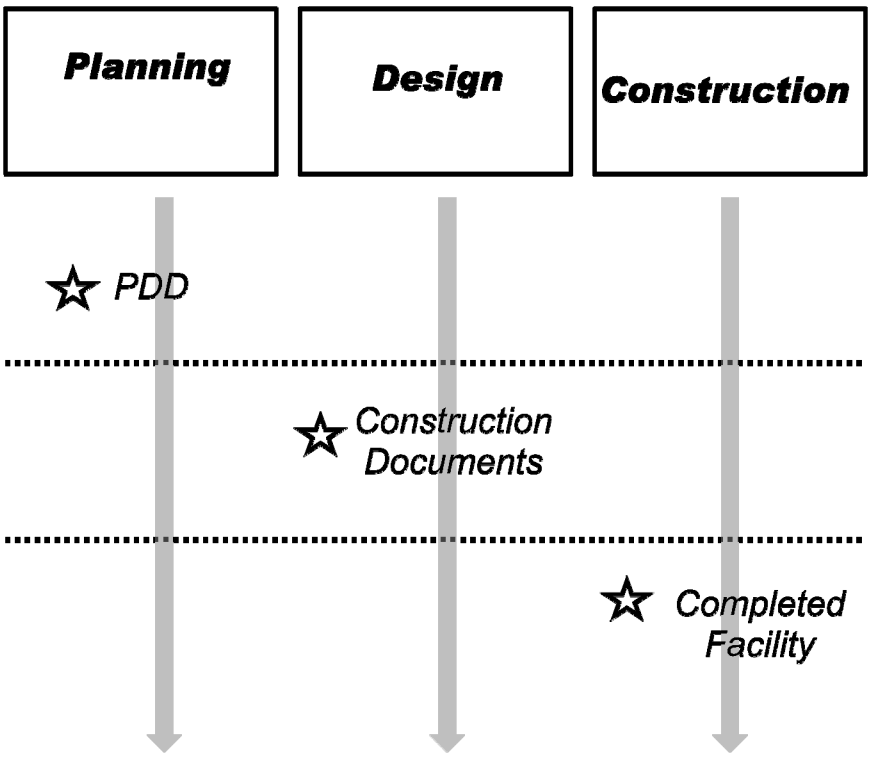
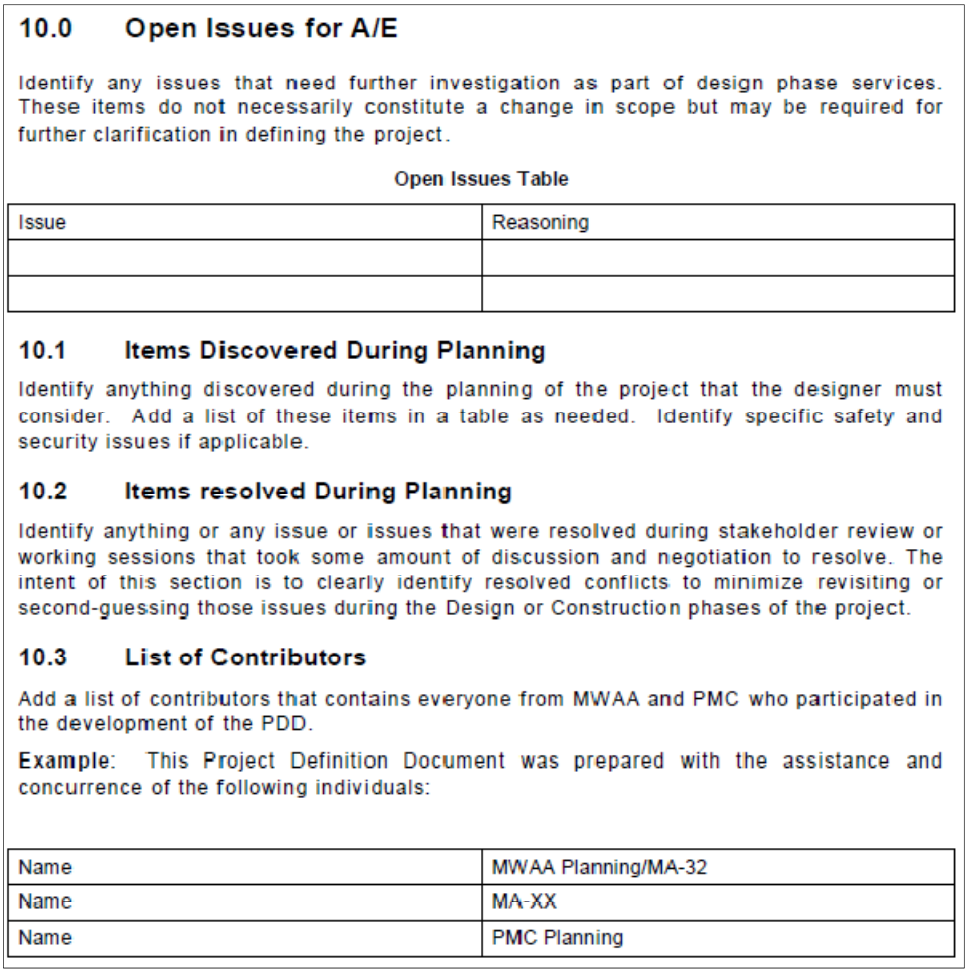



Figure 3. The Collaborative Process for Project Delivery

- Internal reviews and verification of the project's expected cost are expected to reduce the number of negative budget "surprises". The participation of all departments is expected to assure that existing field conditions are accurately portrayed, that project scopes are clearly defined and that no unforeseen contributors to the expense of a project arise.

More accurate estimates of project costs are expected, resulting in fewer budget busts;

- With all stakeholders participating, there should be fewer unidentified issues that would contribute to changes in a project scope following development of the PDD and lead to potential increases in cost. The PDD specifically provides a section where issues among stakeholders that arose during preparation are identified and the method by which each issue was resolved is described (*Fig. 4*);



**10.0 Open Issues for A/E**

Identify any issues that need further investigation as part of design phase services. These items do not necessarily constitute a change in scope but may be required for further clarification in defining the project.

**Open Issues Table**

| Issue | Reasoning |
|---|---|
| | |
| | |

**10.1 Items Discovered During Planning**

Identify anything discovered during the planning of the project that the designer must consider. Add a list of these items in a table as needed. Identify specific safety and security issues if applicable.

**10.2 Items resolved During Planning**

Identify anything or any issue or issues that were resolved during stakeholder review or working sessions that took some amount of discussion and negotiation to resolve. The intent of this section is to clearly identify resolved conflicts to minimize revisiting or second-guessing those issues during the Design or Construction phases of the project.

**10.3 List of Contributors**

Add a list of contributors that contains everyone from MWAA and PMC who participated in the development of the PDD.

**Example**: This Project Definition Document was prepared with the assistance and concurrence of the following individuals:

| | |
|---|---|
| Name | MWAA Planning/MA-32 |
| Name | MA-XX |
| Name | PMC Planning |

Figure 4. Documentation of Consensus in a Typical PDD

- Expansion of the scope of the PDD is also expected to reduce the number of unresolved issues handed over from the planners to the designers and further reduce the likelihood of project changes that would affect schedules and budget. New sections on quality control, regulatory issues, safety, security, sustainability and constructability, added in 2006, provide additional information to the designers who receive the PDD.

**Case Study – PDD Preparation for an RON Apron**

One of the first projects for which a Project Definition Document was prepared collaboratively was a hardstand apron on the site of a former Aircraft Rescue and Fire Fighting Facility (ARFF) at Reagan National Airport (DCA). This apron was intended as a replacement for three aircraft parking positions lost when a new ARFF was constructed south of the original building. In this instance, the PDD was prepared after the budget for the project had been established. *Figure 5* shows the apron.

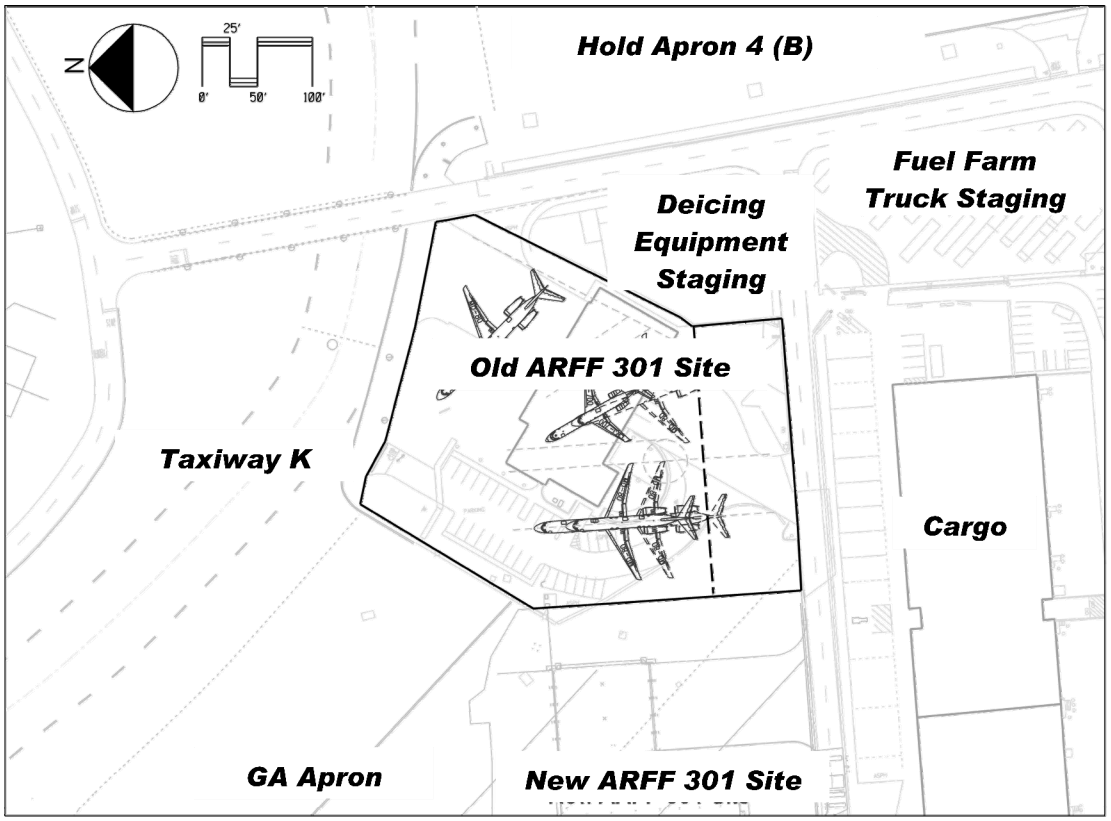



Figure 5. Layout of the RON Apron at Reagan National Airport

Participating in the process were Reagan National's Operations (MA-110) personnel, who provided figures on the number and type of positions required for aircraft remaining overnight (RON), Engineering and Maintenance (MA-120), who provided information on site conditions, and representatives of Properties and Facilities (MA-130) for the Airport and the Metropolitan Washington Airlines Committee. The Design Department (MA-34) and the program management team each provided civil engineers with knowledge of the existing pavements and the locations of utilities. The Construction Department (MA-36) was represented by the designated Contracting Officer's Technical Representative (COTR) and the resident engineer, while the Codes and Environmental Department provided experts on permitting issues. The program management team also was represented by cost estimators and schedulers.

A kickoff meeting was held at the beginning of the PDD preparation process at which the issues particular to the apron design were laid out for the stakeholders. The planner tasked with preparing the PDD described the project goals and objectives,

discussed site conditions and solicited feedback from meeting participants on sources of data and major development issues. The limited project budget and its impact on the scope of the facility was highlighted, and the estimators were apprised of the need to coordinate with the civil engineers to determine all cost elements and identify if the project scope would fit under the funding limits.

After the kickoff meeting, the task manager for the PDD was in contact on a regular basis with all stakeholders as information for the PDD was gathered. The civil engineers and the estimators coordinated the estimate verification process together, and an environmental consultant was brought in to prepare a standard Environmental Checklist. Constructability issues were discussed with the resident engineer and the COTR. When the initial cost estimates for the apron exceeded the defined funding levels, options were explored for modifying the scope of the apron paving while accommodating the two RON positions that were desired by MA-110. In the end, a small reduction in the full-strength pavement area was proposed, and the impact on costs was again reviewed with the cost estimator to assure conformance with the available budget. In the end, the designed cost of the project, at $6.0 million, did not exceed the estimate provided in the PDD, and the cost at completion, aided by favorable construction prices, was $4.0 million.

The design of the apron benefitted from the information included in the PDD, which included recommendations on the relocation of a facility for dispensing of deicing materials, recommendations on storm drainage to allow for future installation of devices to capture spent glycol and information about how to provide access to the jobsite for construction materials that would be brought in from outside the airport. Among the most significant differences between the PDD recommendations and the project design were slight adjustments in the aircraft layout and the provision of a nose-wheel tether for a B757 aircraft, in response to a request from an air cargo tenant. These additions to the project scope had been identified as alternatives during the PDD process, allowing costs for the additions to be quantified in advance. As a result, the estimated cost of the project as designed was similar to the estimates in the PDD. In construction, some creosote piles were found underneath the foundation of the old ARFF station, but otherwise there were no site conditions that were not already documented in the PDD. The apron was completed and opened in March 2010.

**Other Applications**

The Capital Construction Program for Reagan National and Dulles International Airports in Washington, DC, has begun to wind down with the opening of the AeroTrain automated people mover in January 2010, but planning is underway on a new round of projects at both airports, including:

- A hardstand apron with positions for 5 B747 aircraft at Dulles Airport;
- New general aviation and cargo facilities at Dulles Airport;
- Runway Safety Areas for Runways 4-22 and 15-33 at Reagan National Airport;
- A hold apron and taxiway reconfiguration for Runway 1-19 at Reagan National Airport;

- A river rescue facility at Reagan National Airport;
- Relocation of a major electrical and communications vault at Reagan National Airport;
- A perimeter road connecting the North Area of Dulles Airport with the Western Lands and the Airport Support Zone.

Project Definition Documents have been completed or are being prepared for each of these projects, and designs are underway for some. For the hardstand apron at Dulles and the river rescue facility at National, achieving a consensus from all MA-30 departments streamlined the process of getting approval from senior management, and that consensus was fostered by the participation of all stakeholders throughout the PDD preparation process.

**Conclusions**

The collaborative process for developing new projects at the Metropolitan Washington Airports begins with Project Definition Documents. All stakeholders are brought in at the inauguration of project planning, and their consensus, acknowledged by signature in the PDD, assures their satisfaction with the final product. Maintaining a collaborative approach through design and construction is already providing benefits in terms of more predictable costs (as seen in the planning and design of the RON apron at the former ARFF site at Reagan National Airport), more accountability and fewer disagreements about project scopes, schedules and budgets.

# The Federal Aviation Administration's Unique Approach to the O'Hare Modernization Program

Christina Drouet, P.E.; BSCE, MPA; M.ASCE; Manager, Chicago Area Modernization Program Office, Federal Aviation Administration, 2300 E. Devon Ave., Des Plaines, IL, 60018; PH (847) 294-7812; FAX (847) 294-8157; e-mail Christina.Drouet@faa.gov

## ABSTRACT

Shortly after the City of Chicago announced its plans for the O'Hare Modernization Program (OMP) in 2001, the Federal Aviation Administration (FAA) established a project office to oversee the Agency's participation in the program and provide a focal point for the City of Chicago. As the planning and design of the program progress, the FAA's role in the OMP grows and become increasingly complex.

This paper will highlight existing FAA processes that have been adapted to work within the City's aggressive schedule. It will also introduce new positions, functions and processes established in order to meet schedule constraints and ensure compliance with FAA criteria for airport development and operations. The redesign of the Chicago area airspace and its integration with O'Hare airfield changes will be addressed. Lastly, the paper will highlight the complexity of safely operating one of the world's busiest airports while reconstructing that airport at the same time.

✈✈✈✈

## Introduction

Congress has charged the FAA with a myriad of aviation responsibilities for the United States. Chief among these are air commerce and safety. The FAA ensures this by providing an air traffic system for civil and military use. Several key organizations within the FAA contribute to these goals. The Air Traffic Organization (ATO) (approximately 74% of the FAA's workforce) is responsible for air traffic control, and the establishment and operation of air traffic control facilities and supporting infrastructure. The ATO workforce consists primarily of air traffic controllers, technicians and engineers. Aviation Safety comprises the next largest slice of the FAA (15%), and has the mission to ensure safety within the National Airspace System (NAS). The employees in the Aviation Safety organization have professional training in aeronautics, medicine, aviation and aviation safety. The Office of Airports, one of the smallest components of the Agency (<1%), has responsibilities that include airport design standards, airport safety, airport development and grant administration. Most of the remainder of the FAA provides services you would expect to find in any large governmental or corporate structure, such as human resources, legal counsel, and financial services, as primary examples. With respect to airport development projects, the FAA is responsible for establishing

standards for the planning, environmental review, design, construction and safe operation of the NAS and the nation-wide system of public use airports (ACC/FAA, 2008).

**Project and FAA Team Structure**

The Federal Aviation Administration (FAA) established a special project office, the Chicago Area Modernization Program Office (CAMPO) in early 2003 to primarily serve as the focal point for the City of Chicago's O'Hare Modernization Program (OMP). This paper will focus on CAMPO's composition, responsibilities and relationship within the FAA and with the City of Chicago with respect to the OMP. Within the FAA the special project office concept is used on occasion for unique situations such as the establishment of the new Denver Airport that was commissioned in 1995. It was not unusual that the FAA set up CAMPO, but its functioning, breadth of responsibility, and relationships have been, and continue to be, unique.

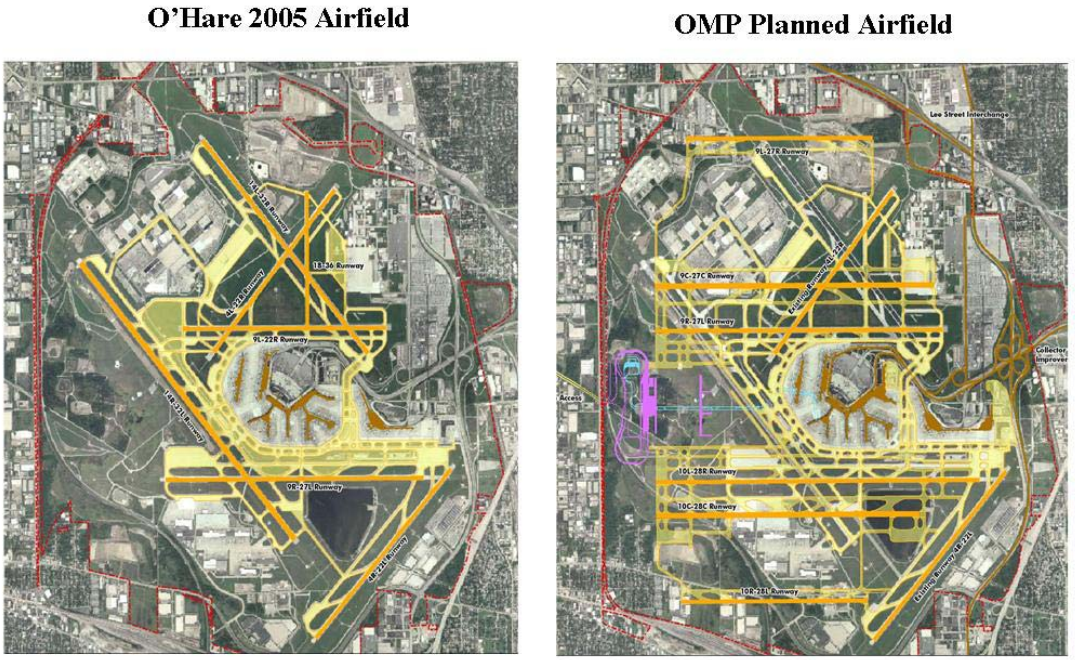



Figure 1. Pre and post O'Hare Modernization Program airfield layouts

In 2001, the City of Chicago proposed the O'Hare Modernization Program. The City is funding the OMP with Airport Improvement Program (AIP) federal grants, Passenger Facility Charges (PFCs) and through an agreement with the air carriers that serve O'Hare. This program includes a complete airfield modernization of O'Hare International Airport. The airfield is currently planned for completion in 2015 at an estimated cost of over $8 billion in infrastructure and airfield improvements. The

Master Plan and approved Airport Layout Plan also include a western terminal complex, still in early planning. The program is not simply lengthening runways or adding runways and taxiways. Rather, the OMP is a total reconfiguration of the airfield from three pairs of parallel runways to six parallel runways in an east-west configuration, and retains the existing northeast-southwest parallel crosswind runways. In effect, the City is rebuilding the airport in its current location while it maintains its full operational capacity. The OMP changes the airfield so much that one central Airport Traffic Control Tower (ATCT) plus two satellite ATCTs will be required for terminal air traffic control when the airfield improvements are completed.

In advising the FAA's management board of the decision to establish the CAMPO in 2003, the FAA Administrator assigned executive responsibility for the OMP to the Great Lakes Regional Administrator. The Regional Administrators do not have direct managerial responsibilities for the majority of FAA employees in their regions. Rather, they have a relatively small and diverse staff that supports Agency-wide functions such as emergency communications and operations, government affairs and aviation outreach. However, Regional Administrators also facilitate intra-agency continuity on complex initiatives, and they act on behalf of the FAA Administrator in this regard. Accordingly, it was then incumbent upon the Regional Administrator to initiate a team and select a manager to lead the team. Once selected, the manager's primary responsibilities were to identify the necessary components of the team and work with the other FAA Lines of Business and Staff Offices (LOBs for simplicity) to provide the needed staffing for the team. The CAMPO manager drafted a team structure and met with the LOBs to discuss the program and anticipated resource needs. The early discussions added to the structure of the team and the process became iterative. As the OMP progresses, team staffing continues to evolve.

Clear messages from the FAA Administrator assisted then, and continue to ensure, LOB support for the program. Initially, team functions and team member interaction were the primary responsibilities of the CAMPO manager. Although not every LOB provided staff for the office as it formed, at the very least they provided a focal point within their organization to work with the team. The most challenging part of organizing the team was establishing the team's role and identity within the well-defined and very structured FAA organization.

Table 1 (page 9) provides the overall FAA organizational structure, by LOB, and Table 2 (page 10) highlights the complexity of the ATO. Organizations with members on the team are highlighted with dashed circles and organizations with occasional or consulting types of relationships to the team are highlighted with dashed squares.

These diagrams are being provided not for comprehension of the entire FAA, but rather to show the complexity of its organizational structure. Added to the challenge of identifying the correct resources needed within the FAA was the additional challenge of a major realignment of the ATO, which began in 2003. Locations,

titles, functionality, responsibilities and reporting structure changed significantly for many of the employees already on the team. During the same timeframe, the Agency embraced the concept of Safety Management Systems (SMS). The SMS provides a common framework to assess safety risks associated with changes to the NAS. It addresses all aspects of Air Traffic Control (ATC) and navigation services, including (but not limited to) airspace changes, air traffic procedures and standards, airport procedures and standards, new and modified equipment (hardware and software), and associated human interactions (FAA, 2008). A change as significant as O'Hare modernization necessitated numerous analyses in accordance with this program. The centralized oversight of the CAMPO ensured resources were assigned to learn the process, risk areas were identified, and the analysis, mitigation and approval processes were completed well in advance of airfield commissioning dates.

The proposed schedule of the OMP gave the FAA an opportunity to implement improvements in the Chicago area airspace. A nationwide airspace analysis had been done in the late 1990's, and components of that study had not yet been implemented in the Chicago area. With the proposed OMP airfield configuration significantly impacting air traffic procedures in the Chicago area, incorporating components of that plan in conjunction with the OMP made optimal sense. The FAA supported the establishment of new air routes and procedures, and the impacts to adjoining air traffic facilities extended from Minnesota to Ohio. Much efficiency was gained, and more will be realized, including additional arrival and departure routes for O'Hare and additional departure routes for Chicago's Midway Airport.

**Promises Made, Promises Kept**

While many of the components of the O'Hare modernization are processes the FAA routinely accomplishes, the magnitude and schedule of the program proposed by Chicago make each task more complex and resource intensive. For comparison, the FAA considers ten years the norm for environmental review and approval, construction and commissioning of a new runway. This program completed the environmental process in three years and commissioned its first runway three years later.

The City and FAA were both keenly aware at the start of the program that the FAA's typical processes would not meet the aggressive timetable the City planned for OMP milestones and completion. Recognizing the importance of the project, and O'Hare's impact on the NAS, the FAA committed itself to the project. The FAA addressed the detailed National Environmental Policy Act (NEPA) requirements and completed a thorough analysis with draft and final Environmental Impact Statements (EISs), which included public outreach. The CAMPO served as the focal point and hub of activity and the Airports and Chief Counsel LOBs provided the technical expertise and staffing to accomplish the work. With the FAA's Record of Decision signed on September 30, 2005, FAA Administrator Marion Blakey, and Secretary of Transportation Norman Mineta, acknowledged the importance of O'Hare

modernization, not just for the Chicago area, but also for the entire U.S. aviation system. (FAA, 2005).

During the early planning and environmental analysis of the project, the City recognized a need to supplement FAA staff in order to meet the aggressive project schedule. They did this through the reimbursable agreement process. A reimbursable agreement is a legally binding contract by which the FAA provides goods and services to a non-FAA party, the associated costs of which are paid by the recipient or project sponsor (FAA Financial Manual, 2007). With this vehicle the FAA was able to hire additional (temporary) employees for the CAMPO, specifically environmental and legal staff. After the environmental process was completed the City continued to recognize the advantage of these FAA positions. The positions currently support expediting the review of the Notice of Actual Construction or Alteration (Form 7460), the design and development of flight procedures, and the processing of Chicago's Engineering Design Review (EDR) packages for the project.

The reimbursable process is also very valuable in assisting Chicago plan and schedule construction. The City is financially responsible to the FAA for many of the proposed facilities, since these FAA facilities are impacted by the airfield modernization. In these situations airport sponsors enter a reimbursable agreement with the FAA to design, construct, install and commission a new Instrument Landing System, as an example. For the OMP, the City decided to design and construct the civil, electrical and mechanical systems, to include the ATCTs, and use the FAA reimbursable agreements to secure FAA support during design, additional design reviews, and FAA construction oversight (not to be confused with Resident Engineering services). The FAA retains responsibility for electronic system design/installation, flight check, and system commissioning, also funded through reimbursable agreements with the City.

By providing the CAMPO as focal point for the City, the FAA has been able to provide a consistent channel for communications. While FAA employees from various LOBs and City staffs meet and exchange information routinely, the CAMPO provides additional value by being the FAA's responsible party to the City. Any question or concern regarding FAA involvement or responsibilities is directed to the CAMPO manager for resolution. Since many issues cross the FAA LOBs, the CAMPO manager brings those parties around the FAA table to discuss and work issues, and then can provide a single FAA position to the City. The involvement varies and has included project schedule changes, design discrepancies, air traffic operations, safety issues and financial procedures. This provides an advantage in that the City does not have to determine whom to speak with in the complex FAA organization. As an example, in early 2008 the City proposed accelerating the commissioning of the Runway 10 extension from November 2008 to September 2008. The City presented the request to the CAMPO manager, who then coordinated with all the FAA LOBs to discuss and analyze the advantages, disadvantages and feasibility of accelerating the schedule. By approaching the request as a team, all of the LOBs were able to hear the pros, cons and impacts to other LOBs, and then could

collectively agree that it was in the FAA's (as well as the City's) best interest to accelerate the schedule. The group discussion also brought to light additional (City and FAA) tasks required to commission early. Taking all input into consideration, the CAMPO manager made the recommendation to the Regional Administrator to accelerate the commissioning. The Runway 10 extension was successfully commissioned on September 25, 2008.

The CAMPO provides advantages within the FAA. With proper coordination through the CAMPO office, all LOBs are aware of work being done in one LOB that might impact another. The office provides a consistent source of information when any date or activity is unclear to team members. Through weekly meetings, established channels of communications and shared office space, internal communications are optimized. LOBs that might only occasionally check in with each other on another project face each other weekly, if not more frequently, to coordinate details from lease and environmental responsibilities to equipment deliveries and design reviews. One of the phrases the FAA uses to describe this concept is horizontal integration, a valuable tool for any project that crosses multiple responsibilities and professions, and especially needed in the tightly structured FAA.

The CAMPO also serves as a centralized office to handle incoming correspondence, Freedom of Information Act (FOIA) requests, calls and emails regarding the FAA's role and involvement in the OMP. Early in the environmental process the FAA established a website for the O'Hare modernization. Environmental data, modeling results, public information and documents released under the Freedom of Information Act (FOIA) were posted. During this time period, the office also received voluminous FOIA requests. CAMPO assisted all involved LOBs with processing the responses, which include over 15,000 documents (7.5 million pages or 120G of data.) Since the opening of the first OMP runway (Runway 9L/27R), local community interest has been addressed through an expanded website that includes frequently asked questions regarding the operation and use of the new runway, and highlights areas of the approved environmental documents that pertain to airfield operations and air traffic configurations. Many times the public poses similar questions and these letters and emails may be addressed to a variety of elected officials and FAA executives. In routing all incoming inquiries through one central office, it provides the advantage of repeatable, consistent responses with input from all pertinent LOBs.

**FAA's Runway Template Action Plan**

With air carrier operations increasing at a steady rate in the 1990's, commercial airport operators and the FAA recognized the need to add capacity. Increased capacity for the NAS is one of FAA's four business plan components (safety, international outreach and organizational excellence being the others). The FAA wanted to ensure it provided timely support to airport operators making the financial investment in new runways and runway extensions. In order to ensure this, the Runway Template Action Plan (RTAP) process was developed. Briefly, it is a Microsoft Project based application that identifies over 200 items that typically

comprise establishment of a new precision approach runway. It includes not only FAA activities, but also airport owner/operator activities. Each Regional Administrator is assigned executive responsibility for the appropriate utilization of the tool for the runway projects within his/her region. The Regional Administrators then determine who to involve and how to accomplish the milestones within the RTAP. For each project, the region assembles the appropriate team with various resources and roles. The most important function of the process is to keep FAA's executive management board informed of the status of each of these projects on a quarterly basis. From 2001 through early 2010, this tool was successfully used to manage the on-time accomplishment of FAA work needed to support commissioning 18 runways or runway extensions at the busiest airports in the country.

The RTAP is a good tool to define accountability. It clearly designates airport owner/operator activities and FAA LOB responsibilities. It provides the framework for the FAA and outside stakeholders closely involved in a runway project to work together in partnership, all focused on a common goal. The RTAP was established at a time when commercial aviation demand overwhelmed existing capacity. The RTAP, along with the horizontal integration team concept, remain useful tools for intra-agency and inter-agency accountability. It is the tool used for each of the OMP runway projects.

**Conclusion**

Since the FAA's involvement in the OMP began in 2002, the team has successfully completed the Environmental Impact Statement and Record of Decision (September 2005); commissioned numerous airspace changes (2007 and 2008); commissioned a 3000' extension to Runway 10/28 and upgraded the Runway 10 Instrument Landing System (ILS) to Category II/III (September 2008); commissioned a new satellite ATCT and new 7500' long Runway 9L/27R with Category I capabilities (November 2008) and Category II/III ILS capabilities by May 2009. The most significant accomplishment to date, however, has been the continual safe operation of O'Hare International Airport during the construction of the modernization. The required vigilance to achieve this significant milestone cannot be understated. The FAA Air Traffic Organization has relied on the Safety Management System tools mentioned earlier for each change to air traffic operations or airfield configuration. The collaborative working relationships with the Chicago Department of Aviation, including O'Hare operations and the OMP organization, and inclusion of FAA in discussions, analyses and design reviews have all contributed to continued safe operations.

The crowning jewel of the FAA's involvement in OMP has been the individual commitment of the employees dedicated to the project. Clearly stating a goal and providing a framework from the top of the FAA was, and continues to be, instrumental to this program's success. Thereafter, putting the right people on the team and providing an environment that allows the team to establish its own culture has led to FAA's successful contribution to the OMP. A shared and cohesive culture,

rather than a clear, well-defined team structure is the force that gives the team its drive (Bolman and Deal, 2008). The different LOBs within the FAA each have their own culture and approach towards their goals. The great diversity that results from including staff from ten different LOBs has created a unique culture within the FAA and contributes to the success of the FAA's OMP team.

The FAA approached the O'Hare modernization as a unique event by establishing a special projects office and allowing it to create its own culture. The FAA continues to demonstrate its abilities to be nimble, responsive in addressing improvements to the NAS and effective at establishing and maintaining collaborative relationships with airport operators. Given the complexity of the program, important lessons learned within the FAA include having dedicated and repeatable communications channels (internal and external), responsive public information, and clear accountability and responsibility. It is also apparent that despite the inevitable personnel and organizational changes, the team assembled to execute the OMP remains focused, and is supported from the executive level.

## References

ACC/FAA. *Improving the Quality of Airport Projects: ACC/FAA Best Practices*. 2008
FAA, Air Traffic Organization, *Safety Management System Manual*, April 2008
FocusFAA, September 30, 2005
FAA, *FAA Financial Manual*, 2007
Bolman, L.G. and Deal, T.E. *Reframing Organizations – Artistry, Choice and Leadership.* San Francisco: Wiley, 2008
Kraus, T.L. *The Federal Aviation Administration: A Historical Perspective, 1903-2008*, Washington D.C., U.S. Department of Transportation, 2008
FAA O'Hare Modernization Program website
http://www.faa.gov/airports/airport_development/omp/

## Acknowledgements

I would have nothing to write about if it were not for the extraordinary team of FAA employees that work diligently every day on the O'Hare modernization while also maintaining the safety of operations at O'Hare. Thank you for your dedication to public service.
My appreciation goes to Barry Cooper, Vincent Siciliano, Mel Banks, Patrick Wells, Charles Prock, Ellen Kijowski, Amy Hanson and Richard Kula of the FAA for their input, assistance and support in writing this paper.
Thanks to Rosemarie Andolino, Commissioner of Aviation and Michael Boland, First Deputy Commissioner of Aviation, City of Chicago, for their exceptional leadership on this very challenging and successful program.
And many thanks to James Chilton of AECOM (Chicago's OMP Program Manager) for his professionalism and focus.

Tables

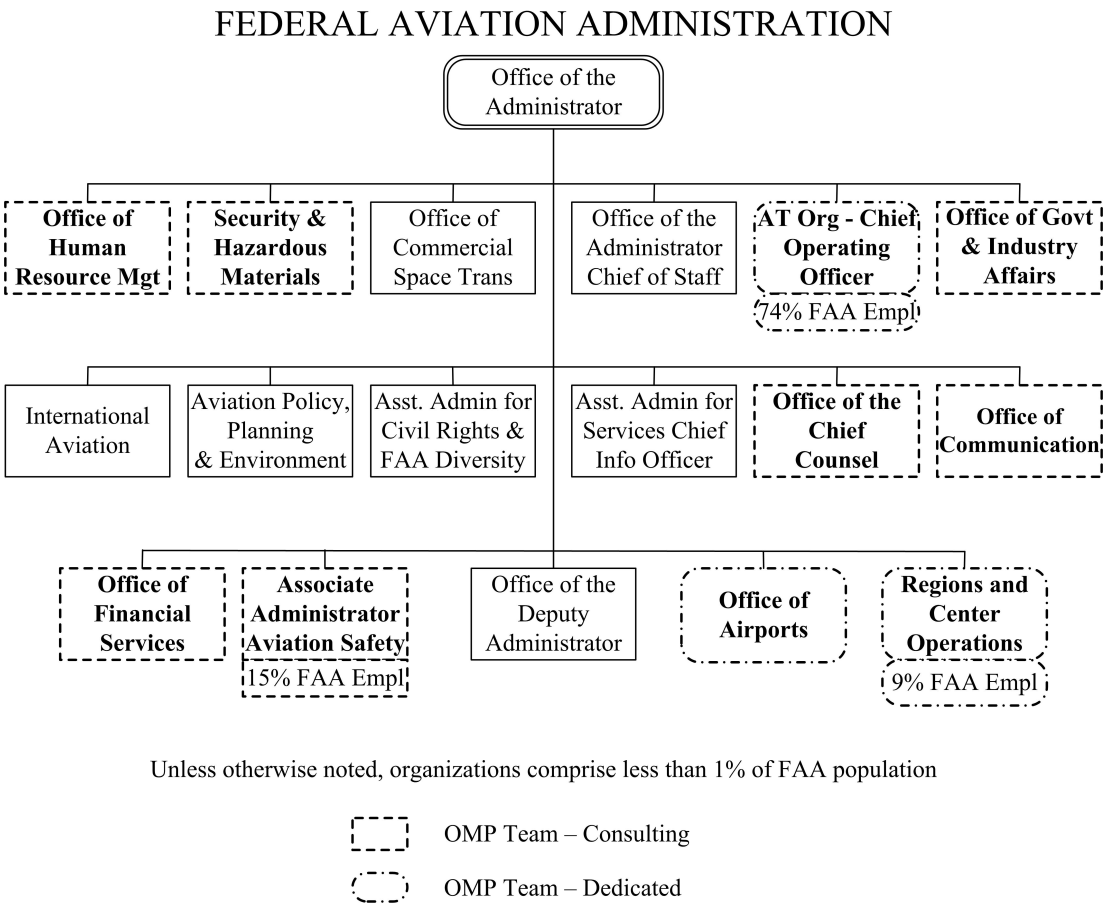



Table 1. Federal Aviation Administration Organizational Structure

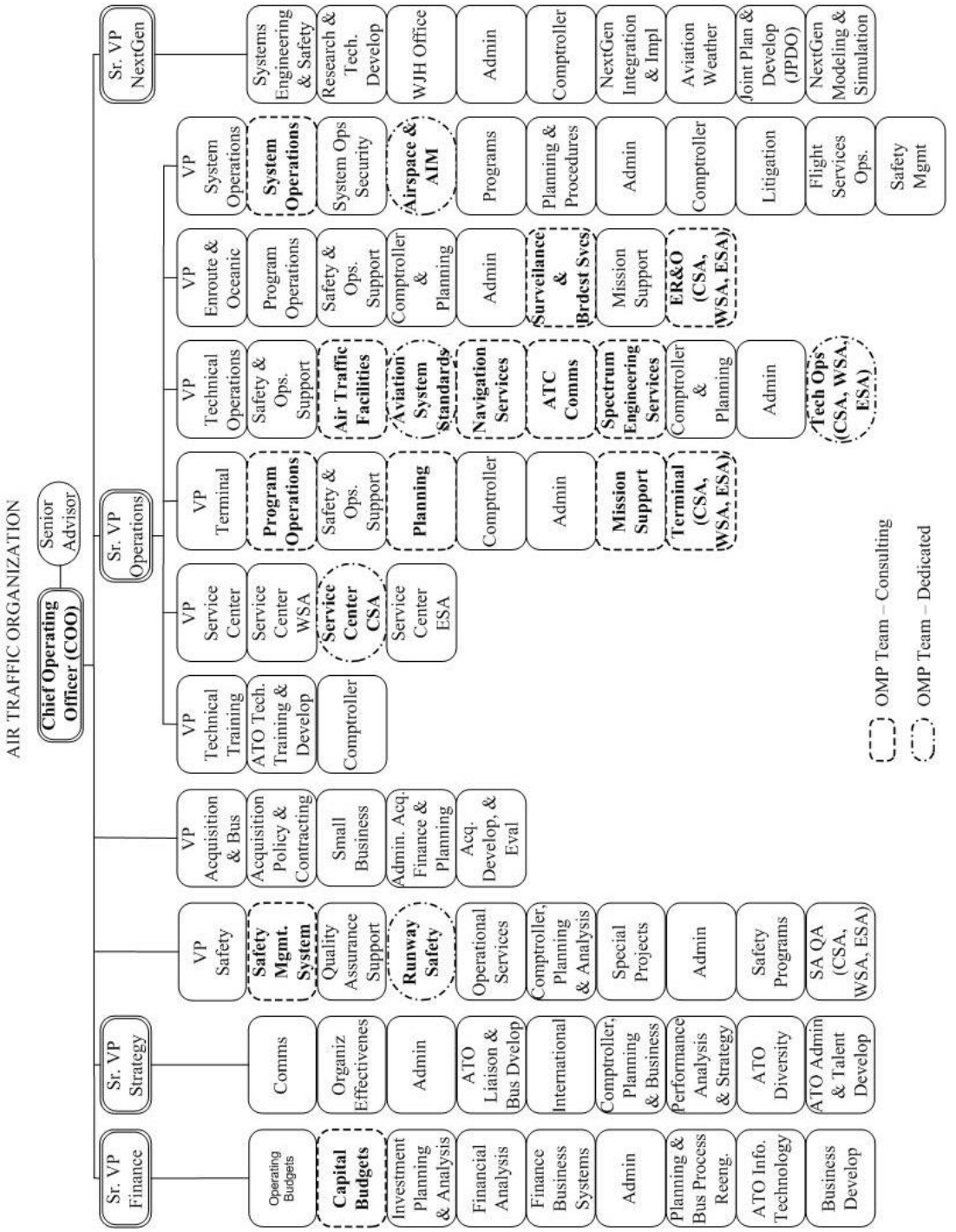



Table 2. Federal Aviation Administration, Air Traffic Organization

Case Study for the Planning, Design and Construction Phase of Runway 13R-31L at John F. Kennedy International Airport.

Guy Zummo, P.E., M.ASCE[1] and Scott D. Murrell, P.E., M.ASCE[2]

## ABSTRACT

At 14,511-feet long, Runway 13R-31L is the longest runway at John F. Kennedy International Airport (JFK), and the second-longest commercial runway in North America.

The runway's last major rehabilitation was in 1993. In anticipation of the next major repaving contract, and the need for JFK to maximize the efficiency of its runway and taxiway system, a multi-year program was initiated. The first step of this program included the development of a conceptual plan for the runway and associated taxiways to minimize delays and have the runway system conform to FAA Design Group VI standards.

Using the conceptual plan a Constructability Study was performed that investigated traditional hot-mix asphalt (HMA) rehabilitation as well the placement of a Portland cement concrete (PCC) overlay on top of the existing HMA surface, and included life-cycle costs, staging options and construction durations. A series of briefings were held with the airlines and the FAA to present the study's findings and to get their buy-in to the final plan.

Development of the contract documents to widen the runway to 200 feet using (PCC) and to construct the new taxiways followed the completion of the Constructability Study. A series of peer review meetings as well as a contractor's workshop were held during the development of the contract documents to ensure that all elements of the construction were addressed.

In June of 2009, this contract was awarded. Construction of a test section was completed in November 2009, and a series of lessons learned meetings followed. On March 1, 2010, the runway was closed to operations for reconstruction. Construction is scheduled to be completed in November 2010, with a significant portion of the runway returned to service during the summer of 2010. This paper presents a case study covering the planning, design and construction of this significant project.

## INTRODUCTION

Runway 13R-31L at John F. Kennedy International Airport (JFK) was originally constructed during the 1940s and lengthened its present length of 14,501 feet under two subsequent contracts (Figure 1). Presently the runway is 150 feet wide. The

[1] Senior Consulting Engineer, Port Authority of NY & NJ, Two Gateway Center – 16th Floor, Newark, NJ 07102 Tel. 973-792-4388 Fax: 973-792-4303 Email: GZummo@panynj.gov
[2] Chief Civil Engineer, Port Authority of NY & NJ, Two Gateway Center – 16th Floor, Newark, NJ 07102 Tel. 973-792-4327 Fax: 973-792-4303 Email: SMurrell@panynj.gov

original pavement section was 12" Portland cement concrete (PCC) on 6" crushed stone screenings. During the 1970s, the runway was overlaid with hot-mix asphalt (HMA). Over the years, the runway has been overlaid number of times and as a result, there are 16-inches of HMA on top of the original PCC surface. The last overlay was placed during 1993. By 2003 the airport's pavement management system projected that Runway 13R-32L would require rehabilitation during 2005. Conceptual planning began for the rehabilitation of Runway 13R-31L.

## CONCEPTUAL PLANNING

Since the 1970s, rehabilitation of Runway 13R-31L has been performed by adjusting or replacing centerline lights and overlaying with HMA. Preliminary design began assuming that a similar approach would be used. However, as the New Large Aircraft of the future became the A-380, it became apparent that Runway 13R-31L needed more than rehabilitation.

JFK is designated a Group V airport under FAA Classifications. The Port Authority of NY & NJ (PANYNJ), which operates JFK, performed multiple studies over the years to determine the airfield modifications necessary to accommodate FAA Group VI designated aircraft in general and the A-380 specifically. Based on these studies and discussions with the FAA a number of required airfield modifications were identified including widening Runway 13R-31L from 150 feet to 200 feet. Larger fillets, widened shoulder and blast pavements were incorporated into the planning. Runway strengthening was not required. The scope of the project changed to rehabilitation and widening.

Another development that would impact the project scope was the significant growth in air traffic operations at JFK starting in 2005. Between 2004 and 2007, plane movements increased from 320,000 to 440,000. This growth lead to additional regional airport delays. In response, the JFK Delay Reduction Program was developed. The program includes eleven potential ground improvement projects for moving aircraft to and from the runways more efficiently. Runway 13R-31L taxiway entrance and exit modifications and relocated runway thresholds were included in this program. The scope of the 13R-31L Rehabilitation and widening project changed again to include delay reduction program components.

During the summer of 2007, a Constructability Study was performed in conjunction with a life cycle cost analysis. Two alternatives were fully developed.

- Rehabilitation using HMA – including a 9" thick overlay
- Reconstruction with PCC – including 18" PCC

The scope of both alternatives included widening the runway to 200 feet; new lighting systems, new drainage systems, re-grading safety areas, new taxiways and alignment changes to existing taxiways.

Two staging options were considered:

- Nightly closures – 10:00 pm to 6:00 am the following morning. The runway reopens to operations each day.
- Staged full closures – Three stages 12,000 feet, the intersection with Runway 4L-22R and the remaining 2000 feet.

The construction duration for each alternative and each staging approach was estimated see Table 1. Airfield lightning and the associated electrical work were on the schedule's critical path.

The life cycle cost analysis was performed using a 3.5% discount rate. The initial cost for the HMA rehabilitation was 3% cheaper than the PCC reconstruction. The life cycle cost for the PCC construction was 35% cheaper than the HMA rehabilitation.

Following a series of meetings with the FAA and the airlines operating at JFK the PCC alternative with staged full closures was selected. Final Design began January 1, 2008.

## DESIGN

In many ways, the design of Runway 13R-31L is typical for a US commercial service runway. FAA Advisory Circular requirements for Airport Design, Airfield Lighting and Pavement Design were followed. However, some significant exceptions to normal practice were incorporated.

### Concrete Mix Design

The Port Authority's Materials Engineering Group developed and tested mix designs to establish specifications, which would result in a durable Portland cement concrete pavement that meets strength requirements. The following adjustments to the specifications were developed:

- Graded Aggregate – See Table 2
- Total Cementitous content not exceeding 550 pound per cubic yard, with a minimum of 40% slag cement
- Permeability ≤1500 coulombs when tested in accordance with AASHTO T277.
- 28 days drying shrinkage less than 0.40% when tested in accordance with ASTM C157.

These additional mix design requirements were approved by the FAA and included in the project specifications.

### Pavement Design

The FAA's FAARFIELD pavement design program does not provide for the design of a concrete overlay of an HMA surface on an earlier concrete pavement. To

overcome this, a sensitivity analysis was performed. The existing pavement was considered an undefined base and a variety of elastic modulus values were used in the analysis. PCC thickness results ranged from 15-20 inches. The 18-inch overlay selected considered the existing asphalt and existing PCC as all asphalt with an elastic modules of 200,000 psi. Eighty-six aircraft types operate at JFK, which exceeds FAARFIELD's input capabilities, so similar aircraft were grouped and a representative aircraft used in the design. Twelve representative aircraft were used for design, including the A-380 (Table 4).

To optimize the surface grades the design called for milling the existing HMA runway surface to 6 inches below existing grade. Since PANYNJ already had good experience placing PCC on asphalt-stabilized base, it was decided that the PCC would be placed directly on the milled surface. The concrete mix requirements limiting drying shrinkage also influenced the decision not to place an asphalt bond breaker on the milled surface. The specifications called for the milled surface to meet a ¼-inch grade tolerance and ¼" in 10-footsmoothing tolerance. Whitewashing the milled surface with a liquid membrane-curing compound prior to placing the PCC was specified to eliminate bonding of the PCC pavement to the milled surface and to lower the surface temperature of the milled surface.

Reuse of Removed Pavement

In addition to milling 50,000-cubic yards of asphalt from the surface of the runway, removal of PCC and HMA surfaced taxiway pavements were also included as a part of this project. Instead of trucking all this material off site for recycling or disposal, the design allowed for reuse of the removed pavement as sub-base materials. This on-site reuse is the most sustainable use possible of removed pavement.

Runway Intersection Grading

Since the design of Runway 13R-31L specifies for the finished surface to be one foot above the existing surface, the grading of the intersection with Runway 4L-22R required special consideration. To ensure that the design did not result in a rough riding pavement the PANYNJ contracted with APR Consultants to verify the design. APR performed a simulation of a variety of aircraft operating across the intersection on Runway 4L-22R, checking for unwanted aircraft response to the proposed profile. The verified design was incorporated into the runway grading plans.

Other Design Phase Activities

During the design phase a peer review of the project was performed. The design, staging approach and time of completion were all scrutinized. Another design phase activity was the prequalification of bidders. In addition to demonstrating the financial capacity to deliver the project, prospective bidders were required to show that they could meet certain production rates for PCC and HMA paving and airfield lightning installation.

A Contractors Forum was then held with the prequalified bidders, The "Forum" included a project briefing by design and construction management staff, a site tour and a question and answer session. The 50-percent complete design drawings were given to the prequalified bidders at the Contractors Forum.

Special Contract provisions were also developed for this project. Damages for delay of up to $6,000 per minute and $300,000 per day were included in the contract. These penalties were significantly higher than in any other PANYNJ runway paving contracts. Additional compensation for early completion was also included in the contract. Up to $10 million can be earned by the Contractor for completing certain stages and the overall contract early. The Contract was awarded June 25, 2009, and the project moved into the construction phase.

## CONSTRUCTION

The construction contract includes:

- 200,000cubic yards of concrete
- 240,000 dowels
- 55 miles of joint seals
- 325,000 tons of HMA
- 12 miles of drainage pipe
- 1500 in-pavement lights

The first major component of the construction undertaken was the construction of the Taxiway "KC" test section. The 800-foot by 100-foot test section was included in the contract so that the means and methods of construction could be evaluated and refined prior to closing Runway 13R-31L for construction. An ERIE Strayer model MG-12CP2 concrete batch plant was erected on-site, and a GOMACO GP-4000 slipform paver was used for "straight line" paving. The curved fillets at taxiway intersections were paved using a Bidwell 4800 paver. Requirements were included to simulate the milled surface of the runway. The test section PCC pavement was grooved, so that all construction operations were proven prior too moving on to the runway. The Taxiway "KC" test section was completed in November 2009. Following the completion of the test section, several lessons learned meetings were held. These sessions identified any concerns and developed corrective actions.

On March 1, 2010, the runway was closed to operations thus beginning the 120-day closure for the construction of Stage I. The work within the Stage I area included approximately 12,000 feet of the runway, taxiways, airfield lightning and signage, and drainage. Construction began with installation of a barrier/security fence enclosing the work area and a roadway leading off the airfield. The barrier/security fence effectively removed the Stage I area from the airfield and minimized the need for airfield operations staff to escort contractor staff and equipment. Even before the site was enclosed, milling and removal operations began. The Contractor quickly

realized that a two-pass milling process provided a surface meeting the specification (Figure 2), and met production requirements. Installation of airfield lightning systems also began on day 1.

Whitewash was applied to the milled surface at a rate of one gallon per 200-square feet. Dowels on chairs were installed at the transverse joints. In pavement lights were installed within cages (Figure 3) as recommended by an Innovative Pavement Research Foundation Study.[3] The same batch plants and pavers used for the test section were used to pave the runway and a GOMACO GHP-2800 slipform paver was used to pave in-fill lanes. A burlap drag finish was applied and liquid membrane compound used. An early entry saw was used to cut the transverse joints. A performance grade (PG) 76-22 binder was used in the HMA placed on taxiways. HMA with PG 64-22 binder was used for the shoulder and blast pavements. The HMA mix design requirements are shown in Table 3. During Stage I paving production rates peaked at 4750-cubic yards per day for PCC and 5000 tons per day for HMA.

All PCC and HMA acceptance testing was performed by PANYNJ Materials Engineering staff. The average 28-day flexural strength of the PCC was 1,105 psi, with a standard deviation of 92 psi. The average permeability was 667 coulombs with a standard deviation of 92. Port Authority surveyors checked final surface grades. Less than 0.4% exceeded the specified grade tolerance of .04 foot prior to grinding Smoothness acceptance was based on profile measurement using a lightweight profilomenter, which was then evaluated using a California Profilograph simulation. The final surface, after limited grinding, had an average profile index of 6.7 inches per mile, using 0.2-inch blanking band. The runway was reopened to air traffic at 11:40 am June 28, 2010, one day ahead of schedule (Figure 4).

The remaining portions of the project are broken into two main stages, the intersection with Runway 4L-22R, and the area east of the intersection. Following the completion of Stage I, construction operations moved to the area east of the intersection. Between September 16 and 29, 2010, Runway 4L-22R was closed for paving of the intersection. Over 1,500 feet of Runway 4L-22R was resurfaced with HMA to meet the PCC on Runway 13R-31L. When the runway paving was completed the true profile was measured using an automated rod and level and an aircraft simulation was performed to verify the smoothness of Runway 4L-22R. Aircraft response was below 0.4 g.

One apparent benefit of the early completion incentives was the unprecedented level of detailed scheduling and daily monitoring that was implemented during all construction stages. The prime contractor and his subcontractors reported their progress, planned construction, issues, and recovery plans at daily meetings intended to maintain the project's schedule.

[3] Sonsteby, O. (2008). "Constructing In-Pavement Lighting, Portland Cement Concrete Pavement". IPRF Report 01-G-002-03-1

Conclusions

With any successful project, careful planning and a quality design result in a successful construction phase. The planning, design and construction of JFK Runway 13R-31L can serve as case study on how to deliver a project of this type successfully. Additional steps taken during the planning and design phase, such as prequalification of contractors, constructability analysis, peer review, optimization of the concrete mix, and getting approvals for completion incentives, proved beneficial to construction. The requirement of a large test section, post test section lessons learned sessions, coupled with superior quality assurance inspection/testing resulted in a quality runway.

**Table 1 – Construction Duration of Staging Alternatives**

| | | **HMA** | **PCC** |
|---|---|---|---|
| **Nightly Closures** | Construction Duration*<br>RW 4L-22R Closure<br>RW 13R-31L Shortened **<br>RW 13R-31L Closure | 24-30 Months<br>10 Days<br>45-90 Days<br>275 Nights | N/A |
| **Staged Full Closures** | Construction Duration*<br>RW 4L-22R Closure<br>RW 13R-31L Shortened **<br>RW 13R-31L Closure | 18-23 Months<br>10 Days<br>70 Days<br>120 Days | 18-23 Months<br>14 Days<br>75 Days<br>120 Days |

**Table 2 – PCC Aggregate Gradation**

| **Sieve Size** | **Percentage by Weight Passing** |
|---|---|
| 2 ½ | 100 |
| 2” | 90-98 |
| 1 ½” | 76-88 |
| 1” | 67-79 |
| ¾” | 65-77 |
| 3/8” | 48-60 |
| No. 4 | 30-42 |
| No. 8 | 27-37 |
| No. 16 | 20-30 |
| No. 30 | 16-22 |
| No. 50 | 4-10 |
| No. 100 | 0-4 |

**Table 3 – HMA Aggregate Gradation**

| Sieve Size | Percentage by Weight Passing Sieve |
|---|---|
| ¾” | 100 |
| ½” | 72-98 |
| 3/8” | 60-82 |
| No. 4 | 40-56 |
| No. 8 | 28-39 |
| No. 16 | 19-24 |
| No. 30 | 13-19 |
| No. 50 | 8-16 |
| No. 100 | 5-10 |
| No. 200 | 3-6 |

Asphalt, Percent Weight of Total Mixture 5.2 - 6.

**Table 4 – Design Aircraft**

| |
|---|
| A320-200 Twin Std |
| B737-800 |
| B737-900 ER |
| B767-300 ER |
| B787-9 |
| A300-600 Std |
| A330-200 Std |
| B777-300 ER |
| A340-300 Std |
| A340-300 Std Belly |
| A340-500 opt |
| A340-500 Belly |
| B747-400 |
| A380-800 |


**Figure 1 – Runway 13R-31L at JFK**


**Figure 2- Whitewashed Milled Surface**


**Figure 3- Reinforcing Cage around Light Can**

**Figure 4- Stage I Completed and Runway Reopened**

# Preserving Our Airfield Pavements

**David K. Hein, P.Eng. and Brian Aho**

Applied Research Associates Inc.
5401 Eglinton Avenue West, Suite 105
Toronto, Ontario, Canada, M9C 5K6,
tel: 416-621-9555 fax: 416-621-4719,
dhein@ara.com
baho@ara.com

## ABSTRACT

The need to preserve our airfield pavement infrastructure is paramount to insuring the viability of transportation of people and goods. Preventive maintenance plays an important role in the preservation of airfield pavement infrastructure. A successful preventive maintenance program cannot function without the support of many features associated with pavement management systems (e.g., pavement inventory, condition assessment, and the framework for the identification and prioritization of pavement preservation treatments).

The purpose of a preventive maintenance treatment is to prevent premature deterioration of the pavement, retard the progression of pavement defects, and cost-effectively extend the life of the pavement. The objective is to slow down the rate of pavement deterioration and effectively increase the useful life of the pavement. A preventive maintenance treatment is not determined by the type of treatment, but by the reason why the treatment is performed.

For cost-effective preventive maintenance it is necessary to apply the right treatment to the right pavement at the right time. The objective is to identify the sections that would benefit most from preventive maintenance (the right pavement), do the identification and apply the treatment in a timely manner (the right time) and to select the most beneficial treatment (the right treatment). The effectiveness of preventive maintenance is largely dependent on the timing of maintenance activities. To ensure that funding for preventive maintenance is available when required, many practitioners advocate the establishment of adequate dedicated funds for preventive maintenance.

The development and implementation of a preventive maintenance program should be done in a collaborative manner, and should be supported by training and educational activities. To succeed, a preventive maintenance program requires a long-term commitment, ongoing improvements, and the documentation and reporting of program benefits.

This paper outlines 7 basic steps involved in developing and implementing a pavement preservation program for airport pavements and provides examples of decision matrices and life-cycle cost analysis procedures to apply and evaluate the success of pavement preservation treatments.

## INTRODUCTION

Every airport operator is faced with the need to maintain airside pavements in good order for safe and efficient aircraft operation using the available budget. This paper describes a rational approach to the development of pavement preservation budgets based on a systematic and

objective documentation of pavement preservation needs. It also describes procedures for prioritization of maintenance and rehabilitation projects.

Decision making for maintenance and rehabilitation of airport pavements consists of two sets of sequential decisions. The first stage involves identifying and prioritizing future pavement preservation needs, treatments, and projects considering the needs and priorities of all airport pavements together. The objective of this stage is to decide at a given time which pavement sections should receive maintenance or rehabilitation (M&R) treatments. The second stage consists of determining, using site-specific engineering considerations, what type of M&R treatment should be carried out on the previously selected sections. There are several publications that provide useful information on pavement management procedures, including pavement condition evaluation, selection of maintenance and rehabilitation treatments, priority analysis and other pavement management topics (ARA 2010, FAA 2006, FAA 2007, Haas at al. 1994 and Shahin 1994). The main components of an Airport Pavement Management System (APMS), grouped into seven main activities, are shown in Figure 1.

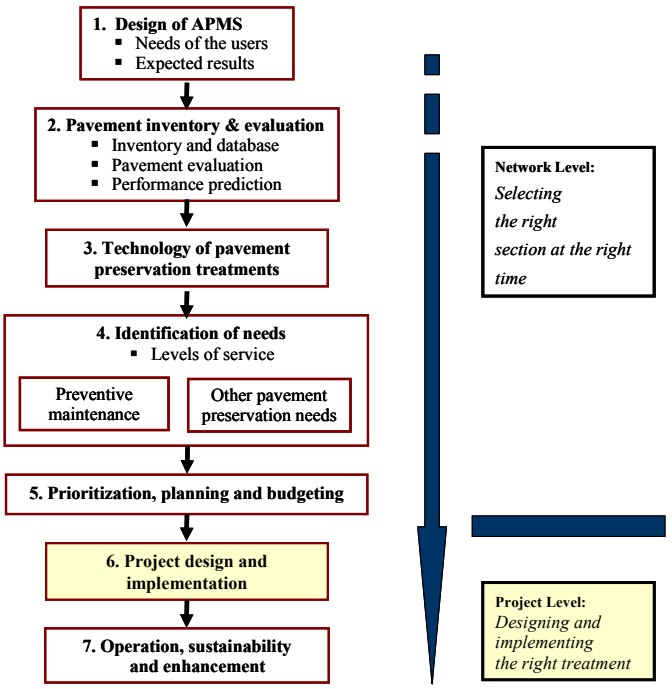



**Figure 1. Primary components of an airport pavement management system.**

## DESIGN OF AIRPORT PAVEMENT MANAGEMENT SYSTEMS

The role of an APMS is to support technical, engineering, and management activities of airport personnel responsible for providing pavement infrastructure for safe and efficient operation of aircraft. The pavement management process provides systematic and objective procedures for maintaining the inventory of pavement infrastructure, monitoring pavement performance, planning and budgeting of pavement preservation activities, and evaluating the cost-effectiveness of past pavement preservation actions. The APMS includes all activities connected with pavement infrastructure, including the initial pavement design and construction, and the subsequent pavement maintenance and rehabilitation activities.

The initial design of APMS is important for ensuring the future use and sustainability of the system. A comprehensive summary of design activities required for successful implementation and operation of APMS is provided in the Transportation Research Circular on Implementation of an Airport Pavement Management System (FAA 2008). Briefly, the design and implementation of an APMS includes the following activities:

- Obtain commitment to establish and operate an APMS, and appropriate funding to do so, from the airport management;
- Identify potential users of the system and determine their needs;
- Decide who will develop and operate the system (internal staff, consultant or a combination of the two);
- Select APMS software;
- Develop database including sectioning of the network and initial pavement condition evaluation;
- Customize software to reflect local input values such as pavement deterioration rates, M&R policies, typical unit costs of M&R treatments, and agency-specific preferences and priorities concerning the selection of M&R treatments;
- Customize software to incorporate agency preferences, such as network condition analysis and the incorporation of Geographic Information Systems (GIS);
- Initial and ongoing staff training; and
- Follow-up plan to ensure that data are updated.

## PAVEMENT INVENTORY AND EVALUATION

An accurate pavement inventory is a prerequisite for a systematic pavement condition evaluation and the selection of M&R treatments. The inventory must include the size and characteristics of pavement assets and their condition. For the purposes of pavement inventory, it is necessary to divide the airport pavement network into homogeneous pavement sections (Shahin 1994). A pavement section should have the same pavement structure and a similar condition throughout. A pavement section is the basic building block for pavement inventory. It is also a basic unit for pavement preservation decision making. If necessary, a maintenance and rehabilitation project can be carried out on a single pavement section.

Pavement condition surveys that evaluate the type, severity and extent of pavement surface distresses are also required for preventive maintenance programs. However, for the selection and application of preventive maintenance, it is also necessary to identify specific pavement conditions and early indicators that trigger the need for preventive maintenance treatments.

Preventive maintenance treatments need to be applied when they are most cost-effective, typically before distresses progress and more expensive corrective treatments are required. For example, routing and sealing of cracks in asphalt concrete pavements should be carried out when the cracks are already well-formed, but before cracks become raveled, have developed into multiple cracks, or before the crack width exceeds about 3/8 of an inch. Consequently, condition surveys of pavement surface distresses for the selection and timing of preventive maintenance treatments should be preferably carried out on candidate pavement sections annually. The first pavement preservation treatments are typically carried out when the pavement surface layer is between 3 and 5 years old. For planning purposes, it is

necessary to estimate future pavement preservation needs. A typical planning period is 5 years, but some large airports may prepare pavement preservation plans for major runways for up to 10 or 15 years. The identification of future pavement preservation needs requires the prediction of pavement performance and storing the results in the APMS database.

Pavement performance depends on the pavement structure, as well as many such as the type and frequency of traffic loads, environment, and subgrade characteristics including drainage. Consequently, pavement performance models are not easily transferable from airport to airport. The selection of performance models depends on available data, agency requirements for estimating future preservation needs, and on the APMS software used.

Typical methods used for pavement performance modeling include:

- Expert modeling for which a trained pavement expert estimates the future condition of the pavement when historical pavement performance data are not available;
- Families of performance curves which are based on the expectation that similar airport pavements exposed to similar traffic will perform in a similar way;
- Extrapolation of existing trends, which is a variation on family modeling and is used when the condition of the pavement was evaluated on only one previous occasion, and the family pattern is extrapolated taking into account the past condition;
- Markov probability models have been used for pavement performance prediction of highway pavements but are not used extensively for airfield pavements; and
- Artificial neural networks (ANN), or neural networks, that can link a large set of data (e.g., a data set describing a pavement and its exposure to the traffic and environment) to a required outcome (e.g., expected life-span of the pavement) without using traditional statistical analysis.

## TECHNOLOGY OF PAVEMENT PRESERVATION TREATMENTS

There is a large amount of information available concerning the technology of pavement preservation treatments. For example there are literally dozens, or in some cases even hundreds, of reports on each of the 24 M&R treatments listed in Table 1 (FPP 2001, FHWA 2006a, FHWA 2006b). Some of the treatments listed can be applied, with only very small modifications, to both AC and PCC pavements; for example, micro-surfacing, and controlled shot blasting. Other treatments are just a variation of the same treatment using different materials. An example for AC pavements would be the crack sealing of AC pavements with hot-poured sealant versus cold-applied sealant. An example for PCC pavements would be the shallow patch repair using PCC material, versus using AC material, or proprietary material.

## IDENTIFICATION OF NEEDS

Identification of pavement preservation needs is based on the results of pavement condition surveys, the prediction of pavement deterioration, and the desirable level of service for airfield pavements. The concept is simple: pavement preservation needs arise when the predicted pavement condition is lower than the recommended or mandated level-of-service. The key for the successful operation of this model is the objective assessment of pavement condition and the establishment of the levels of service that are accepted by decision makers.

**Table 1. Typical Airport Pavement Preservation Treatments.**

| Both Pavement Types | Asphalt Concrete | Portland Cement Concrete |
|---|---|---|
| • Texturization using controlled shot blasting<br>• Diamond grinding<br>• Micro-surfacing | • Sealing and filling of cracks (with hot or cold applied sealants)<br>• Small area patching (using hot mix, cold mix, or proprietary material)<br>• Spray patching (manual chip seal or mechanized spray patching)<br>• Machine patching with AC material<br>• Rejuvenators and seals<br>• Texturization using fine milling<br>• Surface treatment (chip seal, chip seal coat)<br>• Slurry seal<br>• Hot Mix overlay (includes milling of AC pavements)<br>• Hot in-place recycling<br>• Cold in-place recycling<br>• Ultra-thin whitetopping | • Joint and crack sealing (with bituminous, silicone, or compression sealants)<br>• Partial depth repairs (using AC, PCC, or proprietary materials<br>• Full depth repairs (using AC, PCC, or proprietary materials)<br>• Machine patching using hot mix<br>• Slab stabilization and slab-jacking<br>• Load transfer<br>• Crack and joint stitching<br>• Hot mix overlays<br>• Bonded PCC overlay |
| **3 treatments** | **12 treatments** | **9 treatments** |

Pavement preservation needs depend on the level of service the airport pavements are expected to provide. A higher level of service, for the same pavement structure, results in higher M&R needs and thus in higher agency costs. Levels of service for airport pavements are typically expressed in terms of pavement condition index (PCI) values (Tighe 2004).

The target level of service should be specified for different facility types because all facilities do not require the same level. If the target level of service is approved and mandated by the airport management, it can be used to determine the pavement preservation needs required to provide the mandated level of service. In other words, the pavement preservation needs become justified and mandated on the basis of approved criteria. It should be noted that levels of service in terms of PCI depend on several factors:

- Airport type and size – General aviation airports may allow a lower level of service than carrier airports, particularly large carrier airports;
- Facility type – Higher PCI values are typically required for runway pavements than for taxiways or aprons. Also, some airports may use higher target levels for primary facilities (e.g, primary runways) than for secondary or tertiary facilities;
- Number of aircraft operations and aircraft size – Higher PCI values are typically required for facilities serving larger number of aircraft operations or larger and heavier aircraft; and

Type – Some agencies use different levels of service for different pavement types.

Identification of needs on the network level consists of the following steps:

1. Identification of pavement sections that require M&R treatments because of the level of service requirements or because of trigger levels;
2. Selection of M&R treatments for the sections identified in step 1;
3. Estimation of the costs required for the M&R treatments selected in step 2; and
4. Prioritization of projects if the cost of the treatments, estimated in step 3, exceeds the available budget.

Identification of needs also is typically evaluated for two time horizons:

- Short term planning for time horizons of about 5 years or less. For simplicity, it is also assumed that the analytical procedures, used for short term planning, do not include the generation and evaluation of alternative treatments in future years; and
- Long term planning for time horizons more than 5 years where analytical procedures can include the generation and evaluation of alternative treatments in future years.

## PRIORITIZATION, PLANNING AND BUDGETTING

M&R needs are prioritized, scheduled for implementation through programming, and then molded into a budget for a Capital Improvement Program (CIP). To be credible, the process of identification of needs and their prioritization must be consistent, transparent, and logical. The prioritization of needs is described for the same two scenarios used for the identification of needs – for short-term planning and for long-term planning.

Short term planning supports only limited prediction of future network condition without considering alternative future M&R treatments. The predicted condition of the pavement network can be used to evaluate the adequacy of different pavement preservation budgets. It is also possible to use the backlog of projects as an indication of desirable funding levels.

The first step in prioritization is the assignment of a priority level to each M&R treatment representing the unlimited budget. The priority level should reflect the main reason why the treatment is recommended. Priority levels are related to the levels of service used to identify M&R needs, and include safety, PCI, cost-effectiveness, and target priority PCI levels.

Long term planning and prioritization of needs, incorporating incremental cost-effectiveness analysis, has been successfully implemented by many transportation agencies on large highway networks (FHWA 1996). The implementation for airport networks is still in initial phases. A clear example of prioritization using cost-effectiveness analysis for an airport pavement network is provided by Tighe et al. (2004). The reasons for slower implementation include smaller airport pavement networks and greater importance of operational constraints.

Multi-year prioritization analysis need not include projects addressing the safety and critical priority levels as these projects are obligatory. Projects addressing the cost-effectiveness priority level and the target priority pavement condition level can be combined because both are prioritized on the cost-effectiveness basis. Consequently, the analysis has the potential to yield the most cost-effective combination of preventive maintenance projects.

The selection of projects is typically done using incremental cost-effectiveness analysis. This type of analysis has the potential to optimize the selection of M&R projects over a multiyear period and to evaluate the long-term impacts of the selection on the health of the pavement network. The result of multi-year prioritization analysis is a prioritized list of pavement preservation projects for different years that meet specific budget requirements.

Programming activities move projects from the initiation, prioritization, and budget stages to the design stage and to implementation. Budgeting builds on the results of planning and programming activities and produces a budget – a financial document that specifies how the money will be invested in airport infrastructure.

The type of projects included in the airport capital budget depends on local circumstances. Whereas large airports may have a budget dedicated solely to pavement preservation, capital budgets for smaller airports combine all projects concerning airfield infrastructure, and not just pavement preservation projects, to establish a CIP. For example, the budget may also include projects related to the expansion of the airfield pavements, operational improvements, and M&R of other airfield infrastructure, such as buildings and guidance systems. Some authorities prepare a combined budget for a group of airports they manage. Budgeting process should be viewed as part of asset management, the process that strives to manage all airport infrastructure assets together to achieve the efficient allocation of resources.

## PROJECT DESIGN AND IMPLEMENTATION

Project design determines the specific treatment type and design details required for the construction of the project, such as layer types, material properties, and construction details. The selected M&R treatments should address the primary cause of pavement deterioration and not just the distresses seen on the pavement surface during a PCI survey. Compared to the network level identification of needs and prioritization, the project level design requires additional and more detailed data. For complicated projects, the design process consists of a preliminary and final design stages. The preliminary design stage includes: (1) identification of alternatives, (2) design of alternatives and (3) the selection of recommended alternative. The final design stage includes detailed design of the selected alternative.

In order to properly evaluate and select the best treatment from a list of alternative treatments, the candidate M&R treatments need to be well defined and, in some cases, designed. This enables the analyst to estimate the specific costs and benefits of the competing treatments.

Candidate M&R treatments should be ranked according to their estimated benefits and costs. Estimation of benefits for maintenance treatments should be done in terms of the extension of pavement life of the original pavement, the estimation of benefits for rehabilitation treatments should be done in terms of the expected life of the rehabilitation treatment. The main benefit of a maintenance treatment is the difference between the life-span of the original pavement with and without the maintenance treatment. For example, full-depth repairs of PCC pavements may last 15 years or more, but may extend the life of a specific pavement section by only12 years, because the section may fail due to the progression of other distresses.

The methods used to select the best M&R alternative, or the recommended M&R alternative, include life-cycle cost analysis (LCCA), cost-effectiveness evaluation, and ranking analysis. The ranking analysis method is the most comprehensive and is recommended for use for important projects (Stroup-Gardner 2008, Walls et al 1998). If all other things are equal, the

alternative that is least expensive over time should be selected. LCCA facilitates the selection of the least expensive alternative. LCCA can incorporate the costs of not only the initial M&R treatments, but also the subsequent treatments. For example, the installation of retrofitted subdrains may have beneficial effect on more than one rehabilitation cycle.

Maintenance treatments, particularly preventive maintenance treatments, postpone the more expensive rehabilitation treatments. However, the cost of maintenance treatments must be paid much sooner than the cost of any future rehabilitation treatment. The need to pay now rather than later is explicitly recognized in the LCCA by discounting all costs to their present value. The analysis period, the period for which the costs are included in the analysis, should be sufficiently long to take into account all relevant consequences of alternative treatments. The FHWA publication *Life-Cycle Cost Analysis in Pavement Design* (Walls and Smith, 1998) provides a detailed description of the LCCA procedures.

The methodology of LCCA is relatively simple, and consists of the following steps:

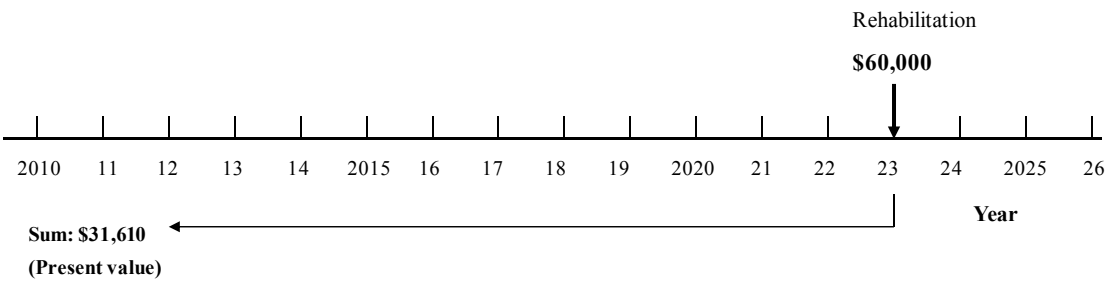
- Inclusion of all viable and practical alternative M&R treatments;
- Determination of agency costs. Agency costs include the initial construction costs and subsequent maintenance and rehabilitation costs throughout the analysis period;
- Determination of user costs. If the construction of M&R alternatives may have a different impact on airport operations and revenues, for example because of the differences in the length of construction, user costs should be included;
- Selection of economic parameters for LCCA (discount rate and analysis period);
- Calculation of the net present value of agency costs and user costs; and
- Selection of the alternative. The alternative with the lowest agency and user costs is the best from the economical point of view.

*Consider the following to determine the life-cycle cost-effectiveness of crack sealing:*

It is proposed to rout and seal a 4-year old asphalt concrete pavement in 2013, and repeat the treatment (reseal) in 2017). It is expected that these preventive maintenance treatments will postpone pavement resurfacing by three years. It is assumed that the cost of sealing is $1,100, the cost of resealing is $1,500, and the cost of rehabilitation (resurfacing) is $60,000.

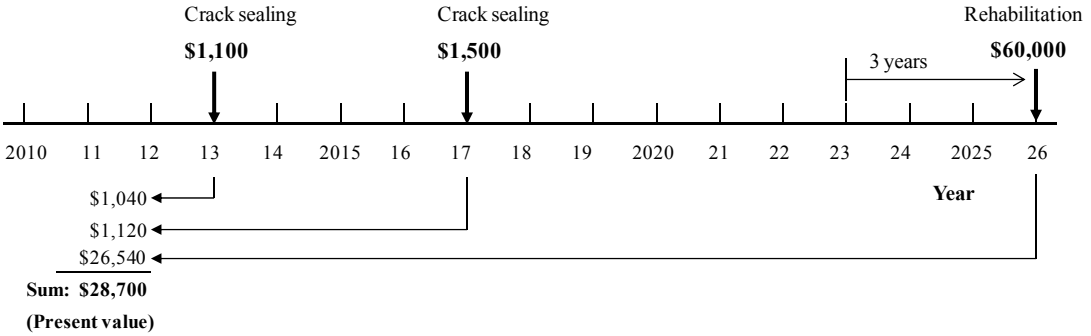
Without Preventive Maintenance

The following diagram assumes that without preventive maintenance a rehabilitation treatment will be required in year 2023. The present value of the rehabilitation treatment, considering a 6% discount rate, is $31,610.



With Preventive Maintenance

The following diagram shows that with the two preventive maintenance treatments, the rehabilitation treatment will be postponed to year 2026. The corresponding present value of this strategy, considering a 6% discount rate, is $28,700.




Comparison of Alternatives

Compared to the alternative without preventive maintenance, the alternative with preventive maintenance results in $2,910 saving ($31,610 – 28,700). Question can be asked, "What type of savings can be expected if the postponement of the rehabilitation treatment is only 2 years? To answer this question, LCCA was repeated and the results were plotted above. According to the figures above, the preventive maintenance alternative will be effective even if the postponement of the rehabilitation alternative is only two years.

**OPERATION, SUSTAINABILIY AND ENHANCEMENT**

Over 80 percent of airport agencies already operate or are developing an APMS (ARA 2010). The challenge for most of the agencies is not to establish an APMS, but to sustain and enhance its operation. The existing APMS operations and the sustainability of the APMS are closely linked: a successful operation of the system is one of the best guarantees of its sustainability. Long-term sustainability of APMS is an on-going process that should be considered during the initial implementation. In addition to ongoing system enhancements, there are situations where a structured comprehensive review of APMS operation is beneficial for improving the current practice and ensuring sustainability. The objectives of the review should be twofold:

- To determine enhancements that are required based on identified user needs; and
- To determine enhancements that may be beneficial based on the best appropriate practice. The best appropriate practice (BAP) is the desired state of practice that meets the particular agency's needs in the most appropriate and efficient way.

The common methodology used for the systematic assessment of the APMS is the gap analysis. As the name suggests the analysis is concerned with identifying the difference between the existing management process and the future desirable process defined as the BAP. The gap analysis consists of three basic steps:

- Assessment of the existing APMS activities against the BAP;
- Identification of activities where the agency has already achieved the BAP; and

- Identification of activities which require improvement to reach the BAP. For these activities, an implementation plan to guide future work should be prepared.

Another method that can help airport agencies to improve their pavement management practices is benchmarking. Benchmarking is similar to gap analysis in the sense that it provides a method for agencies to move from an internal focus to external focus in the search for best management practices. However, as the name suggests, benchmarking seeks to compare the operation of different organizations using objective, agreed-upon measures. In the airport context, the benchmarking measures include outcomes (e.g., average PCI for runways) and recourses (e.g., annualized pavement preservation cost per square yard of pavements).

## CONCLUSIONS

Airport pavement maintenance practices generally follow the objectives, management principles, and methodology of roadway pavement management practices. Airfield and roadway pavements are built and maintained using the same construction technology, materials, and methods.

All types of pavement preservation treatments, including maintenance treatments, preventive maintenance treatments, and rehabilitation treatments, should be considered together when developing pavement preservation strategies for individual pavement sections.

The main challenge facing airport authorities is not which M&R treatment should be used on a particular section, the main challenge is to justify that M&R treatments are necessary using a judicious and objective process, and to obtain funding for their implementation.

To be credible, the process of identification of needs and their prioritization must be consistent, transparent, and logical. The identification of airport pavement preservation needs should be based on approved or mandated levels of service. To facilitate funding requests for pavement preservation, the development of guidelines on appropriate levels of service for airfield pavements (for different types of airports and airport sizes) should be considered.

The selection of the recommended M&R treatments should be based on the life-cycle cost analysis. In addition to the LCCA, a systematic assessment of other treatment attributes, which are important to the customer and the agency, should also be carried out. These attributes may include, agency experience with the performance of the alternatives, impact on airport operations, and environmental and sustainability considerations.

A comprehensive review of APMS can be accomplished using gap analysis that identifies differences between the existing management procedures and those based on the best appropriate practice. The comparison between the results achieved by different APMS can be assessed through benchmarking.

## REFERENCES

ARA. (2010)., “Draft ACRP Synthesis Project 11-03, Common Airport Pavement Maintenance Practices”, Washington, D.C.

FAA. *(2006).,* "Advisory Circular on Airport Pavement Management Program", AC 150/5380-7A, US Department of Transportation, Federal Aviation Administration, Washington, D.C.

FAA. *(2007).,* "Advisory Circular on Guidelines and Procedures for Maintenance of Airport Pavements", AC 150/5380-6b, US Department of Transportation, Federal Aviation Administration, Washington, D.C., 2007.

FPP. (2001)., "Pavement Preventive Maintenance Guidelines", *Foundation for Pavement Preservation*, http://fp2.org.

Haas et al. (1994)., "Modern Pavement Management", Krieger Publishing Company, Malabar, Florida.

AASHTO. *(2001).,* "Pavement Management Guide", AASHTO Task Force on Pavements, ISBN-1-56051-155-9

FHWA. *(2006a).,* "Pavement Preservation Compendium II" , Publication No. FHWA-IF-03-21, U.S. Department of Transportation, Federal Highway Administration, Washington, 2006.

FHWA. (2006b)., "Pavement Preservation Toolbox", Federal Highway Administration, Washington, D.C., www.pavementpreservation.org/toolbox/resources.html.

Shahin. (1994)., "Pavement Management for Airports, Roads, and Parking Lots", Chapman & Hall, Washington, D.C.

Stroup-Gardiner, M., and S. Shatnawi (2008), "The Economics of Flexible Pavement Preservation", TRB 2009 Annual Meeting CD-ROM.

Tighe, S.L., Karim, M., Herring, A., Chee, K., Moughabghab, M., (2004), "Prioritization Methods for Effective Airport Pavement Management: A Canadian Case Study", Proceedings of the 6th International Conference on Managing Pavements, Queensland, Australia.

Walls, J., and Smith, M.R., "Life-Cycle Cost Analysis in Pavement Design – Interim Technical Bulletin", Publication No. FHWA-SA-98-079, U.S. Department of Transportation, Federal Highway Administration, Washington, DC (1998).

# The Benefits of Using a Geospatial Airport Pavement Management System at Denver and Tampa International Airports.

Michael T. McNerney, PhD, PE, M. ASCE
Katherine A. Keegan, PE, M.ENG

## ABSTRACT

Geospatial airfield pavement evaluation and management systems (GAPEMS) have been implemented at Denver and Tampa International Airports that use the power of GIS for data collection, data integration, 100% pavement inspection, and geospatial analysis of remaining life. Surface condition data collected as part of GAPEMS was used to determine traditional network level Pavement Condition Index (PCI) data. PCI data alone, however, was insufficient in terms of understanding the performance of the pavements. A remaining life analysis was developed to integrate additional data into the process to determine project level maintenance and rehabilitation recommendations.

A geospatial pavement evaluation and management system has the advantages of quick inspection time, full pavement coverage, project level data for maintenance and rehabilitation, and the ability to integrate other pavement data into the geospatial remaining life analysis. The benefits documented at both Tampa and Denver International Airports included an understanding of the failure mechanisms of the pavements, analysis of remaining life, and costs savings by shifting priorities to only those pavements in need of repair or replacement rather than pavements that reach a specific condition trigger level.

Denver performed a GAPEMS inspection of the entire airport in 2006 and 2009 and the integrated geospatial remaining life analysis and geospatial data have been invaluable in shifting priorities for reconstruction. Tampa Airport completed an airport inspection in two parts in 2008 and 2009 and the geospatial pavement management system has allowed for the combining of maintenance and operational activities into an integrated airport GIS system. This allows the Operations Department to geospatially identify areas of concern for the Maintenance Department to address and update.

## BACKGROUND

Airport geospatial information systems (GIS) are one of the most popular and effective management systems for airports to manage infrastructure at US airports. The Federal Aviation Administration (FAA) has recently instituted the Airport GIS program that will require all airports to submit new electronic airport layout plans to the FAA in GIS format. The American Association of Airport Executives (AAAE)

has this year instituted a program for storing and managing airport GIS information for subscribing airports.

A successful and effective application for GIS at airports is the management of airport pavements, one of the most important infrastructure items at an airport. The purpose of this paper is to describe the benefits of two implementations of pavement management systems that use a complete integration of GIS for data collection, data integration, pavement evaluation, and capital improvement planning. These two airports serve as a showcase for the capabilities of a true airport geospatial information system and both have achieved monetary and operational efficiencies through the implementation of airport geospatial pavement management systems. In addition, the geospatial pavement management system at Tampa is used as a tool by both Maintenance and Operations Departments to identify pavement and airfield deficiencies and repairs, creating a truly integrated airport GIS system.

The use of GIS with pavement management information systems is not new for airports or highway systems. Wisconsin Department of Transportation published one of the first reports describing the use of GIS for highway pavement management in 1990 [1]. The Federal Highway Administration through workshops and sponsor of research projects such as NCRRP 20-27 encouraged the use of GIS for pavement management [2]. The Federal Aviation Administration has been slower to encourage GIS use. However, authors such as Schwartz and McNerney offered in the 1990s methods for airfield pavement management using GIS [3, 4]. The FAA in its Pavement Management Advisory Circular, FAA 150/5380-7A, suggests the MicroPAVER software that was developed for the US Air Force [7]. MicroPAVER relies on a method of surface distress data collection that reduces data into a single pavement condition index (PCI) for sample sections. MicroPAVER has been updated to allow the display of the PCI in a GIS system, but does not use spatial analysis or GIS for data collection or pavement analysis of remaining life.

## DENVER GEOSPATIAL PAVEMENT MANAGEMENT SYSTEM

Denver International Airport identified a need to perform a more comprehensive pavement evaluation of the airfield pavements than had historically been done. Denver's need was to assess the risk of accelerated pavement deterioration of their pavements in light of concrete durability concerns at a neighboring airport.

The innovative GAPEMS technology was customized to address the airport's needs. The complexity in geospatial technologies as well as the scope and magnitude of the pavement evaluation makes this GAPEMS implementation one of the most unique, challenging and comprehensive airfield pavement management systems.

### *Georeferenced Distress Mapping*

A team of pavement inspectors rated distress in each of over 140,000 pavement slabs using ruggedized tabletPCs, ArcPad GIS, WAAS assisted GPS, and GPS enabled cameras. All inspections were accomplished in daylight hours during regularly scheduled runway and associated taxiway closures. All apron pavements were inspected without any closures. Details of the inspection can be reviewed in the 2007

FAA Worldwide Airport Technology Transfer Conference proceedings [8]. The distress data showed extensive shrinkage, map and D-cracking on specific runways and taxiways and extensive structural cracking in Apron areas.

### *Georeferenced Construction History*

Denver International Airport was constructed in 1992 and 1993 as a green field site with five runways, multiple taxiways, three concourse aprons, and a cargo apron. The concrete pavement was constructed by five different contractors each with their own onsite concrete batching plants. As with all FAA funded pavement construction projects, each project was required to submit FAA documents to close out the grant application through the submission of the final report. These closeout documents included daily inspection reports and other technical information about the mix design and cement certifications. These closeout documents were stored at the airport in paper form. As part of the pavement evaluation, data from these documents were extracted and recorded into GIS including the weather conditions for the actual day that the concrete was placed, the mix design, aggregate source, and the curing method for about 90% of the placements of concrete pavement constructed. Using ArcPAD software, polygons were drawn around and were attributed with the name of the contractor, type of placement, date, contract number, mix design, and any comments. The data collected represented 3,600 polygons of separate construction concrete placements from five contractors.

The construction data was analyzed using the HIPERPAV II software program developed for the Federal Highway Administration (FHWA) [9]. The software analyzes the likelihood of plastic shrinkage cracking and thermal cracking using weather data from cloud cover, wind speed, relative humidity and temperature, and concrete mix design information. The software also calculates the evaporation rate in the first 8 hours of placement.

The combined construction history and distress data collected showed there was a definite pattern of shrinkage cracking and map cracking relative to specific days of construction. Certain days had very high numbers of slabs with observed shrinkage cracking. The data show that two of the construction days each had 201 slabs placed that were showing shrinkage cracking. An analysis showed that 70% of slabs with shrinkage cracking occurred in the top 90 days out 352 construction days.

### *Georeferenced Nondestructive Deflection Testing (NDT)*

A comprehensive deflection testing program was completed for this project. As part of the pavement evaluation and pavement management system, all of the concrete runway and major taxiways were analyzed for structural capacity using deflection testing. Both the traditional method of point deflection testing as described in the FAA Advisory Circular using a heavy falling weight deflectometer (HWD) [10], as well as the Rolling Dynamic Deflectometer (RDD) method of continuous deflection measurements were used.

The results of the NDT indicated very stiff and strong pavement and supporting subgrade layers for runways and taxiways.

The GIS technology was used to graphically present the HWD and RDD data so that it could be related to any portion of the airport by zooming into the location in GAPEMS. Figure 2 is a screen capture of the RDD deflections as shown in the GIS for Runway 17R-35L and an inset photograph of the RDD.

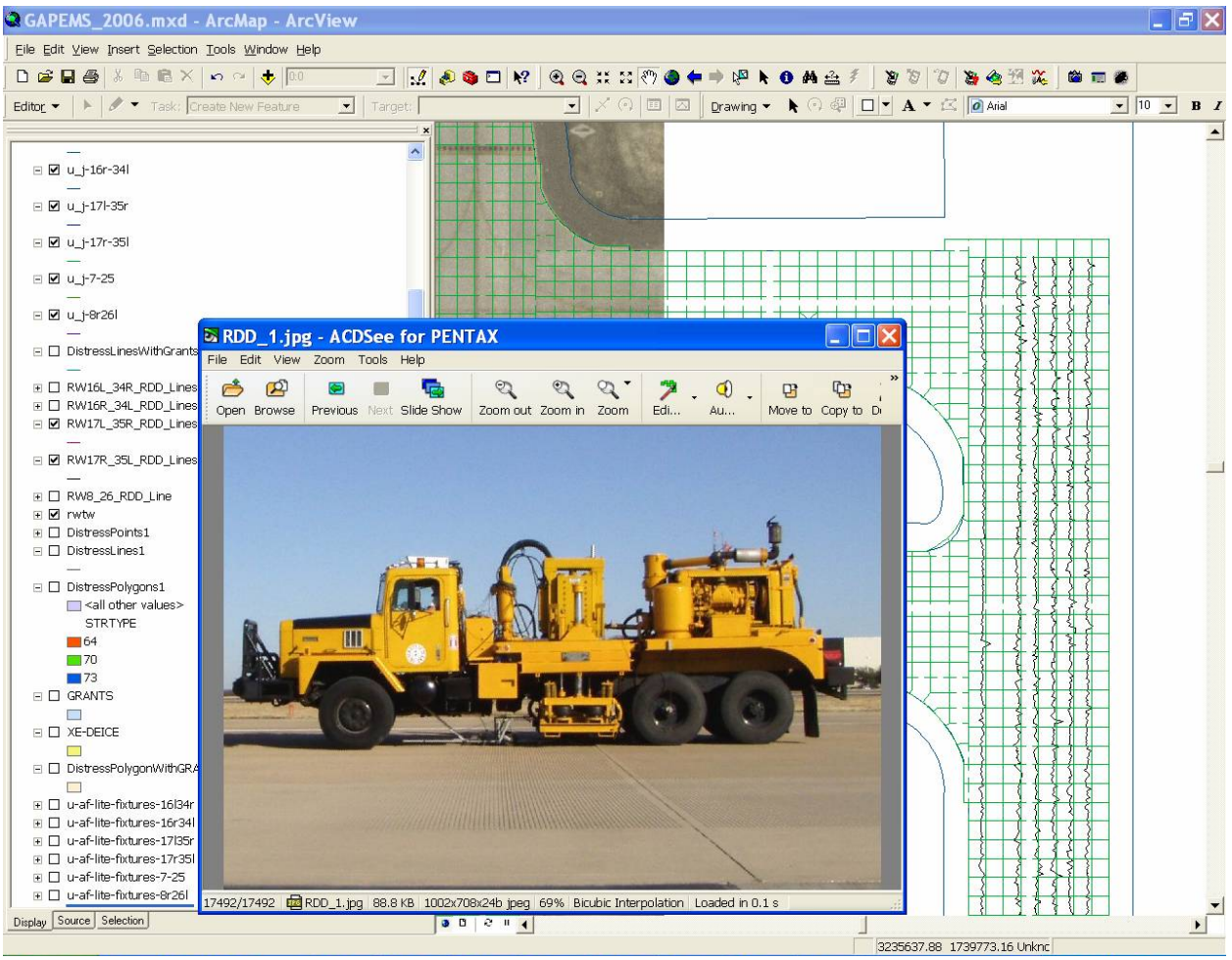



Figure 1 Screen Capture of GAPEMS GIS showing RDD and RDD Deflections at Denver

### *Georeferenced Core samples*

A major objective of the pavement evaluation at Denver is the determination of the remaining life of the concrete pavement with respect to concrete durability in the harsh freeze-thaw and runway deicing environment. The pavement evaluation plan includes making determinations on approximately 200 core samples with petrographic analyses or special testing.

GIS and tabletPCs were used in several unique ways during this project. First, the locations of the cores were selected based upon the georeferenced maps of map cracking, shrinkage cracking and durability cracking distresses.

The tabletPCs were also used to coordinate the review of underground utilities prior to coring the pavement section. GIS layers for in-pavement light fixtures and other electrical systems were integrated into GAPEMS and used to clear each core site. This reduced the time to get the necessary clearance for a specific area.

Finally, the results of the core testing were added to the GIS database after receiving the test reports. The test report number, referencing the comprehensive lab reports, was cross referenced and hyperlinked in the database as well.

Core data overlaid with construction and distress data in GAPEMS were used to assess relationships between surface condition, alkali-silica reaction (ASR) potential and construction details.

### *Geospatial Remaining Life Analysis*

All of the pavement evaluation data was either collected directly into GIS or was converted into data that could be visualized or accessed directly in the GAPEMS.

A new airfield concrete pavement remaining life analysis method was developed that used all of the geospatial data, including distress, construction history, evaporation rate, coarse aggregate properties, operational use, georeferenced photographs and structural capacity to assign an expected remaining life year group to each of the 140,000 slabs. This new method of remaining life analysis, termed FACS, relies on the geospatial analysis tools and is based upon four different primary failure modes of concrete pavements: fatigue [F], ASR [A], crazing and surface cracking [C], and spalling [S]. A complete description of the FACS method is in the ASCE *Civil Engineering* magazine and the ASCE 2008 Pavement Management Conference Proceedings. [11, 12]

### *Custom Geospatial Reporting Tools and Web Browser Interface*

Once a comprehensive remaining life analysis and remediation analyses were completed using the GIS data, customized reporting were developed using geospatial technologies. Any authorized person at the airport is able to use a web browser on the intranet to view all the data in the pavement management system, update the data, perform queries on the data, and print reports. The geospatial pavement management system also allows interaction with all other GIS data on the airport. A custom reporting function allows the user to print map books of pavement areas with any of the data layers displayed including non-pavement layers such as airfield lighting, pavement markings, and underground utilities.

### *Capital Improvement Planning*

A multiyear slab replacement program was developed based upon actual calculated areas of slab replacement by year and the associated costs that were developed years in advance. Using this data the airport changed the priorities of proposed airfield

pavement rehabilitation projects. Denver continues to use this tool to estimate annual funding requirements for pavement maintenance, repair and replacement.

## TAMPA GEOSPATIAL PAVEMENT MANAGEMENT SYSTEM

The Hillsborough County Aviation Authority (HCAA) owns and operates one hub commercial aviation airport, Tampa International Airport (Tampa),-and three general aviation airports all in Hillsborough County. A geospatial pavement distress evaluation and pavement management system were developed in response to a request from HCAA.

The geospatial pavement distress inspection was accomplished in two phases for funding reasons. The project included mapping of concrete and asphalt distresses for over 10 million square feet of pavement, over 13,000 georeferenced photographs of distress, a recommended rehabilitation design for Runway 18L-36R and associated taxiways and aprons, geospatial division of the concrete jointing pattern and dates of construction, and a geospatial remaining life analysis of the concrete pavements using the same method as was used at Denver. The pavement condition index was calculated for both asphalt and concrete pavements using the GIS distress data. Additionally, the Tampa project developed a pavement interface for both airport Operations and Maintenance Departments.

For the project level evaluation, the full strength Runway 18L-36R asphalt pavement GIS polygons were further divided into sample units, approximately 5,000 square feet in area, as required for asphalt testing by the American Society of Testing and Materials (ASTM) Standard D5340 Standard Test Method for Airport Pavement Condition Index Surveys [13]. The data used to develop these polygons was based upon the best available record information provided by the HCAA and aerial photography.

The polygon layer created was used for the data collection of distress mapping. Inspectors collected data using either Motion Computing LE 1700 TabletPC or Dell Latitude XT TabletPC employing the ArcMAP 9.3 GIS program connected by Bluetooth to Garmin GPS receivers and three GPS Ricoh Capilio 500SE digital cameras. The WAAS enabled GPS receivers permitted inspectors to accurately locate distresses within 2 to 6 feet for the pavement slabs or asphalt sample units being inspected.

The Tampa pavement areas inspected during this project were approximately 10,275,000 square feet. All identified inspection areas were rated at 100% coverage except for the additional asphalt taxiways on the east side of Runway 18L-36R and Runway 9-27. Overall the project pavement areas required 31 days of pavement rating. Runway 18L-36R was rated during a scheduled five day closure for maintenance activities.

HCAA selected and purchased two semi-rugged Motion Computing LE 1700 TabletPCs as shown in Figure 3 with Garmin 10 Bluetooth GPS receivers. The

model was selected for its ability to withstand frequent field use and the tablet has a sunlight readable screen.


Figure 2 Motion Computing TabletPC and GPS Camera owned by HCAA used for Data Collection

The most common concrete distress observed was map cracking. At the beginning of the project map cracking distress, which can be indicative of ASR, was not expected to be prevalent on the airfield because there had been no prior indication of ASR. However, map cracking was consistently identified with a distinct pattern throughout the project limits. All map cracking that was observed was of low severity which means that no scaling was occurring on the surface. The major cause of map cracking at Tampa was determined to be drying shrinkage rather than ASR. After Runway 18L-36R and Taxiway C had been inspected, the inspection procedure was customized to allow for a rating of 1 through 5 in the comments sections of the polygon identified as map cracking, with 1 being low severity and 5 being extremely severe map cracking.

The second most common concrete distress at Tampa was shrinkage cracking. Shrinkage cracking can be caused by plastic shrinkage during curing of the concrete, restrained shrinkage from construction, or drying shrinkage after construction. Most of the shrinkage cracking observed was surface cracking s only a few millimeters in depth.

### *Georeferenced Photography at Tampa*

In addition to collecting georeferenced pavement distress data, the inspection team also took GPS enabled photographs of distresses using Ricoh Caplio 500SE digital cameras as shown in

Approximately 13,500 photographs of distresses were taken in Phase 1 and are located in the HCAA database. The extensive photographs were very useful during the analyses period to select areas for further testing for petrographic analyses. The GPS imbedded photographs coupled with the GIS distress maps resulted in huge time savings, avoiding a second trip to determine core locations by having both distress mapping and photographs available for this task. Figure 5 shows a screen capture of the Tampa geospatial system with distresses both a thumbnail version and popup window of a geo-referenced photograph of a crack repair on Runway 18L-36R. The GeoLink software was configured to allow photographs to be imprinted with geographic coordinates, GPS measured altitude, and the date and time when the photographs were taken.

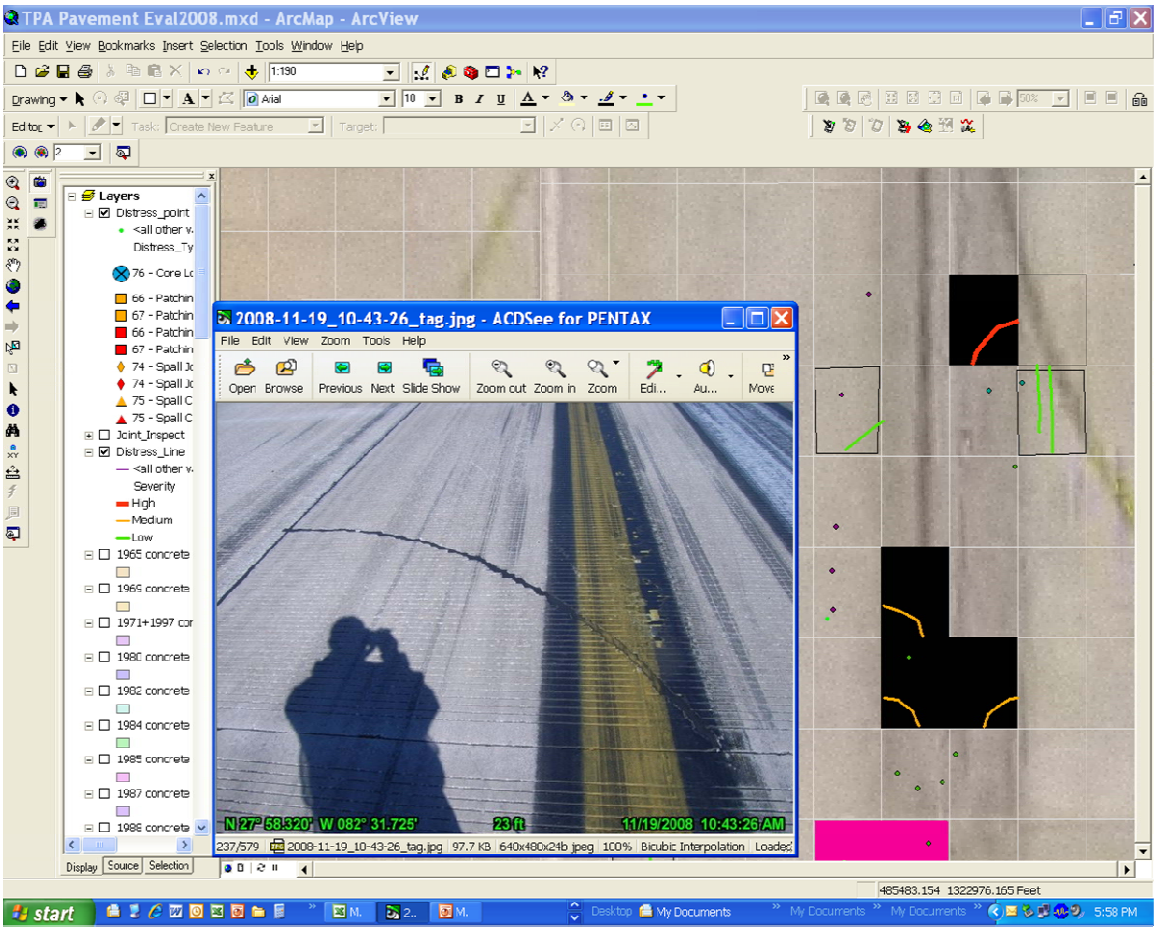



Figure 3 Screen Capture of GAPEMS Showing Georeferenced Photography.

### *Geospatial Calculation of Concrete and Asphalt PCI at Tampa*

AECOM has developed a copyrighted computer program to calculate the concrete PCI in accordance with the procedures of ASTM D5340, which uses the GAPEMS distress data collected in GIS format as the data source. On Runway 18L-36R there were 86 asphalt sample units of approximately 5,000 square feet each. On asphalt taxiways on the east side of Runway 18L-36R, the inspection team rated 59 additional asphalt sample units of approximately 5,000 square feet each. The GAPEMS distress data collected during distress mapping were used to calculate the PCI of these asphalt sample units. Asphalt PCI calculation differs from concrete PCI calculation in that rather than tallying the number of slabs within a sample unit exhibiting a certain distress, the square footage of certain distress and the lengths of cracks must be determined. In a typical manual inspection, the inspector carries a tape measure or rolling wheel to measure these values. In the geospatial method the lengths of cracks and the size of distress polygons are drawn to scale. The GIS calculates these values and also locates them automatically into the associated sample unit. For example, a crack or polygon that exceeds the boundary of a sample unit is recorded and the GIS assigns the values to each sample unit separately. This speeds up the field inspection by not having to be concerned with ‘coloring within

the lines', so to speak. The actual PCI calculations were completed outside of GAPEMS.

***Geospatial Remaining Life Analysis for Concrete Pavements at Tampa***

AECOM has developed the Fatigue, ASR, Cracking/Crazing/Scaling, Spalling (FACS) method of remaining life analysis using GAPEMS GIS data to calculate the remaining life of concrete pavement not at the section or sample unit level, but down to the individual concrete panel level. By mapping all distresses, the remaining life analysis is not limited to declining PCI as the measure of failure. Each of four different failure modes (Fatigue, ASR, Cracking/Crazing/Scaling, Spalling) can be evaluated independently and rehabilitation can be recommended at the slab level data. The remaining life slab assignment can then be put in a data shape file that can be displayed on the TabletPC or desktop computer.

The analysis of remaining life at Tampa using the FACS method provides a much better picture of the slab replacement requirements than does the traditional declining PCI method. The traditional PCI is a good measure of maintenance requirements because of the weight applied to spalling and patching distresses that are indications of maintenance requirements. In the FACS method only those surface distresses at severity levels that are factors in each of the FACS modes of failure are used to determine the remaining life.

The analysis of Runway 18L-36R showed only two sample units rated less than satisfactory (PCI < 71) with a runway average PCI of 92. However, using the analysis of individual slabs with the FACS method resulted in a selection of 13,000 square feet of slab replacements on Runway 18L-36R in the '0-2' years remaining life category. The analysis of Runway 18R-36L showed only one sample unit less than satisfactory PCI with an average PCI of 96. Only 1,250 square feet of slab replacements are identified in the '0-2' years remaining life category with the FACS method.

Table 1 shows the remaining life analysis summary by square feet of each of the facilities. The table shows that of the 3.8 million square feet of concrete pavement analyzed in Phase 2; only 2,930 square feet need replacement in the '0-2' year time frame. The estimated requirements decrease to 1,398 square feet in the '2-5' years time frame. There is only a modest amount of 38,115 square feet in the '5-7' years time frame.

In developing a future construction improvement program (CIP), a significant percentage of these identified pavements will require replacement in future time frames. However, in the next inspection cycle some pavement may show new distresses and some of these pavement in which distress have been identified will either show no further degradation of distress, or will have had maintenance actions performed to delay the need for replacement. Therefore, these estimated replacement numbers for different facilities and their square footages are relative guidelines. The accuracy of these numbers will depend upon the aggressiveness of the maintenance actions and the environmental and traffic factors that affect the

failure modes. The two most prevalent failure modes from the remaining life analysis are full depth cracking and surface deterioration.

Table 1: Distribution of Remaining Life by Facility in Square Feet for Phase 2

| Facility | Total Area SF | 0-2 Yrs SF | 2-5 Yrs SF | 5-7 Yrs SF | 7-10 Yrs SF | 10-15 Yrs SF | 15+ Yrs SF | 20 Yrs SF |
|---|---|---|---|---|---|---|---|---|
| RW 18L-36R | 8,185 | 0 | 0 | 0 | 0 | 0 | 8,028 | 157 |
| RW 9-27 | 6,788 | 0 | 0 | 0 | 0 | 894 | 5,894 | 0 |
| RWY 18R-36L | 1,688,693 | 1,250 | 671 | 18,334 | 79,167 | 208 | 1,588,028 | 1,035 |
| TL V | 15,602 | 0 | 0 | 0 | 0 | 794 | 14,808 | 0 |
| TW D | 174,980 | 0 | 0 | 1,875 | 2,197 | 30,722 | 140,186 | 0 |
| TW E | 44,248 | 0 | 0 | 0 | 1,094 | 13,008 | 30,146 | 0 |
| TW H | 3,414 | 0 | 0 | 0 | 0 | 0 | 3,414 | 0 |
| TW J | 27,821 | 0 | 0 | 0 | 1,241 | 3,693 | 22,886 | 0 |
| TW M | 69,504 | 0 | 0 | 0 | 0 | 2,326 | 67,178 | 0 |
| TW P | 34,207 | 0 | 0 | 2,890 | 0 | 938 | 30,380 | 0 |
| TW Q | 47,479 | 0 | 0 | 1,563 | 0 | 1,874 | 44,043 | 0 |
| TW R | 37,983 | 0 | 0 | 781 | 9,811 | 2,810 | 24,581 | 0 |
| TW S | 150,338 | 0 | 0 | 1,128 | 6,806 | 59,723 | 82,681 | 0 |
| TW V3 | 143,260 | 0 | 0 | 227 | 147 | 169 | 142,717 | 0 |
| TW V4 | 18,377 | 0 | 0 | 205 | 806 | 2,989 | 14,378 | 0 |
| TW V5 | 28,561 | 0 | 0 | 0 | 1,503 | 4,243 | 22,815 | 0 |
| TW V6 | 62,243 | 0 | 0 | 0 | 0 | 1,592 | 60,651 | 0 |
| TW W | 397,179 | 708 | 401 | 208 | 4,155 | 19,540 | 371,334 | 831 |
| TW W1 | 126,377 | 0 | 0 | 0 | 4,354 | 200 | 121,824 | 0 |
| TW W2 | 37,547 | 0 | 0 | 527 | 0 | 0 | 37,020 | 0 |
| TW W3 | 118,573 | 0 | 0 | 0 | 0 | 405 | 118,168 | 0 |
| TW W4 | 100,469 | 0 | 313 | 313 | 6,563 | 2,500 | 90,781 | 0 |
| TW W5 | 84,882 | 0 | 0 | 0 | 4,272 | 0 | 80,610 | 0 |
| TW W6 | 131,723 | 0 | 0 | 625 | 0 | 0 | 131,098 | 0 |
| TW W7 | 112,945 | 800 | 0 | 0 | 0 | 1,063 | 111,082 | 0 |
| TW W8 | 118,157 | 0 | 0 | 9,403 | 738 | 5,528 | 102,488 | 0 |
| TW Y | 33,109 | 171 | 14 | 36 | 0 | 0 | 32,888 | 0 |
| **Totals** | **3,822,645** | **2,930** | **1,398** | **38,115** | **122,853** | **155,219** | **3,500,107** | **2,024** |

The age of the concrete pavement has a significant impact upon the remaining life. The analysis of map cracking shows a higher correlation of distress severity with older pavements. The previous petrographic analysis revealed that microcracking related to drying shrinkage of 18 mm in depth was present in these older pavements with map cracking. Table 2 shows the distribution of remaining pavement service life by pavement age. Although, there are a few pavements in the '30-40' year life category in this project, the amount of pavement is very small.

The normal FAA design life for fatigue is usually 20 years for airports and some airports have justified a 30 year fatigue life. Airfield pavements subjected to air carrier loads that have service lives greater than 40 years are rare. In the remaining life analysis, even in those 30 and 40 year old pavements that showed no signs of fatigue or map cracking distress, it was not considered reasonable to forecast a remaining service life beyond another 10 years. That is not to say that those pavements will fall apart in 10 years, but rather another period of inspection and analysis is needed to reasonably forecast beyond the 10 year period.

Table 2 shows that 99.5% of the project concrete pavements were constructed in 1985 or later. As expected the concrete pavements constructed since 2002 are

performing well. Only 1% of the concrete pavements in the project have replacement issues in the next seven years.

Table 2: Distribution of Remaining Life by Pavement Age

| Year Constructed | Age Years | Area SF | 0-2 Years | 2-5 Years | 5-7 Years | 7-10 Years | 10-15 Years | 15+ Years | Totals |
|---|---|---|---|---|---|---|---|---|---|
| 1968 | 41 | 6,084 | | | | | | 100% | 100% |
| 1969 | 40 | 3,946 | | | | 35% | 21% | 44% | 100% |
| 1974 | 35 | 4,001 | | | | | | 100% | 100% |
| 1980 | 29 | 4,825 | | | | | 9.5% | 90% | 100% |
| 1985 | 24 | 540,092 | 0.2% | 0.04% | 3% | 0.1% | 1% | 95% | 100% |
| 1987 | 22 | 18,544 | | | 1.1% | 4.3% | 15.7% | 79% | 100% |
| 1988 | 21 | 251 | | | | | | 100% | 100% |
| 1989 | 20 | 233,516 | 0.4% | 0.3% | 0.2% | 4.6% | 1.1% | 93% | 100% |
| 1990 | 19 | 507,962 | | | 2% | 4% | 22% | 73% | 100% |
| 1997 | 12 | 1,355,256 | 0.06% | 0.03% | 0.8% | 6% | 0.02% | 93% | 100% |
| 2002 | 7 | 386,828 | | | | 1% | 7% | 92% | 100% |
| 2003 | 6 | 332,618 | | | 0.2% | 1% | 0.06% | 98% | 100% |
| 2006 | 3 | 325,115 | | | 0.07% | 0.5% | 1% | 98% | 100% |
| 2008 | 1 | 104,196 | | | | | 0.7% | 99% | 100% |
| | | | | | | | | | |
| **Totals** | | 3,823,234 | 0.08% | 0.04% | 1% | 3% | 4% | 92% | 100% |

### *Geospatial development of the Slab Replacement Program*

AECOM used the results of the pavement life remaining analysis data to create a GIS file that will allowed the airport to look at the slab replacement and repairs in a spatial view. The airport staff used the GIS data during a field visit to prioritize the slabs to be replaced with three different levels based upon budget availability. The agreed priorities were then designed, bid, and were replaced in 2009.

### *Geospatial Interface with Airport Operations and Maintenance at Tampa*

Tampa International Airport requested that the geospatial airport pavement management system also have the capability for airport operations to identify pavement deficiencies in a GIS layer and enable the maintenance department to view a photograph and to log in GIS when the discrepancy has been repaired. A new motion computing tabletPC model J3400 was purchased by HCAA specifically for airport Operations. The tablet was customized and a separate GIS layer was created for the GIS data. Additional features were added after a trial period that allowed Operations staff to mark with a polygon any areas of the airport that they close either temporarily or by NOTAM for longer periods. In addition, the application was designed to record the data in the field including notes and links for the daily safety inspection report. The report can then be printed from the tablet in the office or connected to the airport network.

The Maintenance Department used a tabletPC to track pavement repairs with the type of repair materials used. The Maintenance tabletPC has not yet been integrated with the work order system at this time.

### *Geospatial Mapbooks of Pavement Data at Tampa*

At Tampa they do not yet have an integrated Airport GIS for several departments but it is something they are planning for the future. Therefore, HCAA chose not to invest in a web browser interface for the GIS. In order to keep pavement management data readily available to the engineers and managers, the HCAA elected to fully implement a system of "Mapbooks", an ArcGIS tool, that cover the airport in 8.5 by 11 inch format in a 1 inch equals 100 foot scale. The airport is covered in an index of mapbooks. Since the primary runways run north south, the mapbooks were ordered in a vertical priority for printing. The engineer has the option of printing only a specific mapbook or mapbooks, or the GIS will print only those mapbooks in order that has the data that has been selected for printing. As Shown in Figure 4, the individual mapbook allows the presentation of distress on each slab at a scale such that each slab is visible.

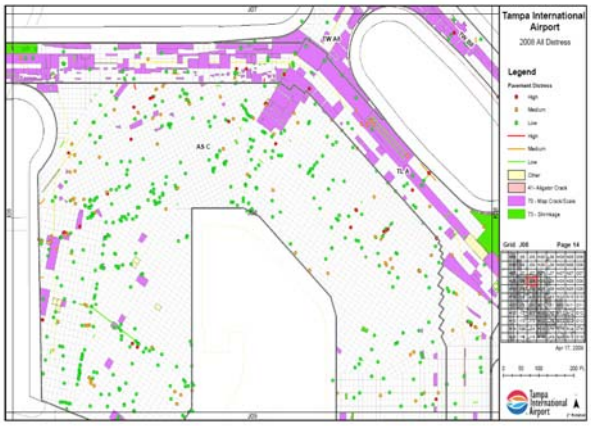



Figure 4 Example Mapbook J08 showing Distress at Tampa

## BENEFITS OF THE DENVER AND TAMPA GEOSPATIAL PAVEMENT MANAGEMENT SYSTEMS

AECOM has implemented similar yet fully customized geospatial pavement evaluation and management systems at Denver and Tampa Airports. Both systems use GIS for data collection, pavement evaluation and Capital Improvement Planning. Denver added more additional testing including petrographic and structural testing to meet specific pavement deterioration concerns. Tampa added an interface for Maintenance and Airport Operations Divisions.

Both airports use distress mapping with TabletPCs, GIS, and GPS that provided a complete picture of the distribution and type and severity of distresses. The TabletPC data collection method is efficient and provides a detailed representation of the condition of the pavements and specific distress distribution, more so than a traditional PCI inspection.

Both airports purchased GPS enabled cameras and have a database of georeferenced photographs of distresses that can be accessed in the geospatial pavement

management system by location or by date. In the pavement remaining life analysis the photographs provided invaluable information to review pavement condition. In the case of Denver, they have a series of photographs for comparison taken in 2006 and 2009 that can be used to help analyze the deterioration of the concrete surface. The georeferenced photographs of the deterioration of surface cracking due to ASR provided visual evidence of the rate of deterioration and justification for slab replacement.

The integration of a geospatial pavement management system with other GIS data is a major benefit. At Denver, it is fully integrated with the airport GIS and all other GIS data is available with the GAPEMS web browser display. For example, the utility data and airfield lighting data can be displayed either on an aerial photograph or CADD airport layout either with or without any layer of the pavement management data.

At Tampa, the mapbooks provide a quick reference of the GIS data. It is considered a great benefit to the airport that a theme of data can be printed in a mapbook form. The theme may be a type of distress, the pavement age, or the PCI of the pavements.

In Denver, the special remaining life analysis was developed because traditional PCI did not provided detailed enough information to identify and address material and construction related deterioration issues. The answers to the extent of ASR deterioration were only solved by a geospatial approach to analyzing distress, petrographic core samples, operational factors and the analysis of construction history by daily slab placement.

Further, in Denver the results of the distress survey and the remaining life analysis caused a shift in the planned reconstruction of Runway 17L-35R. The initial plan was to replace the keel section of the runway since it receives most traffic. However, after the slab by slab remaining life analysis and a review of slab specific failure modes, it was determined that only a few slabs in the keel section needed replacement as well as a a systematic pattern of spalling occurring in the columns of slabs just outside the keel section. Using the GIS data the number of slabs to be replaced was designed and a contract let that resulted in bringing the runway back to almost new condition for $20 million in slab replacement costs. The annual slab replacement program is continually readjusted based upon the constantly updated geospatial data.

At Tampa, the geospatial pavement management system and the full mapping of distress data saved the airport millions of dollars by securing FAA funding. In 2009, less than a year after the geospatial distress data was collected, the airport needed to reconstruct a taxiway in conjunction with adjacent new bridge construction. The FAA did not accept the reconstruction of this taxiway since it had a high PCI value of 80 and was unwilling to provide funding support. Using GAPEMS, AECOM showed that the taxiway had significant map cracking, was the oldest pavement on the airfield and was one of the highest priority pavements to be replaced based upon a remaining life analysis that showed that the failure mode of map cracking had significant potential to result in scaling leading to a failure of the pavement.

Through presentation of geospatial data and the associated narrative report, the FAA reversed its decision and granted funding for the project. This solution fully paid for the decision to implement the geospatial pavement management system at Tampa.

In closing, possibly the biggest benefit of a geospatial airport pavement management system is the ability to continuously capture and present data to solve future problems and optimize pavement decisions. In municipal GIS systems, it was often cited as a benefit that a GIS will pay for itself in the future when inevitably a problem will arise that is unforeseen but because the GIS data is available, a quick GIS analysis can provide the best solution to the problem that otherwise would require a guess. Both Denver and Tampa Airports have already had pay back using GAPEMS by leveraging geospatial data to apply a specific and detailed pavement analysis that could not have been answered by traditional pavement management systems.

REFERENCES

1. "Pavement Management Decision Support Using A Geographic Information System," Wisconsin DOT, Madison, WI, May 1990.
2. "Adaptation of Geographic Information Systems for Transportation, NCHRP Report 359, National Academy Press, Washington DC, 1993.
3. Schwartz, C.W., G.R. Rada, M.W. Witczak, and S.D. Rabinow, "GIS Applications in Airfield Pavement Management," Transportation Research Record 1311, Washington DC, 1991
4. McNerney, M. T., "Airport Pavement Management as Part of an Integrated Airport Geographical Information System (GIS)," proceedings ASCE Transportation Congress, B. Kent Lall and Daniel L. Jones, Jr., Editors, ASCE New York, NY, October 1995, Vol. 1, pp. 573-587.
5. Bay, James A., Kenneth H. Stokoe, II, Michael T. McNerney and B. F. McCullough, Continuous Profiling of Runway and Taxiway Pavements With the Rolling Dynamic Deflectometer (RDD) at the Dallas-Fort Worth International Airport, Transportation Research Record No. 1639, 1998, pp. 102-111
6. Bay, James A., Suttisak Soralump, Kenneth H. Stokoe II, Michael T. McNerney and David VanVleet, "Evaluation of Runway Pavements at the SEA-TAC International Airport Using Continuous Deflection Profiles Measured with the Rolling Dynamic Deflectometer," Transportation Research Record 1716, 2000, pp. 1-9.
7. FAA Advisory Circular AC150/5380-7A, Airport Pavement Management Program, Washington DC, September 1, 2006.
8. McNerney, Michael T. and Mark E. Kelley, "The Use of TabletPCs and Geospatial Technologies for Pavement Evaluation and Management at Denver International Airport," Proceedings of 2007 FAA Worldwide Technology Transfer Conference and Exposition, Atlantic City, NJ, April 15-18, 2007, Paper P07047.
9. HIPERPAV II software, High Performance Paving Software, The Transtec Group, Austin, Texas, available on Line at www.hiperpav.com.

10. FAA Advisory Circular AC150/5370-11A, “Use of Nondestructive Testing in the Evaluation of Airport Pavements,” Washington DC, December 29, 2004.
11. McNerney, Michael T. “Remaining Service Life Analysis of Concrete Airfield Pavements at Denver International Airport using the FACS Method,” Proceedings of the 2008 Transportation and Development Institute Pavement Conference, Bellevue, WA, October 15-18, 2008.
12. McNerney, Michael T., “Precise Prediction,” Civil Engineering Magazine, American Society of Civil Engineers, Washington DC, August 2008 pp56-63, 84.
13. ASTM D5340- 04 Standard Test Method for Airport Pavement Condition Surveys, ASTM International, West Conshohocken, PA

# Airport Planning for Older Air Passengers

William J. Sproule[1]

[1] Department of Civil and Environmental Engineering, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931; phone: 906-487-2568; e-mail: wsproule@mtu.edu

## ABSTRACT

As our population ages, older air passengers face many challenges as they travel. Today there are over 40 million Americans over 65 years and projections are that this number will more than double by 2050 to almost 90 million. This paper defines what is old, examines the characteristics and demographics of older air passengers, describes what it is like to be old, explores the challenges that older passengers face in today's airports, and then discusses the approaches and opportunities as we plan and design future airports.

## INTRODUCTION

Aging is one of the most important demographic dynamics affecting families and societies throughout the world. The growth in our older population is challenging policy makers, families, businesses, health care providers, and others, to meet the needs of aging individuals. The aging of our population has profound implications on our transportation system.

## WHAT IS OLD

Old age is an age at which a person is nearing or surpassing the average life span. Other terms for these persons include seniors, senior citizens, and the elderly. However, the boundary between middle age and old age cannot be defined exactly because it does not have the same meaning in all societies. People can be considered old because of certain changes in their activities or social roles. For example, people may be considered old when they become grandparents or when they begin to do less or different work and they retire.

In the United States, the age of 65 has traditionally marked the beginning of old age because until recently people became eligible to retire at this age with full Social Security benefits. In the 1880s, Otto Von Bismarck picked 65 as the number for old age in preparation for Germany's first pension plan. At that time the average life expectancy was 45 years and he thought the program would not need to cover many citizens if they used twenty years beyond life expectancy. In 1950, life expectancy increased to almost 70 years and by 1990 it was 75 years. A child born today is expected to live to over 80 years and once we reach age 65, we can expect to live 15 to 20 more years.

Today there are more than 40 million people in the United that are 65 years and older, and this represents about 1 in every 8 Americans. Figure 1 shows the growth of older Americans from 1900 to today and projected to 2050 (Federal Interagency

Forum, 2008). The total population in the United States in 2050 is forecasted to be almost 440 million and there will be almost 90 million people 65 years and older. This represents almost 1 in 5 Americans (20%) and the fastest growing group will be 85 years and older.

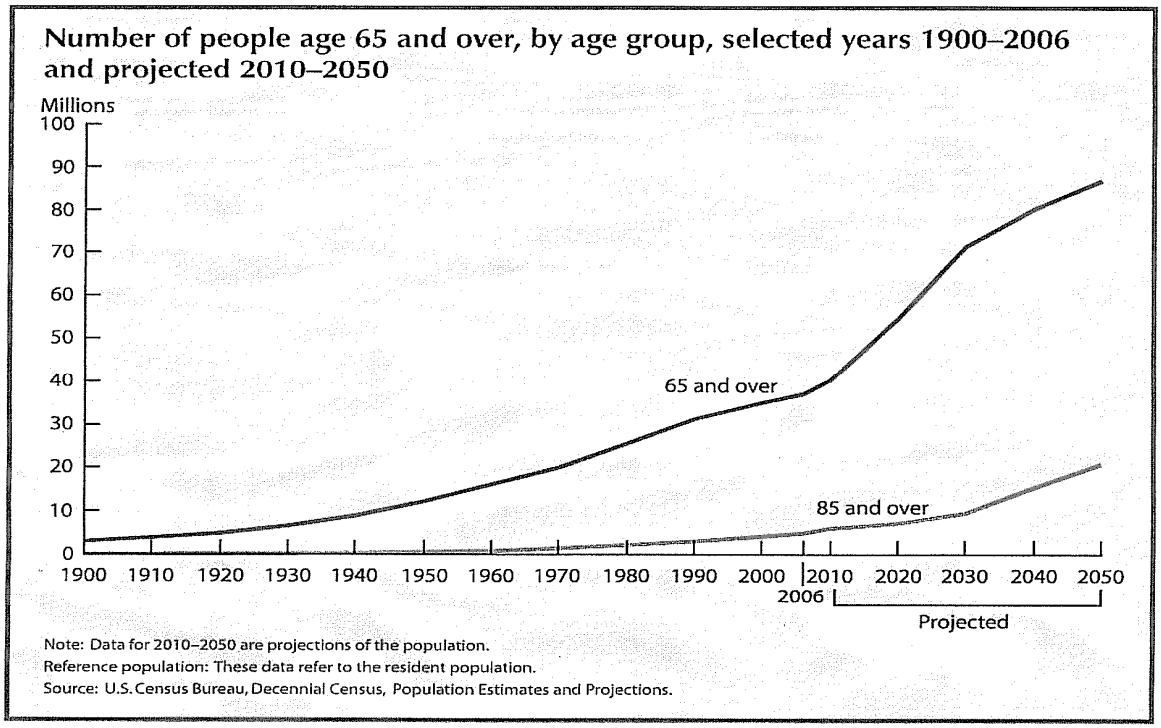



Figure 1 – Number of Older Americans

*Source: Older Americans 2008 – Key Indicators of Well Being, Federal Interagency Forum on Aging Related Statistics, Washington, DC, 2008.*

There will be a substantial growth in the number of older people in the 2010 to 2030 period as the first Baby Boomers (people born between 1946 and 1964) turn 65 in 2011. Recent U.S. Census Bureau reports (U.S. Commerce, 2004, U.S. Commerce, 2005, and U.S. Department of Health, 2009) provide excellent summaries of population aging. The reports examine data for the population 65 and older and actually identify three additional age groups – the young old (those aged 65 to 74), the oldest old (those aged 85 and older) and centenarians (those aged 100 and over). Today the oldest old population in the United States is about 5 million and it is projected to grow to over 20 million by 2050 and this reflects the movement of Baby Boomers into the oldest old category. As the older population grows larger, it will also grow more diverse, reflecting the demographic changes in the U.S. population as a whole over the last several decades.

A valuable graphic to help understand demographics is the population pyramid. Figure 2 shows the population pyramids for the years 2000 and 2050. The relative sizes of age groups are shown as "cohorts". Each pyramid shows males and females cohort as a percentage of the total population. A cohort describes a group of people born in a year or period of time that differentiates them from other groups. For example, persons born in the Depression form a cohort, and baby boomers form another cohort. Cohorts often differ from one another in demographic and social characteristics, and they generally develop attitudes and value systems based on social, economic, and political experiences that are unique to their time and circumstances. People retiring in the 1990s are not only more educated and affluent than previous cohorts of retirees; they are generally healthier and more vigorous and have a more positive attitude toward retirement. One must recognize cohort differences and diversity within the older population so that one responds to meet special needs of the elderly instead of a generalized picture of the "typical older person".

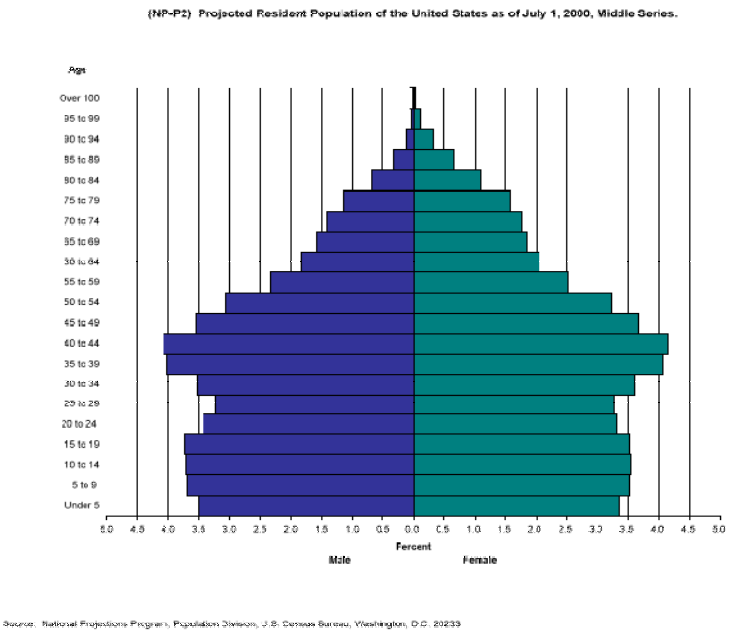
## OLDER AIR PASSENGERS

In the coming decades, planners will be confronted by an unprecedented challenge. American society is aging – not only will there be more people, but they are living longer. This trend has enormous implications on our society.

Last year there were over 700 million air passenger enplanements at U.S. airports. The recent FAA Aerospace Forecast (Federal Aviation Administration, 2010) projected that the domestic passenger enplanements will grow at an average rate of 2.4 per cent over the next twenty years with mainline carriers growing slower than regional carriers (2.2 % versus 3.0% per year, respectively). International passenger enplanements are expected to grow at an average of 4.2 percent per year over the next twenty years, with India, China, and Taiwan forecast to be the fastest growing markets. It is projected that total air passenger enplanements in the United States will exceed one billion in 2023.

To determine the number of older air passengers, results from the National Household Travel Survey (NHTS) for long distance travel in the United States are used (U.S. DOT Bureau of Transportation Statistics, 2006). It is estimated that about 15 million air trips in the United States were made last year by persons 65 years and over. This represents about 0.4 air trips per person. A similar rate has been observed for persons under 25 years old, while for travelers between 25 and 64 the rate is about 1.0 air trip per person.

One has to be careful when using NHTS data and FAA passenger enplanement as enplanements include non-residents and a trip may often include several enplanements. Over 75 percent of air trips are made by persons in the 25-64 age group and about 6 per cent are made by persons 65years and over. The 25-64 age current group represents about 53 per cent of the U.S. population while the over 65 group is about 13 per cent of the U.S. population.



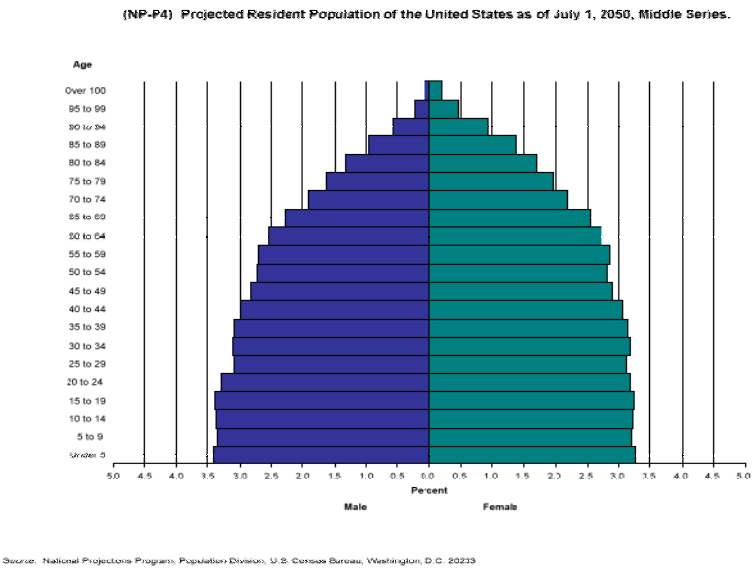




Figure 2 – Population Pyramids for 2000 and 2050

*Source: U.S. Census Bureau, Washington, DC*

Using Census projections, the 65 years and over age group is expected to grow to 16 percent in 2020, 19 percent in 2030, and 20 percent in 2050. If one applied the trip rate directly for all people over 65, there would be 22 million air trips by persons 65 years and over in 2020, over 30 million in 2030, and 35 million in 2050. In the next twenty years, the number of older passengers would be double the current levels. However one must be cautious when applying the rates and realize that the rate probably does not apply to everyone 65 years and over. There will be differences in the cohorts. The young old group makes more air trips than older cohort groups. The proportion of leisure time that older Americans devote to sports, exercise, recreation, travel and socializing declines with age. Americans aged 75 and over spent a higher proportion of their leisure time reading, relaxing, and thinking. However chances are that senior propensity to travel could actually be greater because seniors are healthier and more familiar with air travel than their predecessors. Although there may be some question and uncertainty on specific numbers, everyone agrees that there will be many more older air passengers in the future.

## WHAT IS IT LIKE TO BE OLD

As we age, there is often a general physical decline and we become less active. The medical study of the aging process is gerontology and the study of diseases that afflict the elderly is geriatrics. Aging is a process that affects each individual differently. Some individuals remain relatively healthy while others succumb to one or another age-related disease. Even at age 80, about half of the population is still cognitively normal. A key point to remember is that there is a very wide range in people in any age group or cohort.

It is important to view aging as a normal process. Many of the changes that we associate with aging, however, such as forgetfulness, joint stiffness, or senility may be caused by disease instead of by natural aging. Memory loss and confusion may actually be the result of small strokes, reactions to medication, poor nutrition, depression, or a number of other factors. Sixty percent of the behaviors that appear to be caused by senility accompanying the aging process are related to other causes and have found to be either reversible or treatable.

Other changes, such as vision loss, hearing impairment, slower movement, and loss of bone and body mass, can be expected with age. Many of these changes begin much earlier in the life cycle. For example, hearing loss begins about age 25 and changes in normal vision are generally apparent by age 40. Fortunately, most of these changes are very gradual and there are exceptions to the rule. Research on changing reaction times and learning with age provide good examples. While a decline in reaction time is a common attribute of aging, studies show that older people in their sixties who have remained athletic have faster reaction times than nonathletic people who are in their twenties.

As Frederic Verzar, the Swiss gerontologist once said, “Old age is not an illness; it is a continuation of life with decreasing capabilities for adaption” (Howe, 1994).

A brief overview of aging is described (Centers for Disease Control, 2007) and grouped in three categories – changes in the one's sensory system, physical changes, and mental changes.

*Sensory Changes*

Sensory includes changes in vision, hearing, and other systems. Even with corrective lenses or medical treatments, millions of Americans have less than 20/20 vision. A decline in vision generally begins around age 40 and by age 65 almost everyone will experience some loss in ability to focus, to resolve images, to discern colors, and to adapt to light. By age 75, poor vision is common. Over 60 percent of those considered visually impaired are older persons.

Hearing loss is perhaps the most serious of the sensory impairments that comes with age. The deficit cannot be as easily perceived by others (as can vision loss) and often receives less public sympathy, yet loss of hearing can be very isolating. Hearing loss is also gradual. It begins in middle age and affects more people than any other chronic condition. About 40 percent of all older people suffer a hearing loss that affects their communication and social relationships.

The most common loss is the ability to hear high frequency sounds and voices. The high-frequency sounds of speech become distorted making it difficult to follow conversation. Older people with hearing loss also have a difficult time distinguishing background noise from speech. Noise in general is also most distracting to older people than to younger ones. This may be especially true in public areas like airports where human and voice-synthesized announcements compete with many other noises for a person's attention. The problem of background noise may be exacerbated by the use of a hearing aid which amplifies all sound not just the conversation one would like to hear.

Taste, smell, and sensitivity of the skin also decline somewhat with age. For example, most people over the age of 65 have lost 50 percent of their taste buds and this decline continues with aging.

*Physical Changes*

Muscle strength in the heart is reduced (the degree is dependent on physical activity throughout life), thus reducing output of blood by the heart. Changes in heart rate and in lung capacity during the aging process lead to loss of reserves. Under stress or exertion, older people have less reserve than younger people, resulting in increased tiring and increased time needed to complete tasks.

In old age there is a decrease in strength, agility, and fine motor control. These conditions may also be exacerbated by arthritis, stroke, or other diseases. This decrease results from a loss of muscle mass and degeneration in the joints. Lack of muscle tone and impaired muscle coordination contribute to the loss of ability to balance rapidly or compensate for sudden loss of balance. Negotiating the environment therefore becomes more difficult and the tendency for falls is increased.

*Mental Changes*

Other noticeable changes that seem to come with aging are those that affect mental status, learning, and memory. Although it is not clear that these changes directly relate to cognitive functions, when combined with a slowing of the nervous impulses, they contribute to the older person generally needing longer to perform tasks or to respond to instructions.

Changes in memory are often the source of humor about aging and indeed some memory loss does seem to occur as a natural consequence of age. Research, while not conclusive, indicates that both short-term memory and long-term memory show impairment with age, with short-term memory affected a greater extent. The ages at which these changes occur differ by individual, but in general they will occur by the age of 80.

Senility is often referred to as dementia and is the condition of intellectual and cognitive impairment that is associated with age. These are behaviors caused by either temporary or irreversible organic brain syndromes. Impairments include losses of orientation, of short- and long-term memory, of visual-motor coordination, of learning and retaining spatial arrangement, of the ability to abstract and assimilate new information, and of the ability to carry out sequential tasks or remember time sequences.

*Changes in Confidence*

When an older person cannot function in an environment that was once familiar, self-concept and confidence may suffer. An inability to use transportation or remain involved in the community will often lead to a downward spiral with loss of self-esteem leading to further isolation. Older people, like those in all age groups, cherish their independence.

The American Disabilities Act (ADA) is a landmark law that protects the civil rights of persons with disabilities. ADA Accessibility Guidelines (Federal Aviation Administration, 1999) have been established for airport planning and design. Disabilities can be classified into three main categories – vision impairment or legal blindness, deafness or hardness of hearing, and mobility impairment (Landrum, 2010). Although many of the ADA guidelines are applicable for older passengers there is a significant difference between disabled and aging passengers. Disabled passengers are acutely aware of their limits while older passengers may also experience reduced mobility, agility, and stamina, but they do not want to be considered disabled (Landrum, 2010).

## IMPLICATIONS ON AIRPORTS

Seniors are having problems using airports. Waiting, long walking distances, and difficulties in wayfinding in terminals are anxiety provoking and can deter seniors from using airports (Wolfe, 2003, and Corgan, 2008).

Waiting

A major cause of anxiety and frustration for the airline passenger is the inevitability of having to wait. There are four primary processors at which passengers will often have to wait – check-in, security, gate area, and baggage claim. Passengers are often frustrated by the perceived waste of time and are anxious about missing their flight which can be very troubling to older passengers and inexperienced air travelers. There are also numerous other locations at an airport where waiting is common and these include such activities as restaurant and other food services, concessions, restrooms, stores, intra-airport circulation systems, and ground access services. In addition, waiting generally requires long periods of time standing which can be fatiguing, especially for older air passengers. Waiting has become an expected part of air travel today but considerations should be made for older air passengers by providing more seating particularly in the area of the four primary processors – check-in, security, gate area, and baggage claim.

Walking

As our airport terminals have expanded to handle growing passenger demands, the walking distances have also increased where they have become an obstacle confounding older passengers.

The suggested maximum walking distance between major functions is 1000 feet (300 meters), but greater distances can be accepted provided a form of mechanical assistance is available (International Air Transport Association, 2004). Moving walkways have become a popular alternative in many terminals but they may create difficulties for some older passengers as they require standing and there may be balance issues as they step on or off the device. Automated people mover systems are being used for longer distances but again there may be standing and balance issues. It is suggested that more seating be provided on the vehicles and in waiting areas. A third option is electric passenger assist vehicles/carts. The key issues for older passengers is the ease of boarding and disembarking from the carts, waiting, and restrictions or limitations on where they can be used in the terminal.

The second aspect of walking is the change in levels that occur in most terminals. Stairs are the least desirable option for older passengers. Research indicates that 30 percent of the population 75 years and older have trouble using stairs. Escalators also pose problems to many older passengers as they face balance issues as they step on or off an escalator. The elevator is clearly the preferred option for moving between levels but they are often a hidden option and many older passengers do not realize that an elevator is available.

## Wayfinding

Wayfinding is much more than signage - it actually includes several forms of communication. Wayfinding allows passengers to orient themselves within the facility, provides critical information, and gives passengers clear and simple directions. There is a lot of information that is presented at an airport and finding one's way in airport terminals and getting to and from an airport can pose a challenge for all air travelers but research indicates that way finding is in general more challenging for older people and the risk of disorientation is higher. It is also a challenge for new or inexperienced air travelers.

Most airports rely on maps and information displays to facilitate way finding and special care must be taken in the design of these maps and placement location. Even with excellent maps and displays, many older passengers prefer directions from a person. Volunteer guide programs have been particularly helpful in providing information and reassuring older passengers that they are going in the right direction.

One area that is particularly confusing for passengers is security. The enhancement of security requirements over the past ten years has posed additional difficulties for older passengers. The waiting and standing in line, identification checking, removal of personal items and clothing and shoes and placing them in tubs, maneuvering baggage onto and off of the conveyor belt, and the walking through the passenger screening device. Older passengers are often confused by the many activities that occur in the screening process, the changing procedures, and the verbal commands or information that are provided.

## CONCLUSIONS

People in the United States are living longer and healthier lives than ever before. In the next ten to twenty years a growing group of older air passengers will emerge who will have the time and money to fly, but may be reluctant to do so because airports and air travel make it a very unpleasant experience for older air passengers. We must plan our airports to meet the needs of older air passengers and we must always think about waiting, walking, and wayfinding.

## REFERENCES

Centers for Disease Control and Prevention and the Merck Company Foundation (2007), *The State of Aging and Health in America*, Whitehouse Station, NJ, 2007 (www.cdc.gov/aging).

Corgan Associates and others (2008), *Innovations for Airport Terminal Facilities*, ACRP Report 10, Transportation Research Board, Washington, DC.

Federal Aviation Administration (2010), *FAA Aerospace Forecast, Fiscal Years 2010-2030*, FAA Aviation Policy and Plans, Washington, DC.

Federal Aviation Administration (1999), *Access to Airports by Individuals with Disabilities*, Advisory Circular 150/5360-14, Washington, DC.

Federal Interagency Forum on Aging Related Statistics (2008), *Older Americans 2008 – Key Indicators of Well Being*, Washington, DC.

Howe, Deborah A., Nancy Chapman, and Sharon A. Baggett (1994), *Planning for an Aging Society*, Planning Advisory Service Report Number 451, American Planning Association: Chicago, IL.

International Air Transport Association (2004), *Airport Development Reference Manual*, 9th Edition, Montreal, Canada.

Landrum and Brown, and others (2010), *Airport Passenger Terminal Planning and Design*, ACRP Report 25, Transportation Research Board, Washington, DC.

U.S. Department of Commerce Census Bureau (2004), *We the People: Aging in the United States*, CENSR-19, Washington, DC.

U.S. Department of Commerce (2005), *65+ in the United States*, Census Bureau Report P23-209, Washington, DC.

U.S. Department of Health and Human Services (2009), *A Profile of Older Americans*, Washington, DC.

U.S. DOT Bureau of Transportation Statistics (2006), *Americans on the Go - Findings from the National Household Travel Survey*, Washington, DC.

Wolfe, Harry P. (2003), *Accommodating Aging Population Needs in Airport Terminals*, (www.elderairtravel.com).

# ANALYSIS OF CRITERIA FOR CLOSELY-SPACED PARALLEL RUNWAY APPROACHES: INFLUENCES AND CAPACITY CONSTRAINTS IN A MULTIPLE AIRPORT SYSTEM

Rafael Fraga MSc. - ITA, Brazil.
Carlos Müller PhD. - ITA, Brazil.
Cláudio Jorge Pinto Alves PhD. - ITA, Brazil.

## Abstract

In order to increase arrival rates at many of today's heavily congested airports, simultaneous approaches, whenever possible, are conducted on parallel runways. For many years, the FAA has permitted simultaneous independent instrument approach operations only on parallel runways with a minimum separation of 4300 ft. Nowadays, new procedures have been developed to increase the utilization on closely-spaced runways, PRM/SOIA (Precision Runway Monitor/Simultaneous Offset Instrument Approach), one of the most recent of such procedures, allows simultaneous approaches in systems of runways spaced as close as 750 ft. Fraga et al. (2010) [6] using computer simulation showed the potential application of PRM/SOIA at Sao Paulo/Guarulhos International Airport (GRU) evaluating its impact on the airport runway hourly capacity. The present paper considering GRU as part of the airport system of TMA-SP (São Paulo Terminal Airspace) focuses on the influence of PRM/SOIA implementation at GRU on the system of main airports in the terminal airspace. Evidences were found that TMA-SP current overlapping approach procedures generate interferences on the operations of the airports in the system reducing the overall runway hourly capacity of the airports. In order to quantify this interference simulation was used to evaluate and to compare the airport system runway hourly capacity as opposed to the individual airport runway hourly capacity added up. It was found that TMA-SP airport system runway hourly capacity is about 30% less than the estimated overall runway hourly capacity of the airports alone.

## 1. Introduction

The capacity of any airspace system worldwide is primarily constrained by the landing capacity of its busiest airport. Historically, world air traffic has been increasing at about 5 to 6% per year for many years and will likely continue at that growth rate in the future. Therefore, increases in the number of aircraft landings will approximately triple in 20 years. In order to accommodate these landing rates, it will be necessary to increase the number of airports, increase the usable number of runways at existing airports, or both (Powell 2005 [19] ). As a result, aiming the enhancement of airport's airside capacity, many airports have implemented multiple runways, prioritizing the construction of parallel runways to allow simultaneous use during periods of high demand. Currently, at many heavily congested airports, simultaneous approaches are conducted to parallel runways in order to increase arrival rates, even under instrument flight rules (FAA 2003 [4] ).

There are some major safety concerns during simultaneous parallel approaches, primarily under instrument procedures. One of the main concerns regards the propagation of wake turbulence from the leading aircraft into the path of the trailing aircraft. Aircraft proximity is another main concern.

As a result, during visual simultaneous approaches (that requires visual meteorological conditions - VMC), the trailing aircraft needs to fly not too far behind the leading aircraft. Wind strength and direction playing a key role in the determination of the maximum distance it can fly behind (as seen in Figure 1). Aircraft proximity concern, as a sudden change of direction of one of the aircraft can lead to in flight collision, establishes the minimum distance between the trailing and leading aircraft. Therefore, a Safe Zone between simultaneous approaches can be created (as seen in Figure 1) and the trailing aircraft can be required to fly inside the Safe Zone at all times (Powell 2005 [19] ; Matsuyama and Cugnasca 2008 [16] ). However, this is almost impracticable during instrument procedures (IFR), once the airplanes wouldn't be able to maintain visual contact to each other.

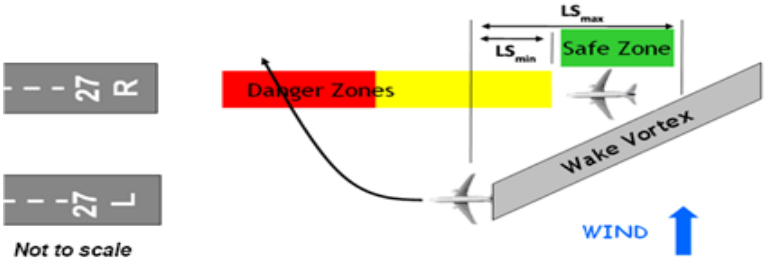



Figure 1. Scheme of zones, longitudinal spacing (LS) and wake vortex [19] .

As a result, for many years, the FAA has permitted simultaneous independent approach operations under instruments procedures for parallel runways with a minimum separation of 4300 ft (FAA 2003 [4] ). These operations can be regularly performed by two aircraft simultaneously flying the standard Instrument Landing System (ILS) procedure using standard displays (Massimini 2006 [15] ).

At some airports, however, limited physical space or even surrounding constructions and terrain topography that may impose violations to the obstacle limitation surfaces may make it impossible to construct parallel runways with the minimum prescribed distance of 4300 ft between center-lines. Hence, some airports that have parallel runways with reduced separations (from 4300 ft to as close as 750 ft apart) may only allow simultaneous approaches under Visual Flight Rules (VFR) and not under Instrument Flight Rules (IFR) when relying in the ILS guidance for landing.

In the 90's, air traffic increase pushed FAA to commission Raytheon to developed a special approach procedure called Precision Runway Monitor (PRM) to allow simultaneous instrument approaches to airports with parallel runways as close as 3000 ft. This procedure relies in the concept of a Safe Zone to be flown by the trailing aircraft and in the use of a high update precision runway monitor radar system with a 1-second update rate (much smaller than the regular radar update rate of 4.8-seconds that was then used). Currently PRM procedure is in use at

Minneapolis, St. Louis, New York, Philadelphia, San Francisco, Atlanta, Sydney, Australia and Chek Lap Kok in Hong Kong.

Pushing even more, in the beginning of 2000's, a special procedure was developed to allow simultaneous instrument approaches to parallel runways as close as 750 ft (FAA 2003 [4] ). This procedure called Precision Runway Monitor/Simultaneous Offset Instrument Approach (PRM/SOIA) is currently in use at the San Francisco International Airport and is said to increase the runway capacity under IFR conditions by as much as 25% with respect to the traditional IFR approach procedure used at the airport (FLYSFO 2003 [5] ).

Therefore, nowadays procedures aiming to improve runway capacity under instrument approaches at any airport with parallel runways as close as 750 ft can be potentially developed.

## 2. Motivation

São Paulo/Guarulhos International Airport (GRU) was originally designed to attend medium and long haul domestic flights, as well as international flights bounded to or from the southern region of South America, typically operated with small to medium size aircraft. Over the years ad hoc arrangements have been made to cope with unplanned demands, mainly regarding the intercontinental flights now operating at the airport with much larger than planned aircraft. As a result, nowadays GRU is Brazil's largest gateway, processing about 70% of all the international passengers bounded to or from the country (Müller and Santana 2008 [17] ), being one of the busiest airports in South America and registering a total of 21.727.649 passengers and 209.636 aircraft movements in 2009 (INFRAERO 2010 [11] ). Hence, GRU has already surpassed its stated terminal capacity of 20.5 million passengers per year (McKinsey & Company, 2010 [14]) and is closely reaching runway capacity.

GRU capacity of expansion to cope with this demand increase is limited. Runway expansion at GRU is very unlikely due to the close proximity of the surrounding communities and the associated environmental challenges. Airport administrators have already stated that plans for a third runway, although present in the airport Master Plan, have been dropped. Terminal and apron expansion is however still considered and very likely in the near future. As a result, congestions are expected in the runway system with the increasing demand at the airport. In fact congestion problems have already been detected at the airport (Pereira et al. 2001 [18] ; Hupalo 2003 [8] ; INFRAERO 2007 [10] and Müller and Santana 2008 [17] ). Furthermore, a study commissioned by the Brazilian National Bank of Development (BNDES) to McKinsey&Company in 2009 [14] to evaluate the Brazilian air transportation industry and just made public in June 2010 concluded that there is a large necessity of investments to increase the Brazilian airports infrastructure in order to supply air transportation services to the estimated demand for the next twenty years. The São Paulo terminal airspace was, of course, of major concern and the study pointed out the urging necessity to increase GRU's runway hourly capacity.

This all ended up motivating Fraga et al. (2010) [6] to study the implementation of the PRM/SOIA procedures at GRU airport in order to evaluate how it could help increase GRU's runway system capacity. Even more so because, except for the Precision Runway Monitor Radar, the ground and onboard technology it requires is already of common use in the current approach procedures (Janic 2008 [13] ).

## 3. São Paulo/Guarulhos International Airport (GRU)

GRU airsided is made of two parallel runways (09L/27R and 09R/27L) with dimensions very similar to the SFO runways. Runway 09L/27R is 12,139 ft long and runway 09R/27L is 9,842 ft long. The runways are closely spaced with runway center-line separation of 1,230 ft, a little larger than SFO runway center-line separation that is only 750 ft. Runway 09L/27R is mainly used for departure operations, and runway 09R/27L is almost entirely dedicated to landing operations (IAC 1999 [9] ; Santana 2002 [20] ). A recent study by Barragán and Dias (2007 [2] ) has shown utilization rates of about 90% for the 09 thresholds in the peak months.

The airport original Master Plan considered a short third parallel runway positioned to the north of the existing runways. Recent reevaluation of this third runway considering current site availability and surrounding urban occupation ended up concluding that it could be some 5,905 to 6,643 ft long with a center-line separation to the closest existing runway (09L/27R) of about 4,796 ft (enough to allow simultaneous instrument operations). Despite the small size and the restrictions it would impose to its operation (mainly dedicated to landings) a study by Santana (2002 [20] ) has clearly measured the impact of this third runway in reducing delays and hence increasing the airport runway capacity.

But, probably due to the limited physical space and to the violations to the aeronautical planning because of the close to the airport urban constructions developed in recent years, the airport and government authorities have already stated that there will not be a third runway in GRU. Therefore, as already mentioned in the last section, the airport complex now operates at the limit of its runway capacity. Therefore, only at the apron and at the airside it is expected that actions will be undertaken in order to expand capacity.

So the question is: What is GRU's actual runway system capacity? The question is even more important when considering that an airport's capacity can be defined as the highest number of aircraft movements a particular aerodrome can accomodate, for specific periods, supported by the airport's infrastructure (Siewerdt 2001 [21] ). According to such definition, decisions taken in order to provide capacity improvements in the infrastructure must be balanced to the entire system (Santana 2002 [20] ), that is, to the landside (e.g. terminals and access routes) and to the airside (e.g. aprons, taxiways and runways).

To answer this question runway capacity must be defined. There are two classic definitions of runway capacity. One definition is for the maximum capacity that reflects the maximum number of operations that the airside can accommodate during

a certain period of time considering continuous demand. It is well accepted that this capacity can only be achieved for short intervals of time, because these values in the long run tend to lead to congestion and to paralyze the operation (Alves C. J. P., unpublished internal report, 2007). Runway practical capacity is the other regular capacity definition and is associated with the number of operations that can be accommodated over a given period of time and that corresponds to a tolerable level of delay. FAA (Advisory Circular 150/5060-5) gives guidance to the evaluation of the practical capacity of an airport (FAA 1983 [3] ). According to this analytical method, considering the mix of aircraft types that regularly operates at GRU, the actual runway system offers a practical capacity in the range of 48 to 72 movements per hour, considering arrivals rate in the range of 60 to 40%, respectively (the capacity increases with the decrease of the arrival rate, and vice-versa). Despite the fact that this evaluation takes into account a series of issues (e.g. runways design and type of operation, types of aircraft in operation, number of taxiways, and so on), the estimated capacity refers to normal conditions only, disregarding issues such as congestion in the local air traffic and meteorological conditions. As a result, the airport can be limited to a smaller capacity than expected by the analytical method due to higher congestion and delay levels. Therefore, simulation plays an important role in capacity studies where non-observed issues in the FAA analytical method are of relevance. Figure 2 shows the observed operations at GRU on January 4th, 2009. It can be seen that even at the moments of peak traffic it always remained below the estimated capacity using the FAA method. Even so, delays were reported in the real operations of that day.

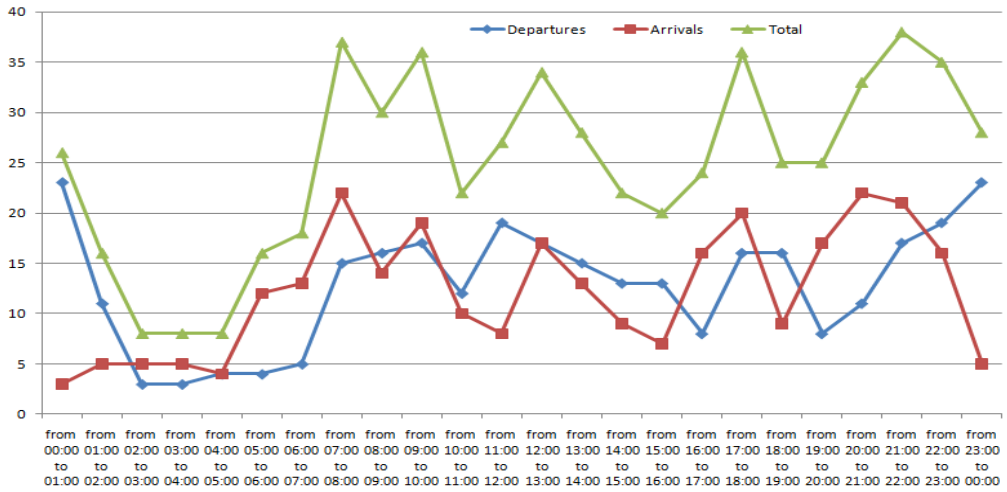



Figure 2. Movements on January 4th, 2009: the peak day in January [12] .

A particular aspect of the operation at GRU that is also not considered in the FAA analytical method is that GRU is inserted in a busy terminal area, TMA-SP, and traffic bounded to or from GRU inches its way through traffic associated with Sao Paulo/Congonhas Airport (CGH) and Campinas/Viracopos Airport (VCP) also at TMA-SP suffering interference and in flight delays from them. In order to study the impact of the IFR traffic of these three airports at their runway capacity a simulation of the overall TMA-SP was developed and the results are shown in this paper.

## 4. São Paulo Terminal Area (TMA-SP)

A Terminal Airspace is the control area delegated by the Air Route Traffic Control Centers (ARTCCs) to the Terminal Radar Approach Control (TRACON) for the provision of approach and departure sequencing, typically within the range of a "fast-sweep" radar sensor (Andre et al. 2003 [1] ). TMA-SP is such a control area and is shaped by two circles, one centered on CGH with a radius of 42 NM, and the other centered on VCP with a radius of 27 NM; extending vertically from 5,500 ft to approximately 19,500 ft. There are eight airports under this area, three of them are major airports having instrument procedures: CGH, VCP and GRU. These are the airports considered in this study.

## 5. São Paulo/Congonhas (CGH) and Campinas/Viracopos (VCP)

CGH is a downtown airport, totally surrounded by Sao Paulo urban area. It is a very busy domestic airport with two parallel runways (17R/35L: 6,364 ft long and 17L/35R: 4,708 ft long). Runway 17L/35R, the shortest, is mainly dedicated to general aviation (GA) operations. CGH demand is mainly related to business passengers and shows a very unique pattern, that is almost constant during the entire day (CGH operates from 6am to 11pm). Authorities have slotted CGH with respect to commercial airplane operations to a maximum of 33 movements per hour, with the possibility for opportunity slots for GA operations in accordance with air traffic controller judgment. Therefore, even with the slot restriction, it is reported that the hourly traffic reaches values higher than 33 movements (according to [12] ). VCP is one of the most important cargo airports of the country. It has only one runway (15/33) that is 10,629 ft long. In the 90's, with the modernization of cargo handling and customs clearance facilities, the airport began to operate intensively with international air cargo and became a reference of the segment in the country (INFRAERO 2009 [11] ). Even so, current demand is still low for this runway: according to [12] , the hourly movements in the peak day on July 2009 (one of the most recently congested periods) registered only 20 aircraft operations. However, VCP Master Plan is very ambitious, it plans for three parallel runways that could make this airport the biggest within TMA-SP in the near future.

## 6. Methodology

The simulation methodology used to generate the results to this study is detailed in Fraga et al. (2010) [6] . All the information regarding the aeronautical documents (Standard Terminal Approach Routes – STAR, Standard Instrument Departure Routes – SID, and the Aerodrome Description Chart – ADC) for the three airports considered in TMA-SP, as well as the flight information data-bank and the steps of the modeling process (verification, validation, assumptions and limitations of the scenarios) are presented and described. A reasoning for the selection and use of RAMS Plus simulation software is also included.

## 7. Results

Fraga et al. (2010) [6] considering current demand levels showed that the implementation of PRM/SOIA approaches at GRU can bring significant improvements, such as 45 to 51% airborne delay reductions at TMA-SP flights. It also showed that there is already an intense relationship between the traffic of GRU, CGH and VCP for the current demand level.

This paper now focuses on the evaluation and comparison of the TMA-SP overall capacity against the evaluation of the independent capacity of each of the three airports considered. The maximum arrival hourly capacity of the TMA-SP system is explored for two scenarios: one representing the current procedures (assumed as baseline) and another representing the availability of PRM/SOIA procedures at GRU, both using stressed demand. The stressed demand was created maintaining the same mix of aircraft observed at each airport during the typical day as in [6] (03/06/2008). It includes more than 100 arrivals per hour per airport. That is more than twice of the capacity calculated by the FAA method ([3] ) for the highest arrival rate percentile of the method (60%), as in Table 1.

Table 1. Runway system IFR hourly capacity (HC) calculated by FAA method ([3] ).

| Airport | HC for 40% of arrivals | HC for 50% of arrivals | HC for 60% of arrivals |
|---|---|---|---|
| **GRU** | 71.78 | 58.20 | 48.50 |
| **CGH** | 48.06 | 46.44 | 43.68 |
| **VCP** | 48.06 | 46.44 | 43.68 |

Table 2 shows the maximum hourly arrival movements the simulation registered at the three airports together when considering just arrivals. Arrivals per airport were not presented since they can vary in accordance with the arrival priority the air traffic control gives to each airport, which in fact varies throughout the day in order to optimize the airborne delay. Therefore, the results shown in Table 2 represent capacity for the entire system, the summation of the hourly arrival movements at each airport considering the highest value as the maximum capacity of the TMA-SP system.

Table 2. Peak arrivals per hour considering the overall system (entire TMA-SP).

| Scenarios | Entire TMA-SP arrivals/hour |
|---|---|
| Baseline | 54 |
| PRM/SOIA | 63 |

Table 2 shows an increase in the arrivals of approximately 18% with the implementation of the PRM/SOIA procedures at GRU.

The capacity of each airport alone was also evaluated by means of the same simulation process not taking into account traffic from the other two airports. These isolated airport simulations considered stressed demand and only arrival traffic. GRU only simulation resulted a peak of 28 arrivals per hour at GRU, CGH only simulation resulted a peak of 26 arrivals per hour at CGH, and VCP only simulation resulted a peak of 24 arrivals per hour at VCP.

For validation matters, these results were compared to the ones generated by Gualda and Moser (2007) [7] and Barragán (2009) [2].

Gualda and Moser (2007) [7] using the FAA Airfield Capacity Model have evaluated theoretical capacities for GRU runway system for a range of in flight separation requirements. They considered that the set of separation requirements that better represents the procedures actually used at GRU was (APP-4NM and DEP-90s; see in Figure 3). They evaluated GRU runway capacity at about 28 movements per hour (considering 100% of arrivals), a result very close to the one we obtained for the isolated hourly capacity at GRU.

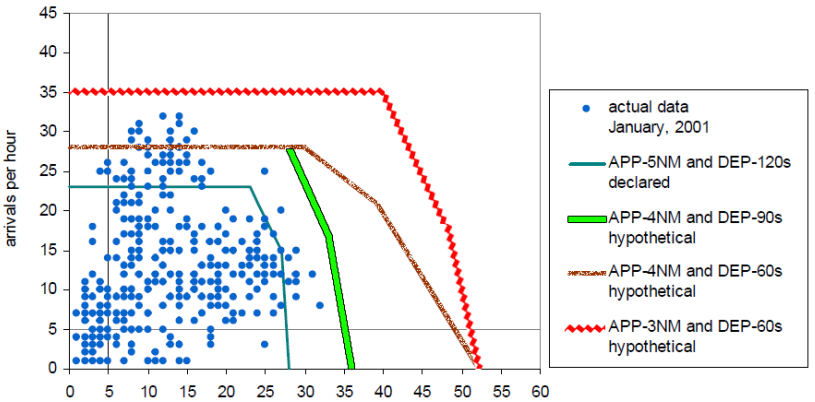



Figure 3. GRU runway system capacity in ACM model for different separations [7] .

Barragán (2009) [2] simulated both GRU and CGH using isolated approaches with ARENA and REDIN simulation software. He obtained an hourly capacity of 29 arrivals for GRU and 26 arrivals for CGH (also using 4NM separation for approaches).

Hence, considering these studies there is a good confidence on our evaluation of isolated airport capacity with RAMS Plus simulation.

An estimated system capacity obtained by the summation of the isolated airport capacities would indicate a TMA-SP arrival capacity of about 78 arrivals per hour (using current approach procedures). The capacity estimated by means of the simulation of the entire TMA-SP and with the in-flight interferences it imposes show a smaller capacity, about only 54 arrivals per hour (also using current approach procedures) in accordance with the scenario considered (Table 2). That means an evaluated hourly capacity difference for the entire TMA-SP about 30 %.

The results show the intense relationship between the airports at TMA-SP, pointing out the importance of considering airport runway capacity in a situation of airports under the same terminal area as an airport system capacity, rather than the capacity of isolated airports since the overall capacity is limited by the interferences that may prevail between the operations of these airports.

## 8. Conclusions

Despite the benefits of the implementation of the PRM/SOIA at GRU showed in Fraga et al. (2010) [6] , that could generates an estimated 18% increase in the hourly arrivals capacity for the entire TMA-SP; this work found evidences that the current approach procedures generates interferences between the traffic bounded to the airports. It is considered that this interference is caused by the overlapping of the approach routes inside the terminal airspace. This interference was evaluated and it possibly results in a reduction in the overall hourly capacity of the airport system of 30% (when considering the current approach procedures).

Finally, it can be said, from the evidences already captured, that even with the eventual construction of a new forth airport in the Sao Paulo terminal area (TMA-SP) the overall hourly system capacity will be very much constrained by in flight interferences.

This strong relationship between the airports in TMA-SP makes it clear that there is the necessity of a new design for the terminal routes (i.e. avoiding overlapping procedures for the different airports); it also points out to the importance of considering airport runway capacity as an airport system capacity issue for airports under the same terminal area.

**References**

[1] Andre A. D., Capozzi B., Krozel J., and Smith P. (2003) “The future national airspace system: design requirements imposed by weather constraints.” AIAA Guidance, Navigation, and Control Conf., Austin, TX.

[2] Barragán, G. A., and Dias, D. P. (2007) “Análise da variação do tempo de ocupação de pista para operações de pouso no Aeroporto Internacional de São Paulo/Guarulhos.” M.S. thesis, ITA, S. J. dos Campos, Brazil (in portuguese).

[3] FAA (1983) “Airport capacity and delay.” AC 150/5060-5. Washington. v. 1995.

[4] FAA (2003) “ILS/PRM Approach for Air Carriers Video.” <http://www.faa.gov/training_testing/training/prm/> (July 3, 2009).

[5] FLYSFO (2003) “SFO and United Airlines Successfully Test New Landing System.” <http://www.flysfo.com/web/export/sites/default/download/about/news/pressrel/pdf/SF-03-64.pdf> (Nov. 1, 2008).

[6] Fraga. R., Müller, C., and Alves, C.J.P. (2010) “Analysis of criteria for closely-spaced parallel runway approaches: the potential application of PRM/SOIA at São Paulo International Airport.” 12th WCTR , Lisbon, Portugal.

[7] Gualda, N. D. F., and Moser, R. F. (2007) “Simulation and analysis of airport configurations using visual simmod: application at São Paulo/Guarulhos International Airport.” USP, São Paulo, Brazil.

[8] Hupalo, M. F. (2003) “Análise da área de controle terminal São Paulo por meio de simulação.” M.S. thesis, ITA, S. J. dos Campos, Brazil (in portuguese).

[9] IAC (1999) “407/DPT-4/99.” DCA, Rio de Janeiro, Brazil (in portuguese).

[10] INFRAERO (2007) “Movimento operacional acumulado da rede” <http://www.infraero.gov.br/upload/arquivos/movi/mov.operac._1207_revisado.pdf> (April 8, 2008) (in portuguese).
[11] INFRAERO (2010) “Aeroportos Brasileiros.” <http://www.infraero.gov.br/aero.php> (October 5, 2010) (in portuguese).
[12] INFRAERO (2009) SGTC databank - Control Tower Management System.
[13] Janic, M. (2008) “Modelling the capacity of closely-spaced parallel runways using innovative approach procedures.” Transportation Research, Part C 16, 704–730.
[14] McKinsey & Company (2010) “Estudo do Setor de Transporte Aéreo do Brasil: Relatório Consolidado.” Rio de Janeiro, Brazil (in portuguese).
[15] Massimini, S. V. (2006) “Simultaneous Independent and Dependent Parallel Instrument Approaches: Assumptions, Analysis, and Rationale.” MITRE, Center for Advanced Aviation System Development.
[16] Matsuyama R. T., and Cugnasca P. S. (2008) “Análise comparativa de técnicas de pouso em pistas paralelas e suas aplicações para o cenário brasileiro.” XII Sitraer, 370-381, São Paulo, Brazil (in portuguese).
[17] Müller, C., and Santana, E. S. M. (2008) “Analysis of flight-operating costs and delays: The São Paulo terminal maneuvering area.” Journal of Air Transport Management 14, 293-296.
[18] Pereira, B.D., Silva, C.M., Guedes, E.P., and Feitosa, M.V.M. (2001) “Estudo dos atrasos na area terminal São Paulo: Situação atual e futura via modelo de simulação.” XV ANPET vol. 3, Campinas, Brazil (in portuguese).
[19] Powell, J. D. (2005) “Use of ADS-B and perspective displays to enhance airport capacity.” 24th Digital Avionics Systems Conference.
[20] Santana, E. S. M. (2002) “Análise de Novos Cenários Operaionais para o Aeroporto Internacional de São Paulo/Guarulhos.” M.S. thesis, ITA, S. J. dos Campos, Brazil (in portuguese).
[21] Siewerdt, E. (2001) “Gestão de Aeroportos e Espaços Aéreos no Limite de suas Capacidades“, II SITRAER, ITA, S. J. dos Campos, Brazil (in portuguese).

# Development of a computer program for Airport runway location, orientation, and length design in Iraq

**Saad Issa Sarsam***
**Hiba Akram Ateia****
*** Assistant Professor in Transportation Engineering,**
**Email: saadisasarsam3@hotmail.com**
****M.Sc. Student**
**Civil Engineering Department, Baghdad University, Baghdad - IRAQ**

**Abstract**
The orientation of Airport runway requires information on wind speed, wind direction, and wind speed intensity and duration. On the other hand, the runway location design requires information on the latitude and longitude of the site, temperature variation at the site throughout 20 years period, and the altitude of the airport site, so that the x and y coordinates could be corrected.
The design of runway requires information on the expected traffic volume, type of aircraft intending to use it, and the class of the airport as per the ICAO specifications.
In this work, a computer program has been developed, with data base constructed inside, such data regarding the wind and temperature and the traffic data as explained above was fed to the data base. The latitude, longitude, and altitude of each of the eighteen provinces in Iraq were also fed to the system.
A methodology for drawing the wind rose diagram and the conditions for selecting the best runway orientation through optimization process has been included in the data base.
The required calculation procedures for basic runway length and the corrections usually adopted for variation of temperature and altitude, landing and takeoff situations were also fed to the data base in the system.
An input sheet was designed through visual basic environment so that if an airport is required anywhere in Iraq, the system will ask for the basic data of latitude and longitude and will suggest the best orientation of the runway using the optimization technique.
After suggesting the orientation, another input sheet will ask about the airport class required, the type of air craft, and the traffic volume and modes of takeoff and landing, and then the system will suggest the basic runway length required, and the required corrections to be considered for the environmental conditions.
In order to check the validation of the software, Data of selected existing airports in Iraq were obtained and fed to the software so that a comparison of its output with the existing condition could be obtained.
It was felt that such program will help in the management of airport design, and provide instant information regarding the runway orientation, location, and design.

**Kew words:** Length design; Location; Orientation; Runway; Wind rose.

## 1.0 Introduction

The requirements of the airport runway system are analyzed for its ability to meet the needs of users throughout the planning period. The main objective of this effort is to provide software for runway system that meets FAA and ICAO standards, and provides for a safe and efficient airfield. For the operational safety and efficiency of an airport, it is desirable for the primary runway to be oriented as close as possible to the direction of the prevailing wind. This reduces the impact of crosswind components during landing or takeoff (ICAO).

The recommended length for a primary runway at an airport is determined by considering either the family of airplanes having similar performance characteristics, or a specific aircraft requiring the longest runway. Additional factors considered include critical aircraft approach speed, its maximum certificated takeoff weight, useful load and length of haul, the airport's field elevation above sea level, the mean daily maximum temperature at the airfield, and typical runway surface conditions, such as wet and slippery ( FAA, 2000).

## 2.0 Background

The orientation of the runway is an important consideration in airport planning and design. The goal is to define the runway orientation that maximizes the possible use of the runway throughout the year accounting for a wide variety of wind conditions

FAA and ICAO regulations establish rules about runway orientation and their expected coverage

Ideally, all aircraft operations on a runway should be conducted against the wind, since wind conditions vary from hour to hour, a careful examination of prevailing wind conditions at the airport site is required.

**Jia et al (2004)** presented the GIS- based wind rose method to determine the optimal orientation of a runway. The method uses a set of customized GIS functions, and the data base management tools to solve both of the partial coverage problems and for optimization of runway orientation.

**Mosa and Mumayiz (2000)** presents a computer model for optimizing the runway orientation based on a given wind data and allowable crosswind components. The model is based on a mathematical formulation which transfers circles and radial lines of the wind rose method into points with numeric coordinates. The development of the model is presented and discussed in details in their paper. In determining the wind coverage for a given runway direction, the model calculates exactly the areas of sectors covered, fully or partially, by the runway template and determines a factor for adjusting the wind data of covered sectors. The optimization is achieved through an "exhaustive search" by evaluating the coverage at small increments of orientation angle covering all possible directions.

**Falls and Brown (1972)** in their report for NASA presented two procedures for obtaining the optimum runway orientation relative to minimizing a specific crosswind speed. They stated that either the empirical or the theoretical methods described in their report may be used to determine the runway orientation that minimizes the probability of crosswinds. The empirical procedure requires only hand calculation on an ordinary wind rose, while the theoretical method utilizes wind statistics computed after the bivariate normal elliptical distribution is applied to a data sample of

component winds. It requires the assumption that the wind components are bivariate normally distributed.

**FAA Advisory Circular (2000)** states that an airport's runways should be oriented such that aircraft can take-off and land into the prevailing wind with minimal crosswind exposure. The runway system should provide 95% wind coverage. Thus, the goal is to achieve 95% coverage or better.

### 3.0 Development of the software

This work represents one of the early attempts to develop a system for analyzing runway orientation problems using relational data bases in Iraq, and had expert processing logic as its managed by system which consists of groups of programs, which manage, control, save, and recall data, making available to many users to access and deal with the information. Figure 1 illustrates the starting page of the software.

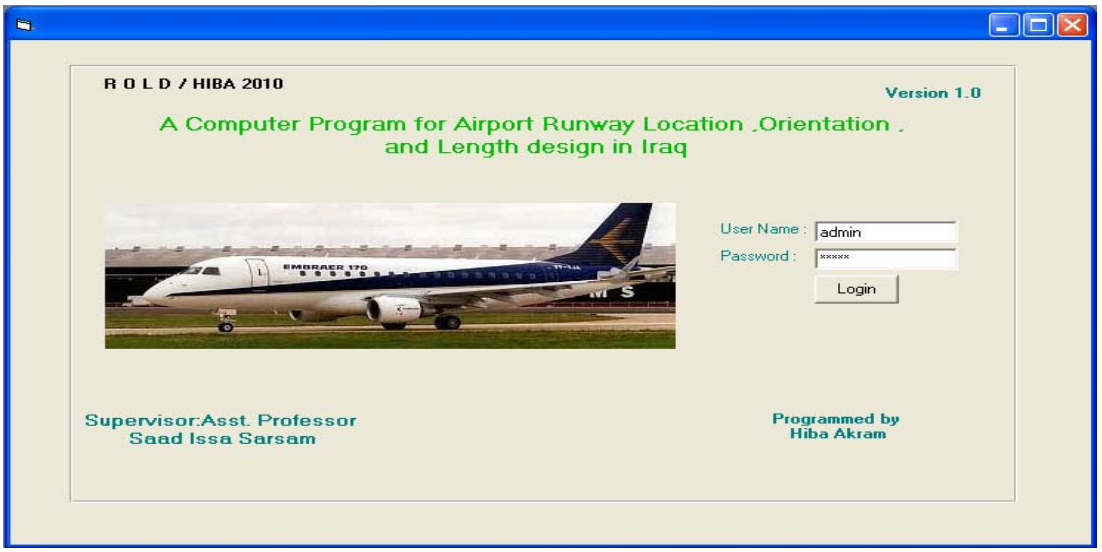



**FIGURE 1: The starting page of the software**

**3.1 Metrological data** Iraqi Metrological organization data are available from 24 stations across the country; these stations have daily records on climatic conditions for the regions on which they exist. Such data covers the past 20 year's period for the whole country.

The wind data which includes wind direction, speed and percentage intensity (duration) have been collected, the air temperature data which includes the daily maximum and minimum air temperature of the hottest months in each year was also obtained and calculated. Such data were fed and stored in the data base of the developed system.

### 3.2 Wind Coverage

Wind coverage is calculated using a wind rose, which graphically depicts wind data. Two systems of the wind rose have been implemented; the first one takes the wind direction and duration into consideration. The wind rose is plotted with its radial lines

representing both wind intensity and direction. Figure (2-a) illustrates the work sheet of the method, while figure (2-b) demonstrates the wind rose diagram. In the second method, the wind rose is essentially a compass rose with graduated concentric circles representing wind speed. Each box in the wind rose represents a compass direction and, when filled, indicates the percentage of time wind travels in that direction at that speed.

The wind rose template has a polar coordinate system that is made of circles and radial lines Circles on the template represent the wind speed, while the radial lines illustrate the angles or the wind blowing directions. Each cell bounded by two circle segments and two radial lines stores the percentage of time that the winds correspond to a given direction and velocity range. The circle with label represents the speed of crosswinds the runway will experience during its operations is less than that of the allowed crosswinds.

A transparent runway template is placed on the wind rose to represent the proposed runway that accommodates the size and operating characteristics of aircraft. The template is rotated around the center of the wind rose in order to search for an optimal runway orientation. At each rotating angle, the total percentage of allowable crosswinds in the wind rose that are covered by the template is calculated, and a best angle that can give the maximum percentage of coverage is determined. Figure (3-a) shows the work sheet of the method, while figure (3-b) illustrates the wind rose diagram. The decision to use each of the systems depends mainly on the type of the input data available.

### 3.3 Design criteria implemented

FAA-2000 and Aerodrome design manual (Volume 1 for ICAO standards) employ the most critical aircraft expected to operate in the airfield (in this context the most critical is the largest, it recommends to provide a runway orientation that satisfies 95% coverage (i.e., crosswinds below a critical value) considering yearly wind conditions. If one runway does not meet the 95% criteria, a second crosswind runway which is perpendicular to the first runway is suggested.

### 3.4 Airport Design Aircraft

Critical aircraft is that aircraft with the most demanding (i.e. largest) critical dimensions and highest approach speed that consistently (at least 500 operations per year) uses the airport (FAA-2005). Each aircraft has a uniquely stated maximum crosswind component (derived from flight test experiment and supplied by the manufacturer). The challenge for the designer is to accommodate all of the aircraft that will use the facility in a reliable and reasonable manner. FAA coding system is used to relate airport design criteria to the operational and physical characteristics of the aircraft projected to use the airport, while ICAO standards considers takeoff weight, airport altitude, and the required takeoff length.

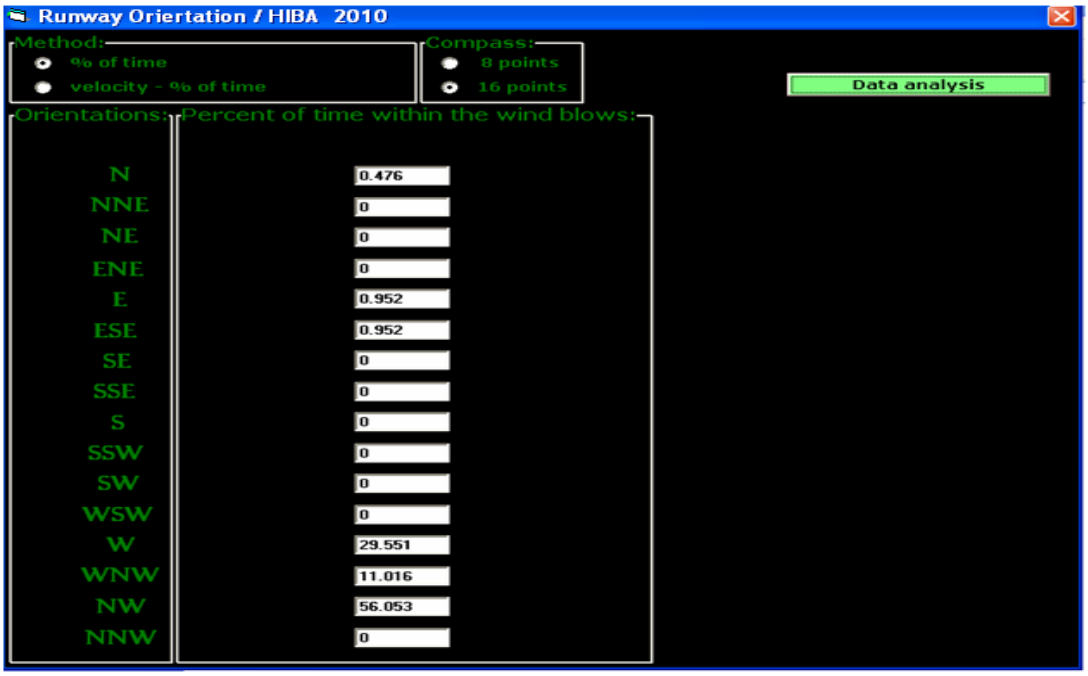



**FIGURE 2-a: Work sheet for runway orientation using wind direction and intensity**

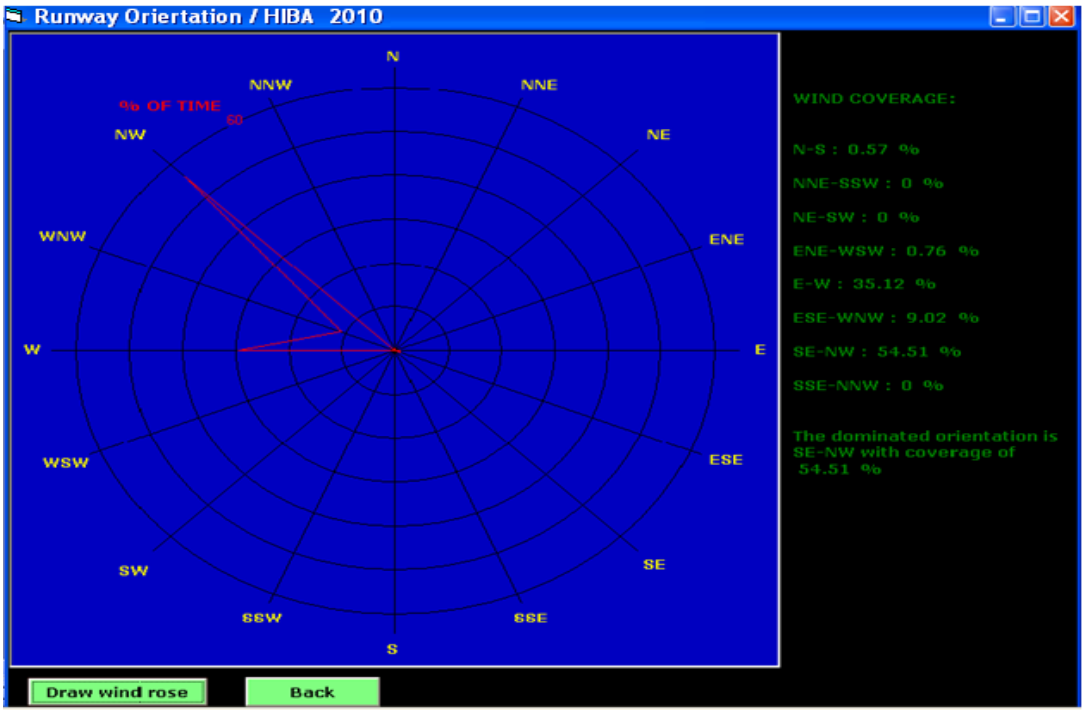



**FIGURE 2-b: Wind rose diagram using wind direction and intensity**

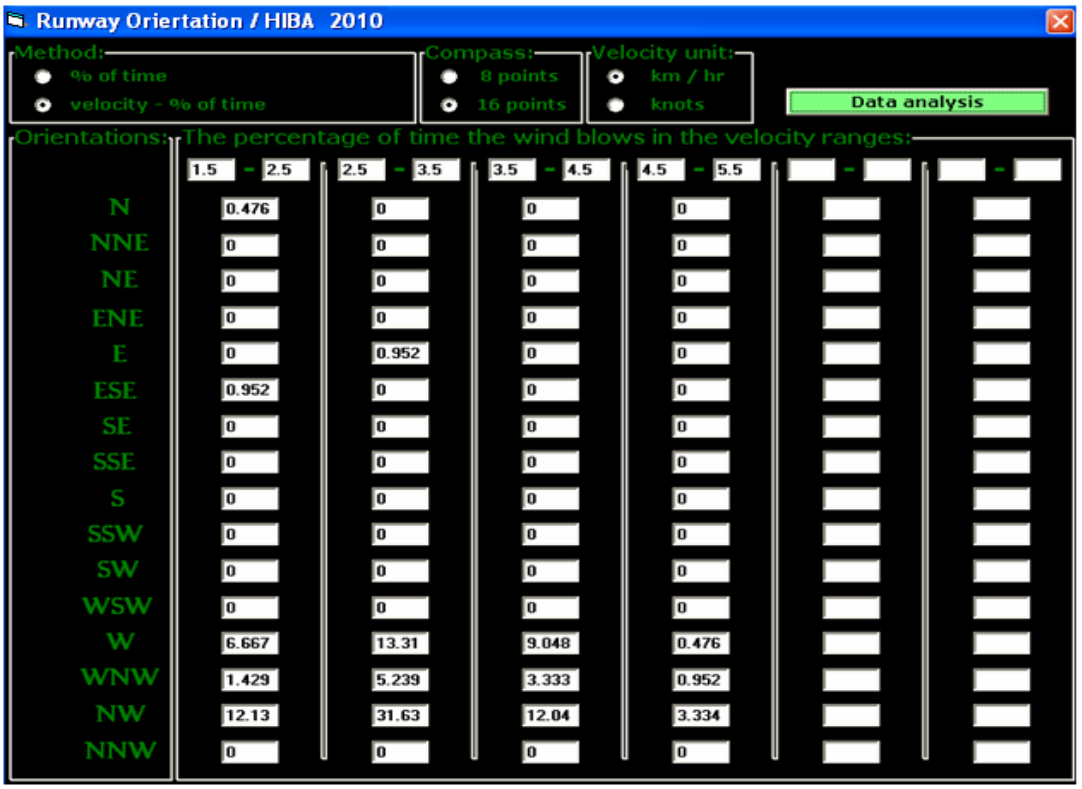



**FIGURE 3-a: Work sheet for runway orientation using wind direction, speed and intensity**

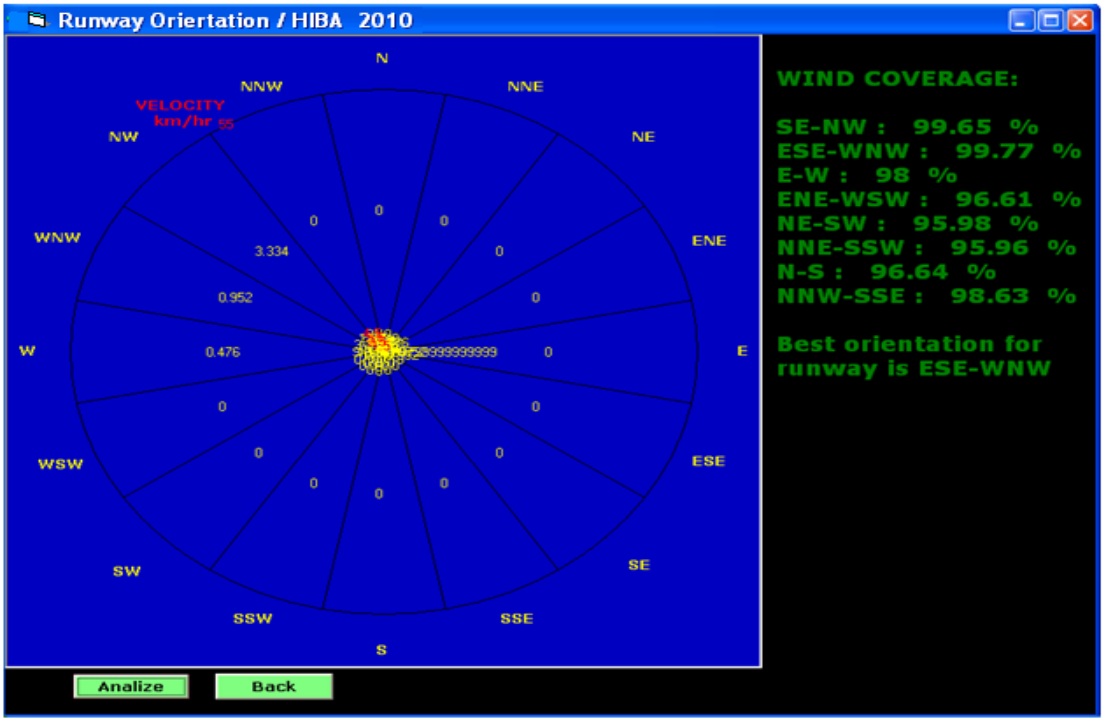



**FIGURE 3-b: Wind rose diagram using wind direction, speed, and intensity**

### 3.5 Runway Length Requirements

The runway length required for an airport is based on standards presented in FAA AC 150/5300-13, and FAA AC 150/5325-4A. Runway length requirements for an airport are based on several factors such as airport elevation, mean maximum temperature of the hottest month, runway gradient, airplane operating weights, runway surface conditions (*i.e.*, wet or dry), and others.

The recommended basic length for a primary runway at an airport is determined by considering either the family of airplanes having similar performance characteristics, or a specific aircraft requiring the longest runway. This need is based on the aircraft or family of aircraft that intended to use the airport on a regular basis, where regular basis is typically defined as a minimum 500 itinerant operations per year (FAA-2005). Additional factors considered include critical aircraft approach speed, its maximum certificated takeoff weight, useful load and length of haul, the airport's field takeoff elevation above sea level, the mean daily maximum temperature at the airfield, and typical runway surface conditions, such as wet and slippery. Both the Advisory Circular (FAA -2005), as well as the FAA's Airport Design and FAA software (FAA-2000) classify aircraft based on weight. The program includes an aircraft fleet profile designed to be representative of the small and large aircraft that comprise the general aviation aircraft fleet in the United States.

ICAO standards take three cases into consideration for estimating the basic runway length, the normal takeoff, landing, and engine failure continued or aborted takeoff. The required information concerning different types of aircrafts behavior in the above mentioned cases are obtained from the web sites of the manufacturers (as for Airbus, Boing etc.). The basic runway length is then corrected for gradient, elevation, and temperature variations. Finally, the longer runway length among the above three cases is taken into consideration as per ICAO. The runway length requirements in this work are defined in accordance with the Aircraft Characteristics for Airport Planning Manuals distributed by the corresponding aircraft manufacturers. These manuals provide consideration for most factors that influence the basic runway length required for aircraft operations. The ICAO standards for runway length are considered in the developed software. Figure 4 demonstrate the sample calculation work sheets for calculating the basic runway length, while figure 5 illustrates the work sheets for correction of basic runway length.

### 3.6 Verification of the software

The software was checked for output accuracy using the data of three existing national and international airports in Iraq (Baghdad, Mosul, and Basra), the runway orientation varied by 3-7 angle degrees only, this could be attributed to that such airports have been constructed during the mid-seventies and eighteens, the runway orientation was based on wind data taken through the fifties and sixties. The metrological conditions, especially those for wind speed and direction, have been changed throughout the past forty years.

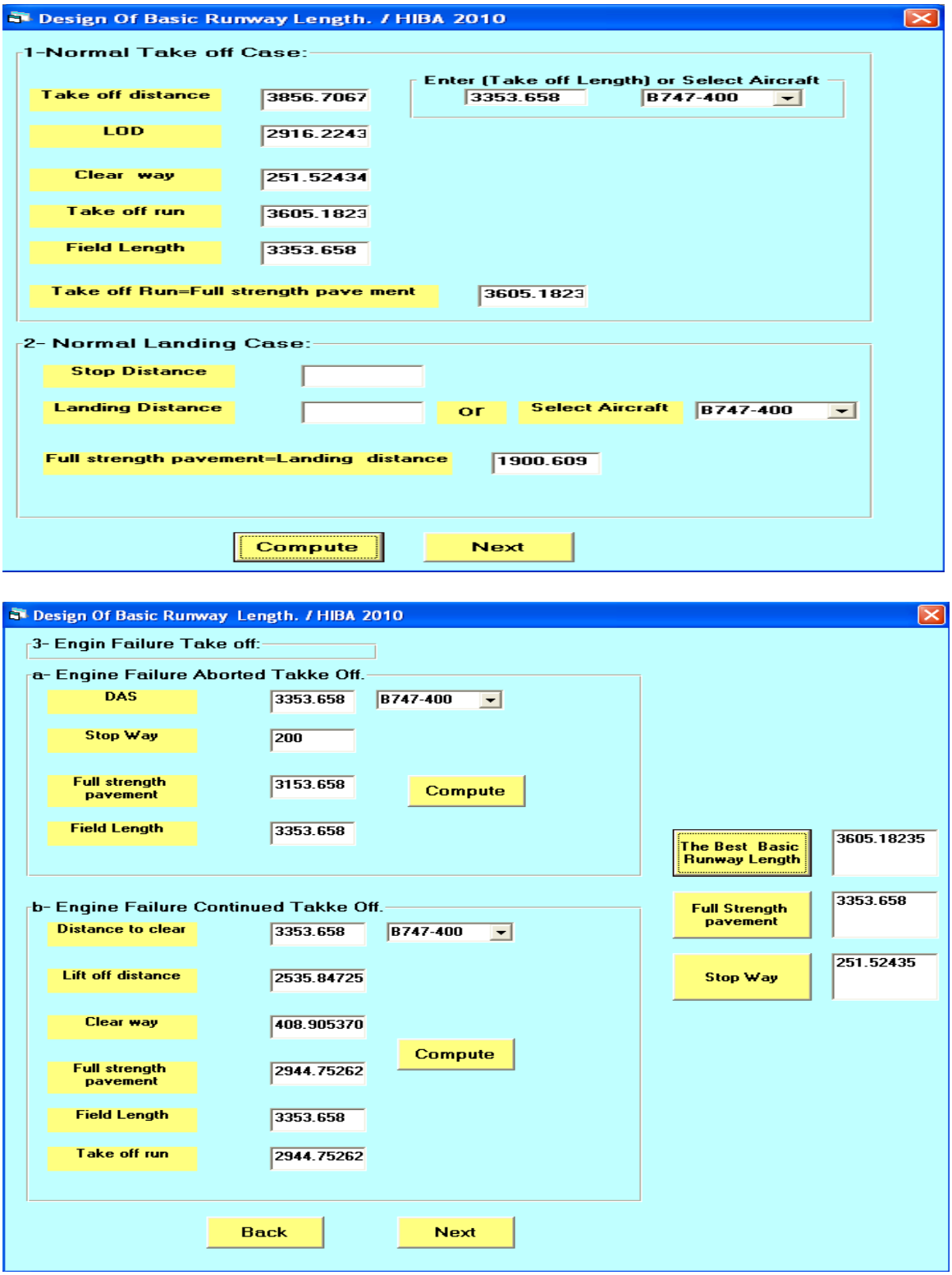



**FIGURE 4: Work sheets for basic runway length calculation**

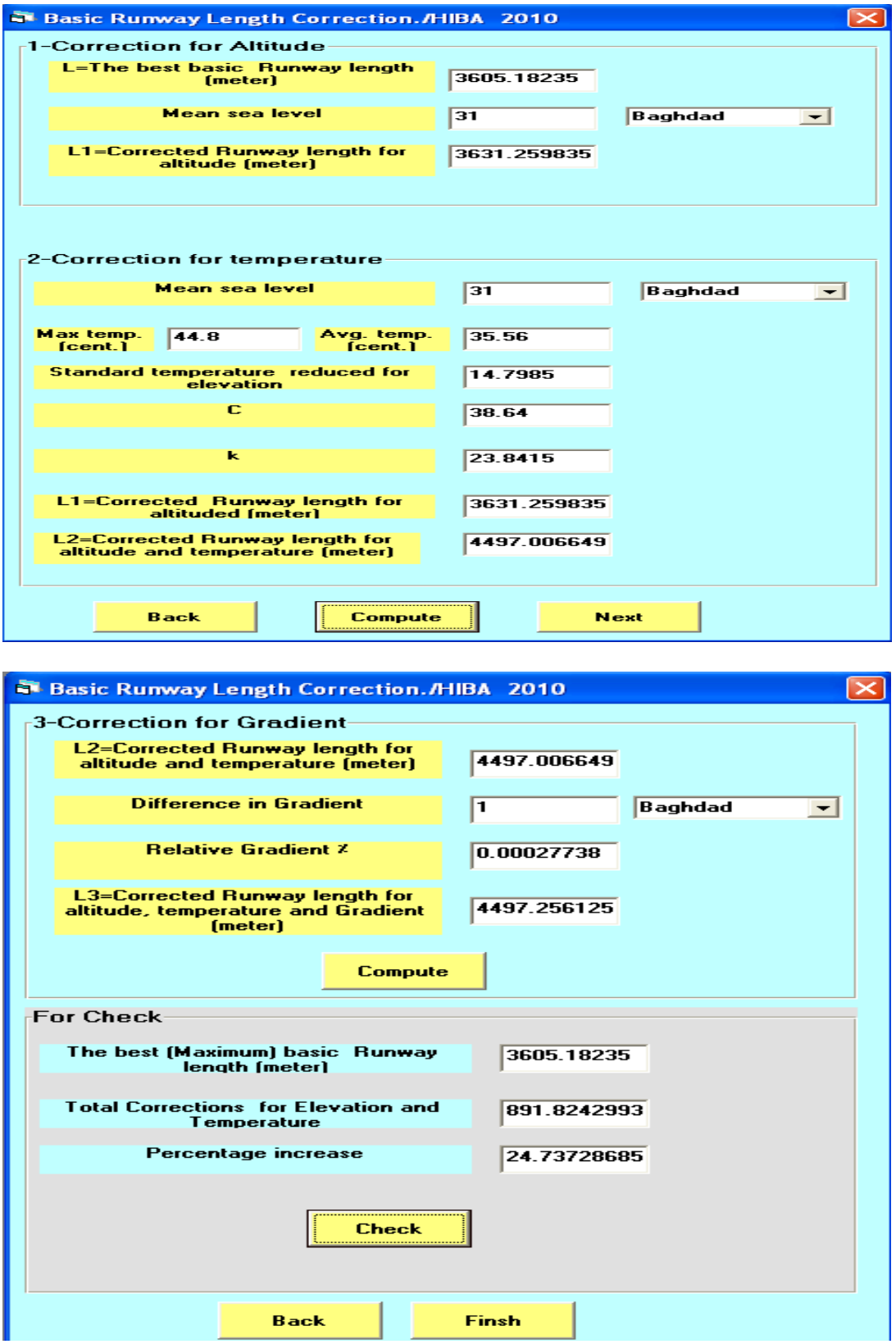



**FIGURE 5: Work sheets for correction of basic runway length**

## 4.0 Conclusions

Based on the developed software, it could be concluded that such program will help in the management of airport design, and provide instant information regarding the runway orientation, location, and length design.

## References

**1-** Mousa R. and Mumayiz S. (2000) "Optimization of Runway Orientation" Journal of Transportation Engineering, Volume 126, Issue 3, pp. 228-236 (May/June 2000).

**2-** Jia X., Chung D., Huang J. , Petrilli M. and The L.(2004) " ARO: Geographic Information Systems-Based System for Optimizing Airport Runway Orientation" Journal of Transportation Engineering Volume 130, Issue 5, (pp. 555-559) (September/October 2004).

**3-** Falls L.W. and Brown S.C.(1972) "Optimum runway orientation relative to crosswind" NASA Technical note TND-6930, National aeronautics and space administration, Washington DC, September- 1972.

**4-** US DoT- FAA(2005) "Runway length requirements for airport design" Advisory circular AC 150/5325-4B, 2005.

**5-** US DoT- FAA(2000) "Airport design" Advisory circular AC 150/5300-13, 2000.

**6-** US Army corps of Engineers (2008) "Airfield and heliports planning and design" Department of defense, Unified facilities criteria UFC3-260-01, 2008.

**7-** International Civil Aviation Organization (ICAO) Annex 14- Aerodrome Design and Operations.

8- Airbus S.A.S., Customer Services, Technical Data Support and Services, A380 Airplane
Characteristics (AC) manual,
http://www.content.airbusworld.com/SITES/Technical_Data/docs/AC/DATA_CONSUL T/AC_A380.pdf

9- Boeing 747 Airplane Characteristics for Airport Planning, Document D6-58326-1 (Revision D, December 2002) http://www.boeing.com/commercial/airports/747.htm3.3

# Land-Based Innovations for Stopping Commercial Aircraft

By

Hugh K. DeLong, III, AAE, PE, M.ASCE

M.ASCE, BS United States Air Force Academy, MS Air Force Institute of Technology, Accredited Airport Executive, Licensed Professional Engineer, Airport Civil Engineering Officer at ESCO, CEO at ESCO Engineering & Services, PC
ESCO – Zodiac Aerospace, Airport Civil Engineering Office, 2239 High Hill Road, Logan, New Jersey 08085, Ph (856)241-8620 x 453, Fax (856) 241-8621, email: hugh.delong@zodiacaerospace.com

## Table of Contents

## Abstract

There have been many improvements and innovations in the specialized technology involved with the emergency stopping of commercial aircraft when they have had an unplanned excursion from the landing or take-off runway. While these overruns and undershoots do not occur at a high frequency, when they occur the results can be disastrous. For more than fifteen years, ESCO-Zodiac Aerospace has been a pioneer in the commercial aircraft stopping business. ESCO's Engineered Material Arresting Systems (EMAS) has led the way with a system that has gone beyond Federal Aviation Administration (FAA) acceptance in 1996 to become a proven life saver. In the more than half dozen times that an aircraft emergency has resulted in an aircraft in an EMAS bed, the system has worked successfully. More than 200 people have been protected from serious injury or death because of the presence of an EMAS in the Runway Safety Area (RSA) at the end of a runway. All aircraft that have entered an EMAS have subsequently flown from the airport in less than two weeks and continued on in revenue service. Because of a continuous improvement perspective on the part of the designer and manufacturer, there has been a continuing stream of innovations for the EMAS process. The purpose of this paper is to provide an overview of the current state of the practice of EMAS design by highlighting the numerous recent innovations in the product, design and construction of EMAS systems. With as much detail as can be provided for a proprietary product, the innovations of the basic EMAS block will first be reviewed. Second, the EMAS design process improvements will be summarized. Finally, the alternatives for construction of an EMAS bed will be revealed. The paper will summarize these past innovations and also offer a quick look into the future of this well-developed safety system, touching on some of the possible future innovations.


## Historic Review of Product Development Innovations

In the mid-1990s, with strong support of the Port Authority of New York and New Jersey, the Federal Aviation Administration's Technical Center and ESCO, a new land-based product for stopping commercial aircraft was created (Bennett, October, 1999). ESCO established the design of the unique core material for an EMAS block and patented the idea (Mahal -341727, Feb 7, 1997) along with its companion patented block test method (Mahal -341731, Feb 7, 1997). The original core material strength indicators of "80" and "60" were later supplemented with a "50" strength level to better adjust EMAS beds to the fleet mix of aircraft that might need to use them, as lighter weight aircraft need lower strength material in order to initiate the crushing action which decelerates the aircraft. Predicting an aircraft's reaction to the presence of an EMAS bed required the development of a new computer model for EMAS (Mahal -341730, Feb 7, 1997).

Protection of the core block from jet blast, weather and the environment moved from a directly applied, thin cementatious layer to individual block packaging using thin cement boards with a scrim fabric and into the current system of a plastic top with embedded scrim (Mahal, Oct 19, 2005, ) and a plastic bottom with forklift tine holes for ease of handling (Mahal, Apr 6, 2007).

The seams between the EMAS blocks evolved from a basic caulk and backer rod combination to an extruded silicone tape being adhered to the block edges. The sides of EMAS beds moved from being sealed to being intentionally vented at each block joint to allow the bed to breathe. The front of the EMAS beds (closest to the runway end) used an evolving form of metal debris deflector to protect the EMAS bed from jet or prop driven particle scour and erosion while also aiding the entry of aircraft wheels into the EMAS bed. As part of the shift to individually finished blocks, the need to accurately position the correct height block in its planned location led to the use of a surveyed grid being painted onto the EMAS support pavement. To support northern tier EMAS installations, an extreme low ground pressure snow blower specification was evolved to permit the safe removal of any snow and ice that interfered with navigational devices or lights.

To support maintenance activities, an equipment access ramp was integrated into the front or back of EMAS beds to facilitate light-weight, wheeled carts for individual block replacement efforts (lower left photo). As more EMAS beds were fielded, the need to be able to test the blocks for their current strength surfaced (lower right photo). ESCO created a customized field strength test method (Mahal, Sep 24, 2009) which was subsequently approved by the FAA (Marinelli July 15, 2009). This method has already been used to verify the condition of several fielded EMAS beds.

To expand the benefits of better block protection from the newer plastic tops, a retrofit package was created to permit the upgrading of earlier beds to the current form of weather and jet blast protection (upper left photo). FAA approval came in the summer of 2009 (Marinelli July 15, 2009). For their research value, several partial retrofits were conducted at a handful of bed sites and the inaugural complete retrofit bed was installed at Binghamton Regional Airport in the fall of 2009 using an FAA grant (upper right photo) and a second retrofit occurred in Summer of 2010.

## Review of Current Design Process Improvements

While EMAS design has always been provided as a sub-consultant to the prime engineer or a joint venture partner, the actual design process has grown from a basic portrayal of EMAS deceleration predictions for aircraft in the 1990s into an early design, decision-making tool for funding agencies. By current FAA policy (Bennett Sep 30, 2005), the Preliminary EMAS Design Report documents the EMAS options available for use at the specific project site. The FAA headquarters Office of Airports Safety and Standards (Lang March 15, 2004) makes the selection of the EMAS option that will be funded for that project, setting the detailed scope of the EMAS support pavement and the limits of the rest of the project. Because the EMAS design process is new and unique to most engineering firms, an expanding EMAS design manual has grown to fill the need of providing a basic overview for prime engineers as they seek to become the firm selected to design and implement the full project scope. Another unique aspect of the actual installation is the fact that ESCO provides Instructor/Inspectors to ensure continuous surveillance inspection of the EMAS bed installation. This process assures the airport owner that the EMAS bed is installed properly. A companion component of the design process is the fact that the EMAS installation specification is focused upon the extensive use of means and methods; defining manpower, materials, equipment, methods and a schedule. This type of specification allows an inexperienced contractor to successfully bid on and install an EMAS bed, keeping the financial impact of the grants localized to the airport area.

Early modeling results for the first EMAS beds led us to the conclusion that different aircraft need different EMAS block heights. The low hanging engines on some

aircraft preclude the use of optimum block heights. Specifically, the early Boeing 737 engines would touch any EMAS block higher than a nominally sized "twenty" block. This observation led to the design process where any bed serving a B-737 would have to remain smaller than 20 inches (50 cm) until the point in the bed is reached where the B-737 would come to a stop after exiting the runway at 70 knots.

Another thing learned from early modeling efforts was that the rate of raising the EMAS block height could have a significant affect on the forces transmitted to the aircraft landing gear. Since the flat EMAS block tops are either level or inclined upward at a rate of one inch (2.5 cm) in four feet (120 cm), the EMAS bed designer needs to make choices about the front ramping of block heights. A ramp using only inclined tops gets a bed to the highest level in the shortest distance, but a "quick" ramp like this might very well cause the onset of resistive forces that would fail a nose landing gear. A predictive performance computer model is used to assess the ramp design. A gentler ramp that alternates between flat and inclined blocks would slow the rate at which decelertive forces are applied to the aircraft, helping to avoid a nose landing gear failure. Calculating the resistive forces that an EMAS bed could produce is important for figuring out deceleration of the aircraft, but knowing the ultimate failure loads of aircraft landing gear requires extensive aircraft manufacturer data to avoid breaking the aircraft.

## Alternatives for Construction of an Aircraft Arresting Bed

While many Runway Safety Areas (RSAs) are large, rectangular and uncluttered where an EMAS bed would be placed, others have Approach Lighting Systems (see photos below), Far Field Monitors (upper left photo on next page), immoveable utility access points, and conflicting geometric shapes. To adjust for these situations, EMAS beds have been built with truncated rear corners, removable individual blocks for utility access, and cut-outs around ALS conduits and Far Field Monitors.

Overseas, many Localizers have frangibly mounted signal sensors “in front” of the Localizer Array and the EMAS installation in China had to adjust for a localizer signal sensor (upper right photo). The pre-painted grids for block placement provided the added value of depicting which parts of a block would need to be field-modified for conduits or towers, allowing a specific block to be altered well ahead of its scheduled placement time. This minimized the delays associated with setting a non-standard EMAS block. To meet the intent of frangibility requirements within an RSA, conduits within an EMAS bed were adjusted to incorporate two frangible couplers, one at the ground level and another at the top of the surrounding blocks. To weather proof a vertical penetration of an EMAS block, non-metallic pipe seals normally found on flat roofs were used with excellent results.

In one installation at Little Rock National Airport in Arkansas, two separate RSAs intersected in the location where one EMAS bed was planned, causing the jet blast deflector to shift into an “L-shaped” design in lieu of the standard straight line (lower left photo).

In another installation at Boston’s Logan International Airport, four-engine heavy aircraft taxiing past the planned EMAS would subject the sides of the bed to unusual jet blast, so the debris deflector was again revised into a “U-shaped” deflector to protect the foremost side blocks. This created the necessity to establish an equipment access ramp at the rear of the bed (lower right photo).

Where EMAS beds were planned to be positioned in a 100-year flood plain, the method of adhering the blocks to the support pavement had to be revised to reduce the chance that a submergence event would severely damage or move the bed (upper left photo – Mississippi River). Since EMAS blocks weigh about 25 pounds per cubic foot (400 kgs/cubic meter) and water weighs 62.4 pounds per cubic foot (1000 kgs/cubic meter), the EMAS blocks float easily and every part of a flooded block must resist buoyant forces of nearly 40 pounds per cubic foot (600 kgs/cubic meter). Given block lengths and widths of four feet (1.2 meters) and nominal block heights ranging from 3 inches (7.5 cm)to 26 inches (65 cm), a large block could have well over 32 cubic feet (0.9 cubic meters) of volume, creating a buoyant force of well over 1000 pounds (45 kilograms) of uplift. The early practice of just drizzling about a quart (0.95 liter) of melted asphalt cement onto the asphalt pavement to set and locate the block was upgraded to require the placement of about 2 gallons (7.5 liters) of melted asphalt cement for every block (upper right photo). Additionally, to avoid the embrittlement that comes with asphalt aging, the asphalt cement was required to be modified with a 25% content of a crack-sealing material to maintain flexibility throughout the life of the bed. In fact, the actual pavement material was changed from asphalt to concrete so that the weight and strength of the concrete pavement could resist the uplift forces being applied by the adhered and buoyant EMAS blocks.

Another pavement design change brought about by the need to support an EMAS bed was the multi-layered Moisture Mitigation Method which involved the use of four standard products for non-standard purposes within an asphalt pavement system. At the base of the system, a geotextile fabric is used to help condense upward-moving water vapor and redirect it away from the EMAS at the bottom side of the geotextile. Next, following the placement of compacted aggregate, an asphalt prime coat was applied at twice the normal rate to create a second barrier layer for water vapor. Then tack coats were applied between asphalt lifts, again at twice the normal application rate, as a third water vapor barrier. Finally, the melted, block-holding asphalt cement was applied at about a half quart (0.9 liter) per square foot (930 square centimeters) with over 90% coverage of the block location, forming the fourth and final Moisture Mitigation Method used in a typical EMAS support pavement.

At the Wilkes-Barre/Scranton Airport in northeast Pennsylvania, the airport wanted to temporarily place a small EMAS in a blast pad area for several years before the planned fill project would permit the placement of a larger, standard EMAS and maybe a short runway extension (upper left photo). As a hilltop airport, the chance of flooding was nil, so a block placement design with the ability to relocate most of the blocks was incorporated into the project (upper right photo).

## Future Innovations Possible for an Aircraft Arresting Bed

While the EMAS product, design and installation are fairly well developed and ESCO embraces a continuous improvement philosophy under multiple ISO certifications, there are still some possible future innovations for EMAS. The product improvements to date have focused on the four areas of improving product durability, refining the speed-predicting computer model, reducing product installed cost and reducing airport maintenance costs. While significant improvements have been made in each of these four areas, ESCO's future focus will adjust to the companion goals of reducing costs while reducing green house gases during production and shipping. These new developments might be accomplished by use of an alternative core material or the possibility of on-site fabrication. The software basis for the computer model might be updated and further research on the block joint sealing materials might also reduce costs in some of the improvement areas cited above. In the future, as the push for aviation safety continues to increase, EMAS beds might even be installed atop wide RSA "bridge" decks or on floatable RSAs. The full impact of future EMAS innovation is exciting but hard to predict!

## References

Bennett, David L., Director of Airport Safety and Standards, AAS-100. *Runway Safety Area Program*. FAA Order 5200.8. October 1, 1999.

Bennett, David L., Director of Airport Safety and Standards, AAS-100. *Engineered Material Arresting Systems (EMAS) for Aircraft Overruns*. FAA Advisory Circular 150/5220-22A. September 30, 2005.

Lang, Catherine M., Deputy Associate Administrator for Airports, AAS-100. *Financial Feasibility and Equivalency of Runway Safety Area Improvements and Engineered Material Arresting Systems.* FAA Order 5200.9. March 15, 2004.

Mahal, P., Patent number 56611-341727, EMAS blocks & bed, Feb 27, 1997.

Mahal, P., Patent number 56611-341731, EMAS test method, Feb 27, 1997.

Mahal, P., Patent number 56611-341730, EMAS computer model, Feb 27, 1997.

Mahal, P., Patent number 56611-337962, Revised (502) blocks, Oct 19, 2005.

Mahal, P., Patent number 56611-382175, Field Strength Test method, Sep 24, 2009

Mahal, P., Patent number 56611-337966, Plastic Bottom trays, Apr 6, 2007.

Marinelli, Rick, Manager, Airport Engineering Division, AAS-100. *Engineered Material Arresting Systems (EMAS) Recent Developments: Field Strength test (FST), Lid Retrofit.* Federal Aviation Administration Memorandum. July 15, 2009.

## Additional Information

1. DOT/FAA/PM-87/27, *Soft Ground Arresting Systems,* Final Report, September 1986 – August 1987, published August 1987 by R. F. Cook, Universal Energy Systems, Inc., Dayton, OH.
2. DOT/FAA.CT-93/4, *Soft Ground Arresting Systems for Commercial Aircraft,* Interim Report, February 1993 by Robert Cook.
3. DOT/FAA/CT-93/80, *Soft Ground Arresting Systems for Airports,* Final Report, December 1993 by Jim White, Satish K. Agrawal and Robert Cook.
4. DOT/FAA/AOV 90-1, *Location of Commercial Aircraft Accidents/Incidents Relative to Runways,* July 1990.
5. UDR-TR-88-07, *Evaluation of a Foam Arrestor Bed for Aircraft Safety Overrun Areas,* 1988 by Cook, R. F., University of Dayton Research Institute, Dayton, OH.

# Braking Performance of Towbarless Towing Vehicles During Maintenance Tows

Gordon F. Hayhoe

Airport Technology R&D Team, AJP-6310
FAA William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405
(609) 485-8555
gordon.hayhoe@faa.gov

## ABSTRACT

Using a medium-sized Towbarless Towing Vehicle (TLTV) (16.8 tonnes (37,000 lb)) as an example, a comparative analysis is given of braking performance when applying either the TLTV brakes alone or the aircraft brakes alone when bringing a TLTV/aircraft combination to a complete stop from high speed and at maximum braking. It is shown that the stopping distance with tractor braking alone can be up to five times longer than the stopping distance with aircraft braking alone. It is also shown that the nose gear forces generated with aircraft braking alone can be excessive for lighter aircraft but allow a reasonable safety factor for heavier aircraft. Therefore, depending on the tractor/aircraft combination, it is possible to allow an operator in the cockpit of an aircraft being towed to apply the aircraft's brakes during an emergency. Each specific combination of tractor and aircraft should be analyzed separately before making such a determination for that combination.

## INTRODUCTION

Towbarless towing vehicles (TLTV) are frequently used for maintenance towing operations in which an aircraft is towed between a maintenance hangar and a terminal. A TLTV consists of a tractor in which the tires on the nose gear of the aircraft being towed are held in a cradle within the wheelbase of the tractor and with the tires held completely clear of the ground, see figure 1. Towing speed can be as high as 32 km/h (20 mph) and towing distances are frequently longer than 1.6 km (1 mile). Aircraft manufacturers typically recommend that personnel riding in the cockpit of a towed aircraft be instructed to not apply the aircraft's brakes during a tow so as to avoid damaging the nose gear of the aircraft. However, if an emergency occurs, the stopping distance when braking only the tractor may not be sufficient to allow the driver to avoid whatever caused the need for emergency braking. Using a medium-sized tractor (16.8 tonnes (37,000 lb)) as an example, a theoretical analysis is given of tractor/aircraft response when applying either the tractor brakes alone or the aircraft brakes alone. Maximum expected nose gear drag forces are then computed during aircraft braking for a range of aircraft and the results compared with the aircraft manufacturer's specified maximum allowable nose gear drag forces.


Figure 1. TLTV with the nose gear of an aircraft to be towed captured in the nose gear cradle (photograph courtesy of Goldhofer Airport Technology).

## ANALYSIS OF TRACTOR AND AIRCRAFT RESPONSE DURING BRAKING.

Nose gear braking forces under tractor and aircraft braking were calculated using a simple two degrees of freedom simulation model. The purpose of the simulation study was not to provide a precise representation of the motion of any specific aircraft and tractor combination, but to illustrate the major differences between braking with the tractor only and braking with the aircraft only. It also served to illustrate the effect of different braking strategies and to indicate what the worst-case nose gear force is likely to be during braking.

The two degrees of freedom in the simulation model are longitudinal motion of the aircraft and longitudinal motion of the tractor. The two bodies are connected in the longitudinal direction by a linear spring representing the nose gear and wheel-tire assemblies. The nose gear assembly is a flexible metal structure in the longitudinal direction and damping from this source will be very low unless the longitudinal motion of the nose gear couples into motion of the strut. Damping from flexure of the nose wheel tires within the tractor's cradle is difficult to estimate without direct measurement. In the absence of a direct measurement, a value of longitudinal damping of 7.3 kN/m/s (500 lb/ft/s) was assumed to provide some damping in the system. This value represents a damping factor of 0.017 for the tractor oscillating as a single degree of freedom system on the nose gear. A higher level of damping would cause the system oscillations to decay more rapidly, but it is unlikely that the true damping would be sufficient to significantly affect the conclusions drawn from the simulation study. Aircraft gross weight was set at 186 tonnes (410,000 lb) (typical of a Boeing 777-200 under tow) and the maximum braking effort from the tractor brakes

was set at 11.9 tonnes (26,240 lb). The maximum aircraft braking effort was set at 0.6 times the static weight on the aircraft main gear, and it was assumed that 90 percent of the static weight of the aircraft was supported by the main gear. For tractor braking only, the rolling resistance of the aircraft tires was set at 0.005 times the weight on the main gear. Tractor weight was set at 16.8 tonnes (37,000 lb), and the longitudinal stiffness of the nose gear and tires in combination was estimated to be 133 kN/m (180,000 lb/ft). For aircraft braking only, the rolling resistance of the tractor tires was set at 0.005 times the tractor weight plus the vertical force on the nose gear.

Figure 2 shows the aircraft speed, tractor speed, and tractor speed relative to the aircraft for full braking effort applied by the tractor as a step input (maximum braking from time zero). The initial speed was set at 32 km/h (20 mph). The sudden application of the tractor's brakes starts the tractor and aircraft oscillating relative to each other. The magnitude of oscillation of the aircraft is small compared to that of the tractor.

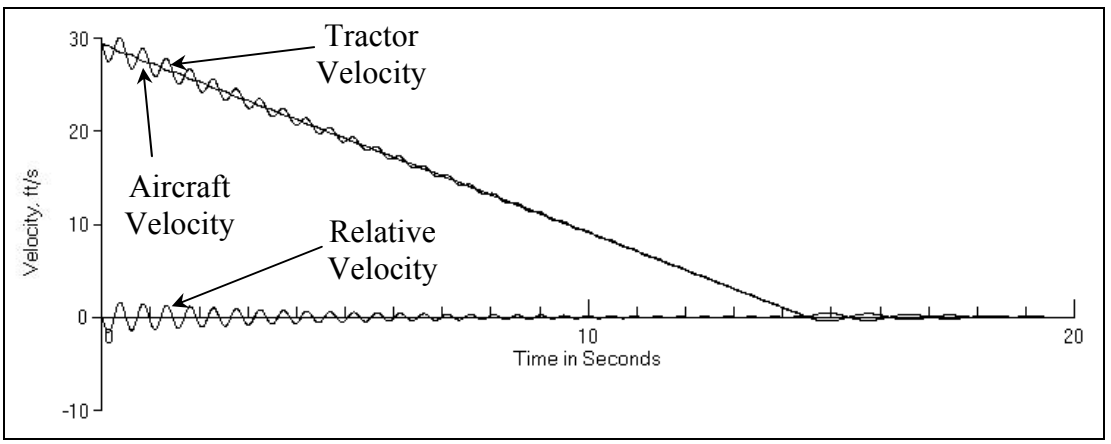



Figure 2. Velocities of Aircraft, Tractor, and Tractor Relative to Aircraft for a Step Input of Tractor Braking at an Initial Speed of 32 km/h (20 mph), Aircraft Weight = 186 tonnes (410,000 lb), and Stopping Distance = 64.77 m (212.5 ft), 1 lb = 4.45 N

Figure 3 shows the horizontal nose gear force. The sudden application of the tractor brakes results in a large amplitude oscillation of the nose gear force. This is almost damped out by the time the tractor has stopped. After the tractor has stopped, an oscillation of the aircraft on the nose gear is excited at a lower frequency. The aircraft oscillation decays according to the magnitude of the rolling resistance of the aircraft tires. The maximum nose gear force was 207 kN (46,500 lb), which occurred at the first peak of the tractor oscillation after brake application.

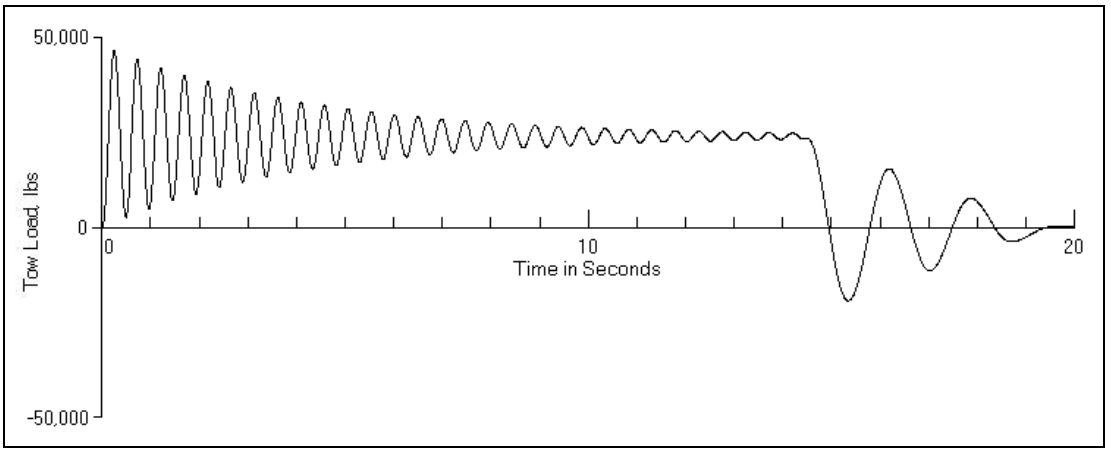



Figure 3. Horizontal Nose Gear Tow Load Force for a Step Input of Tractor Braking at an Initial Speed of 32 km/h (20 mph), Aircraft Weight = 186 tonnes (410,000 lb), and Maximum Tow Load = 206.9 kN (46,500 lb), 1 lb = 4.45 N

Figure 4 shows a time history of tractor speed and nose gear force recorded by Goldhofer Airport Technology (Goldhofer, 1999) during tests to satisfy the requirements of (Boeing, 2007). The trace shows two complete maximum acceleration and maximum braking tests for a Goldhofer AST-2 towing a Boeing 767 at 161 tonnes (355,000 lb) GVW. During the first test, braking starts at about 50 seconds and the braking force (tow load) builds up quite slowly and in stages. This is probably due to a combination of driver behavior and braking system delays. The large oscillations that occurred after brake application in the simulation, noted in figure 3, are not evident. When the tractor stops at about 72 seconds, a low-frequency oscillation similar to the simulation result noted in figure 3 is evident. In the second test, the braking force builds up much more rapidly. But, again, the large oscillations noted in figure 3 are not evident. The maximum tow load (and therefore nose gear force) was about 75.6 kN (17,000 lb).

As noted above, the simulation model did not reproduce the response shown in figure 4 for the time period just after brake application. Therefore, a ramp input of tractor braking force over 1 second was applied in the simulation instead of a step input. The results are shown in figures 5 and 6. The characteristics of the response are much closer to those of figure 4 than with the step input, particularly when compared to the second test in figure 4. The oscillation of the tractor relative to the aircraft after brake application is almost completely eliminated, but the oscillation of the aircraft, after the tractor stopped, is still present

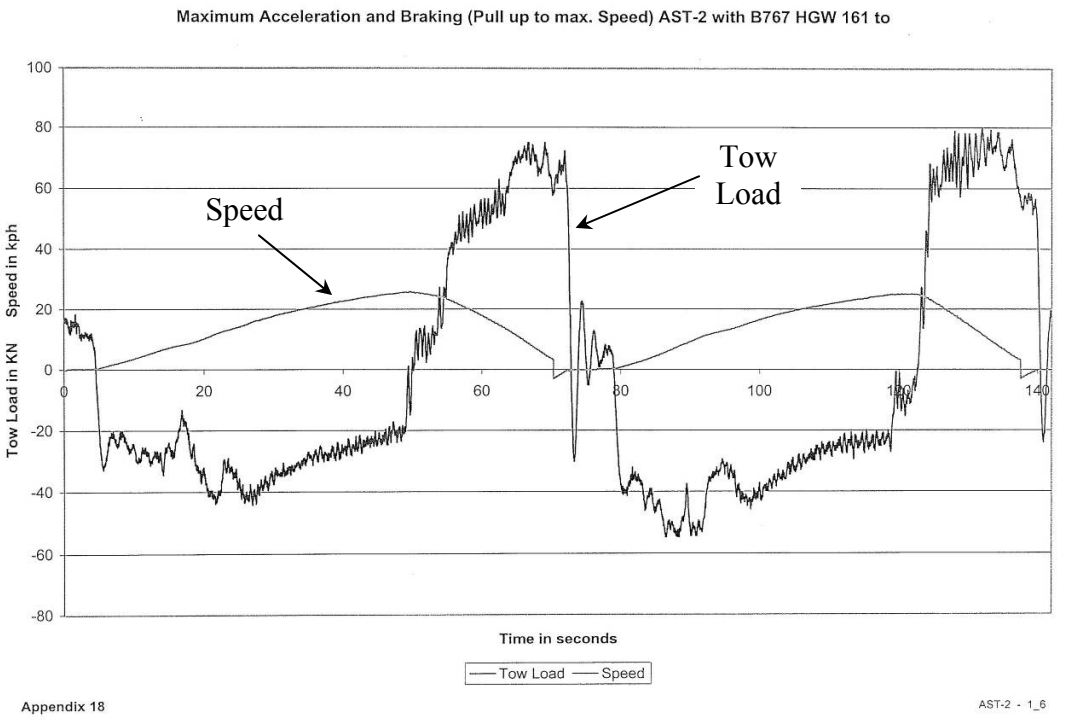



Figure 4. Time History of Tractor Speed and Nose Gear Force (Tow Load). From (Goldhofer, 1999). 1 km/h = 0.621 mph = 0.91 ft/s.

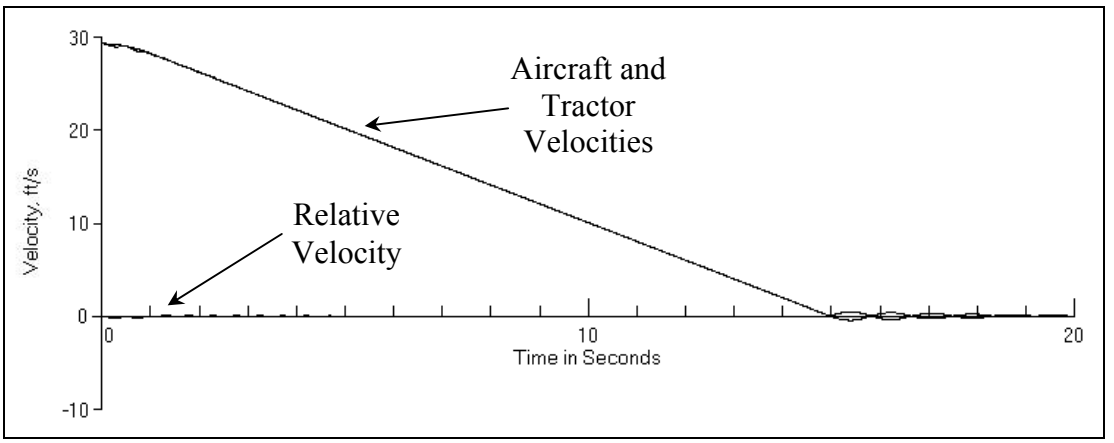



Figure 5. Velocities of Aircraft, Tractor, and Tractor Relative to Aircraft for a 1-Second Ramp Input of Tractor Braking at an Initial Speed of 32 km/h (20 mph), Aircraft Weight = 186 tonnes (410,000 lb), and Stopping Distance = 68.9 m (226.1 ft), 1 lb = 4.45 N

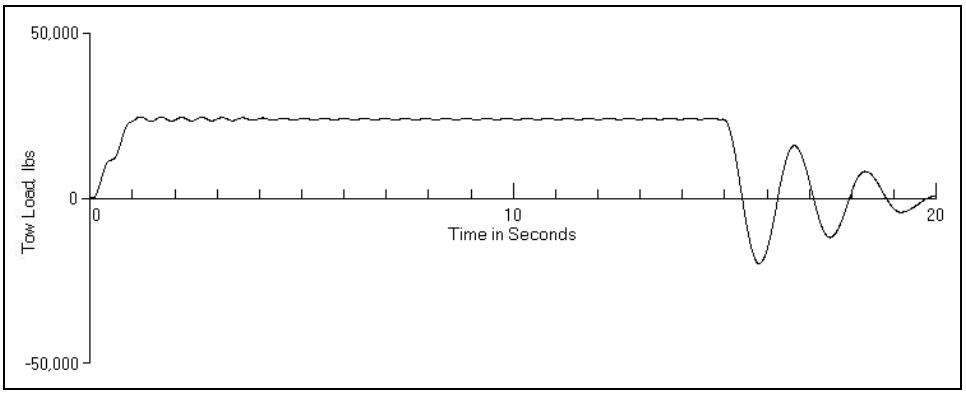



Figure 6. Horizontal Nose Gear Tow Load Force for a 1-Second Ramp Input of Tractor Braking at an Initial Speed of 32 km/h (20 mph), Aircraft Weight = 186 tonnes (410,000 lb), and Maximum Tow Load = 109.7 kN (24,660 lb), 1 lb = 4.45 N

Figures 7 and 8 show the response predicted by the simulation for a 1-second ramp input of aircraft braking, instead of tractor braking as before, and with only tire-rolling resistance acting on the tractor's wheels. The stopping distance for the combination vehicle is much shorter with aircraft braking only than with tractor braking only, at approximately one fifth the distance. But the tow load is 25 percent less at 81.7 kN (18,360 lb). The maximum force is composed primarily of the inertia force acting at the center of gravity of the tractor due to the deceleration caused by the aircraft braking force. Because the mass of the tractor is small compared to the mass of the aircraft, this component is, in this case, very close to the aircraft braking force divided by the weight of the aircraft, multiplied by the weight of the tractor ($0.6 \times 0.9 \times 16.8 \times 9.807 = 89.0$ kN (20,000 lb)). Figure 8 shows the tractor oscillation after the aircraft has stopped, as indicated by the higher frequency of vibration than for tractor braking. For comparison, the same simulation was run with a step input of aircraft braking. Figure 9 shows the nose gear force response. An oscillation of the tractor relative to the aircraft is excited after brake application, as before, and the maximum force almost doubles to 155.5 kN (34,940 lb). The approximate doubling of the force is explained by the fact that the force oscillates about the steady, completely damped, force. Since the oscillation starts at zero and damping is low, the first peak must be close to twice the steady force.

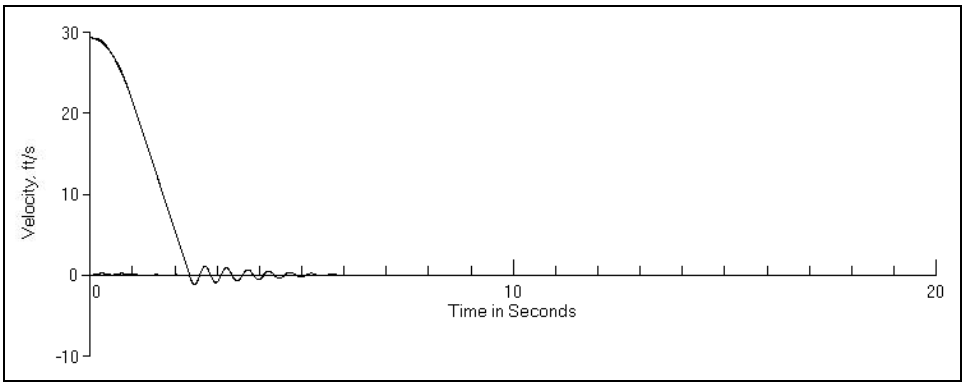



Figure 7. Velocities of Aircraft, Tractor, and Tractor Relative to Aircraft for a 1-Second Ramp Input of Aircraft Braking at an Initial Speed of 32 km/h (20 mph), Aircraft Weight = 186 tonnes (410,000 lb), and Stopping Distance = 12.47 m (40.9 ft), 1 lb = 4.45 N

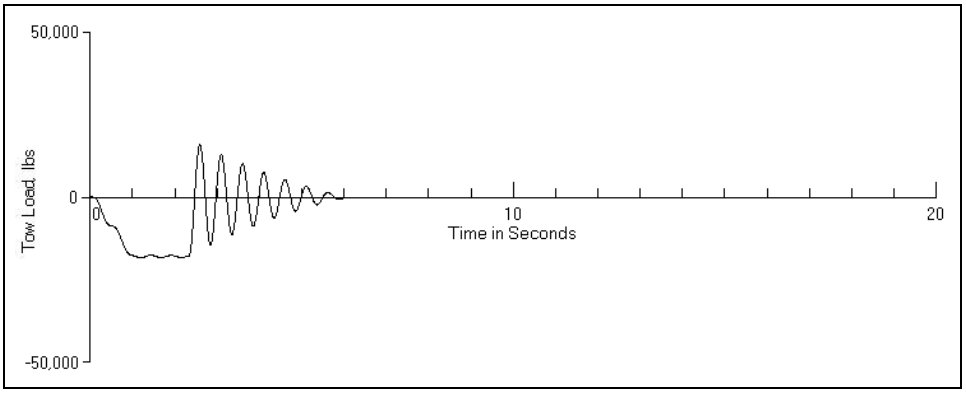



Figure 8. Horizontal Nose Gear Tow Load Force for a 1-Second Ramp Input of Aircraft Braking at an Initial Speed of 32 km/h (20 mph), Aircraft Weight = 186 tonnes (410,000 lb), and Maximum Tow Load = 81.7 kN (18,360 lb), 1 lb = 4.45 N

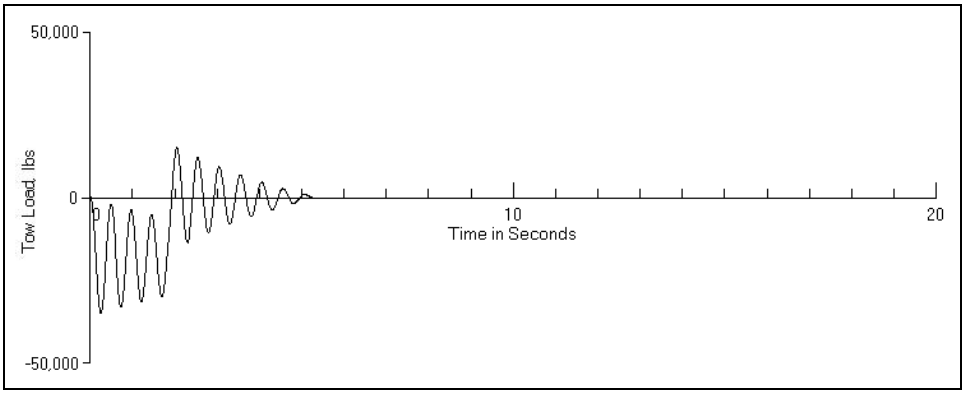



Figure 9. Horizontal Nose Gear Tow Load Force for a Step Input of Aircraft Braking at an Initial Speed of 32 km/h (20 mph), Aircraft Weight = 186 tonnes (410,000 lb), Maximum Tow Load = 34,940 lb, and Stopping Distance = 8.2 m (26.9 ft), 1 lb = 4.45 N

This analysis is essentially the same as the response spectrum analysis for shock loading on vibrating systems described in most textbooks on mechanical vibrations (see, for example, Thomson, 1981). In this case, the lower limit of predicted tow load is approximately equal to the steady tow load after the startup transient vibration has damped out, and the upper limit, resulting from the addition of the startup transient to the steady load, is approximately twice the lower limit. The steady tow load can be found from a static analysis of the inertia forces acting on the aircraft and the tractor. If it is assumed, as before, that the aircraft braking force generated by the main gear is 0.6 times 0.9 times the weight of the aircraft, and that the external forces acting on the tractor are zero, then the total inertia force from the deceleration of the complete tractor/aircraft combination is equal to the aircraft braking force, or

$$F_{A+N} = 0.6 \times 0.9 \times W_A = \frac{(W_A + W_T)}{g} \times a$$

where:

$W_A$ = weight of the aircraft
$W_T$ = weight of the tractor
$g$ = gravitational constant
$a$ = deceleration of the tractor/aircraft combination

therefore,

$$a = \frac{0.6 \times 0.9 \times W_A \times g}{(W_A + W_T)}$$

and the inertia force acting on the tractor is

$$F_N = \frac{W_T}{g} \times a = W_T \times 0.6 \times 0.9 \times \frac{W_A}{(W_A + W_T)}$$

$F_N$ is also equal to the horizontal nose gear force (the tow load). Substituting the tractor and aircraft weights of figures 7 and 8 gives a value of 81.6 kN (18,340 lb) for $F_N$. This is very close to the value of maximum nose gear force of 81.7 kN (18,360 lb) from the simulation results shown in figure 8. Figure 10 shows the range of predicted maximum nose gear forces compared to the tow load limits specified in (Boeing, 2007) for all Boeing and McDonald-Douglas aircraft rated for use with a Goldhofer AST-2 tractor, which has a gross operating weight of approximately 16.8 tonnes (37,000 lb) (see Hayhoe and Patterson, 2009). The 777 tow load limit is significantly higher than the trend line for all the aircraft tow load limits plotted in figure 10.

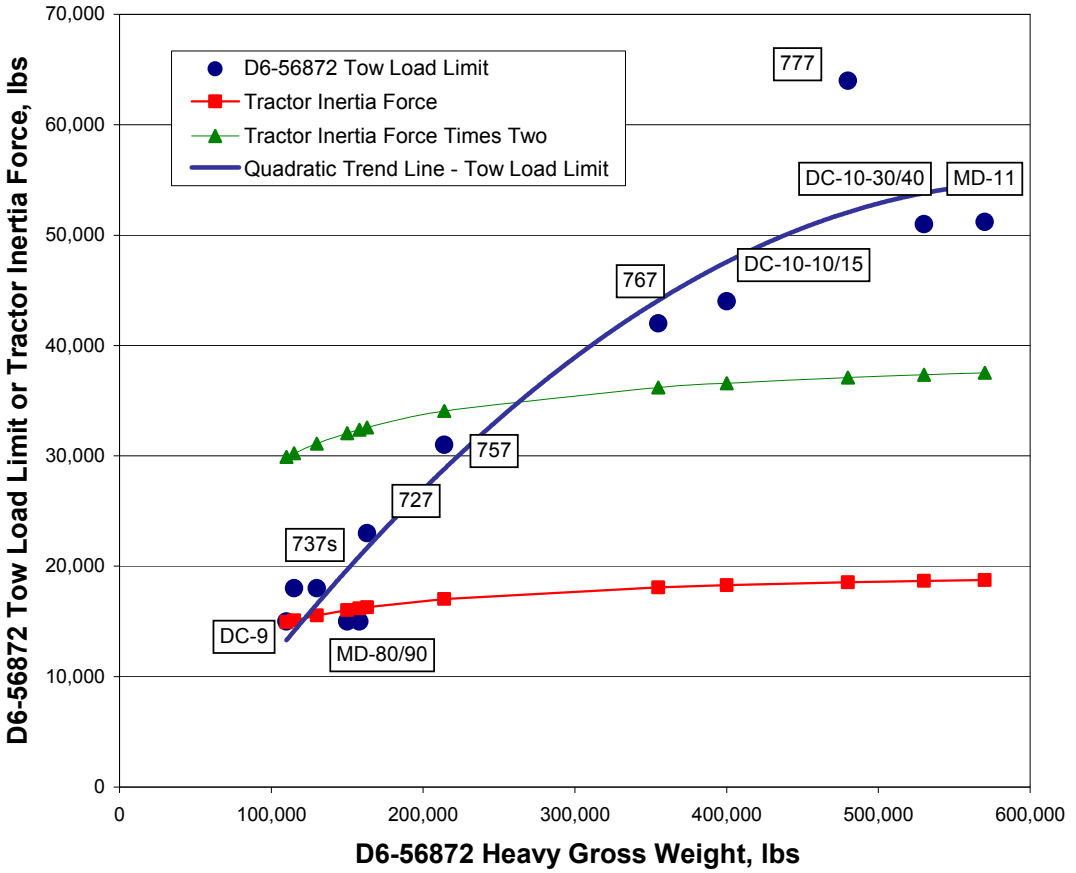



Figure 10. Predicted Nose Gear Force due to Aircraft Braking Compared With the Tow Load Limits Specified in Boeing Document D6-56872 for the Range of Aircraft Rated for Use With a Goldhofer AST-2 Tractor (1 lb = 4.45 N)

## CONCLUSIONS

Allowing an operator in the cockpit of an aircraft to apply the aircraft brakes in an emergency during a TLTV towing operation would allow the tractor/aircraft combination to be brought to a stop very quickly and greatly reduce the risk of injury to the tractor operators and reduce the risk of violating operating limits and requirements while maintenance towing on an airfield. A simulation analysis indicated that this is a viable option for normal operating procedures for medium sized tractors towing 777 aircraft, even though the Boeing Service Letter on TLTV operations (Boeing, 2007) recommends that aircraft braking not be allowed during TLTV maintenance towing. The reason that aircraft braking may be a viable option for certain tractor/aircraft combinations is that, for a given model of TLTV, deceleration of the tractor/aircraft combination under aircraft braking (and hence, the nose gear tow load) does not increase very much as the weight of the aircraft increases. However, the strength of the nose gear rises approximately in proportion to

the maximum gross weight of the aircraft, with the 777 higher than the trend line compared to earlier model aircraft.

## ACKNOWLEDGEMENTS

The work described in this paper was supported by the FAA Airport Technology Research and Development Team, AJP-6310, Dr. Satish K. Agrawal, Manager. The contents of the paper reflect the views of the author, who is responsible for the facts and accuracy of the data presented within. The contents do not necessarily reflect the official views and policies of the FAA. The paper does not constitute a standard, specification, or regulation. Thanks are also due to Goldhofer for providing information on TLTV operational characteristics and specifications and to the reviewers for their comments.

## REFERENCES

Boeing Commercial Aviation Services (2007), Service Letter ATA: 0910-00, "Towbarless Towing," April 6, 2007, includes Boeing Document D6-56872, "Towbarless Towing Vehicle Assessment Criteria," as Attachment II.

Goldhofer Airport Technology (1999), "Towing Tests With B767 and AST-2," March 12, 1999.

Hayhoe, Gordon F. and Patterson, James Jr. (2009), "Towbarless Towing Vehicle Operations – Evaluation of Braking Action and Vehicle Conspicuity," Report DOT/FAA/AR-TN09/44, U.S. Department of Transportation, Federal Aviation Administration, August, 2009.

Thomson, William T, "Theory of Vibrations with Applications," Prentice-Hall, N.J., 2nd Edition, 1981.

# Soil Improvement of Runway STRIP and Runway End Safety Area (RESA) Through an Innovative Methodology

M. Crispino[1], F. Giustozzi[2] and G. Paracchini[3]

[1] Professor, Politecnico di Milano, DIIAR –Section "Transport Infrastructures", 20133, Milan, Italy, email: maurizio.crispino@polimi.it
[2] Ph.D candidate, Politecnico di Milano, DIIAR –Section "Transport Infrastructures", 20133, Milan, Italy, email: filippo.giustozzi@polimi.it or fgiustozzi@vt.edu
[3] Responsible for Works, SEA Aeroporti di Milano, Milan, Italy, email: Paracchini.Giordano@sea-aeroportimilano.it

## ABSTRACT

National and international regulations define requirements for bearing capacity of runway strip and runway end safety areas. Bearing capacity requirements imply the reduction of damages to an overrunning aircraft, facilitating its deceleration and allowing the easy access of emergency vehicles. In particular, international regulations from ICAO impose a maximum sinking of the aircraft nose landing gear into the soil. Hence, any functional and structural intervention in such areas should take into account this goal. Furthermore, interventions in airports infrastructures should imply minimum interference to airport activities and should comply with sustainability and environmental requirements.

In this context, an innovative methodology, called *"Controlled Compaction Energy Approach"* was developed to stabilize soils in situ. The innovative approach does not have as main goal the maximization of density and bearing capacity as is usually the goal in traditional in situ stabilization procedures; the achievement of regulation requirements is based on the control of the compaction energy applied after stabilization, in order to develop only the required bearing capacity to obtain a surface that is not too stiff to be able to dissipate enough aircraft energy. The compaction energy needed strongly depends on different variables, such as the interstitial voids and water content and the amount of binder.

A full scale test area was developed in order to set up the technique; afterwards, a real scale application of the method was carried out at Linate Airport of Milan (Italy) within the functional requalification of the runway end safety area (RESA). Such RESA is currently under operation for about two years.

## INTRODUCTION

Airport areas are subject to regulation, standards and recommendations by international organizations, such as the International Civil Aviation Organization (ICAO) that are often adopted by local governments and aviation authorities.

ICAO Aircraft Accident/Incident Data Reports (ADREP) indicated that aircraft undershooting or overrunning the runway during landings and take-offs suffer significant damages.

Runway STRIP and Runway End Safety Area (RESA) constitute an area, surrounding the runway, long and wide enough to contain aircraft overruns and undershoots resulting from a reasonably probable combination of adverse operational factors. They should also provide a graded area on a certain inner portion, free from obstacles and with a bearing capacity that avoids damages to an overrunning aircraft, enhance the airplane deceleration and facilitate the movement of rescue and fire fighting vehicles. Furthermore, interventions on airports infrastructures should imply minimum interference with airport activities and should comply with sustainability and environmental requirements. For M&R (maintenance and rehabilitation) or new interventions in airports areas reuse and valorization of existing soils and recycled material coming from the reconversion of waste products should be implemented according to sustainability.

In this context, the present paper presents an innovative methodology aimed at improving the bearing capacity required on particular and sensitive areas; the main goal is not the maximization of density and strength but the achievement of the current regulations in terms of bearing capacity.

## DEFINITIONS AND STANDARDS

ICAO Annex 14 and the Aerodrome Design Manual (DOC. 9157 – Part 1, Runways) establish geometrical and structural requirements for runway STRIP and RESA. Geometrical standards set up length, width and slopes of the predicted areas; structural requirements, instead, concern load-bearing capacity. In particular, the above mentioned areas should be able to prevent the collapse of the nose landing gear of an aircraft. The surface should provide enough drag to an overrunning aircraft and it should have sufficient bearing strength to avoid damages. Aircraft manufacturers consider that 15 cm is the maximum depth to which the nose landing gear may sink without collapsing (ICAO, DOC. 9157 – Part 1, Runways). Therefore, the standards recommend that the soil at a depth of 15 cm below the finished strip surface should be prepared to have a bearing strength in terms of CBR between 15 and 20. The top 15 cm may be of lesser strength which would facilitate deceleration of the aircraft.

The standards clearly define a *differential* strength to be achieved for the soils involved. They do not impose the total lack of settlements but just the maximum limit in accordance with aircraft manufacturers' instructions; that's because the aircraft, in case of overrun/undershoots, should be able to face the resistances offered by the soil in order to slow down and stop.

## CEMENTITIOUS STABILIZATION

The stabilization treatment of in-situ soil allows to reduce materials supplying and new soil consumption, limiting in this way the soils handling across the

construction site and increasing the productivity; moreover, the reuse and the valorization of existing soils with the enhancement of their characteristics through innovative treatments lead to limit the use of new virgin aggregates and wasted material production.

Using cementitious materials generally allows the enhancement of in situ soils to levels consistent with the requirements of a given design application.Through the years, cement has been declared to be effective in stabilizing a wide variety of soils, including granular materials, waste materials (RAP) and crushed concrete, although a preventive mineralogy test is always suggested; the usual cement content varies between 2.5 and 4.5 percent by weight of dry soil (Little 2000).

Portland cement is made by calcium silicates and aluminates that, if blended with water, could react hydrating themselves to create the cementing complex of calcium-silicate-hydrate and calcium-aluminate-hydrate. The outcome is a moisture-resistant material that provides high durability and resistance to leaching over time (National Lime Association 2004). The usual enhancements achieved through a stabilization process consist of a better soil gradation, reduction of swelling potential or plasticity index, improvements in durability and strength, increase of the shrinkage limit, reduction of clay/silt-sized particle and raise the resilient modulus (Bahar 2004).

The construction stage of a cement modified soils is usually a fast, straightforward process. Indeed, cement can be merged into soil in a great number of ways. The most frequent method consists of spreading dry cement in measured amounts on a graded soil and blends it with a mixer to a specified depth. Besides all technical and economic advantages (speed up of construction stages, lower costs of materials supplied, etc.) a huge saving of non-renewable resources is performed by stabilization.

The present paper presents a new kind of stabilization method in order to achieve the particular regulation requirements in terms of bearing capacity of sensible airport areas, previously mentioned in the previous paragraph. In fact, an unusual binder content, much lower than the ones usually adopted, was used as later on described. This kind of *"soft stabilization"* approach should be able to afford bearing capacity requirements reported by the standards.

## CONTROLLED COMPACTION ENERGY APPROACH

A flow of dense granular material involves various physical processes such as friction and particle collisions. Compaction of granular systems is a complex phenomenon that involves motion inside a certain volume, particle mobility and caging effect. Grains mobility is also deeply controlled by the composition of the soil mixture (Ludewig 2006).

An efficient compaction (95-100 % of the standard/modified Proctor test) makes possible to substantially improve the bearing capacity and stability of a granular mixture, increase the impermeability and, in most cases, to minimize future settlements. In in-situ stabilizations, after the cement binder is spread over the soil

and blended with a mixer, compaction is required to reach the requested bearing capacity.

In this context the present research, basing on a well-established state-of-the-art, was aimed at finding a way to control compaction and not to achieve the highest degree of compaction, as generally occurs or it is needed. It's known that for cement modified soil, a density increase of 5 percent will increase the strength up to 50 percent.

Influence of frequency and amplitude on the compaction effect has been discussed as long as vibratory rollers have been used. As a rule the compaction effect of granular mixture and soils has a maximum value at frequencies between 25 and 50 Hz, but most of the frequency curves are usually flat. An increase in amplitude gives a pronounced improvement in compaction and an "in-depth" effect in the entire frequency range. For stabilized soils the optimum amplitude is usually established from 0.4 to 0.8 mm and the suitable frequency ranges from 33 to 50 Hz (Forssblad 1981). Several methods of calculation indicate that frequency variations have a limited effect on the vibrating force. The energy transmitted to the soil and the compaction effort per unit volume (energy) can approximately be expressed as:

$$Compaction\ energy = f_1(L) + f_2(a\,f\,/\,v) \quad (1)$$

where: $f_1$, $f_2$ are functions; $L$ is the static linear load; $a$ is the amplitude; $f$ is the frequency and $v$ the roller speed (Forssblad 1981).

The compaction method adopted is based on a smaller number of roller passages with high amplitude of compaction and low/medium rollers' weight. Through an experimental full scale test section, later on described, it was established a relationship between the deformation modulus $M_d$ (Static Load Plate Test) of the cement-modified soil and the number of roller passages for each combination of compaction procedure (vibratory roller vs. pneumatic roller, dynamic load vs. static load, low amplitude vs. high amplitude, etc.). The deformation modulus $M_d$ was then related to the CBR index, computing in this way the compaction energy due to a single passage and the energy gap included between two following passages; for each compaction combination a diagram relating CBR index with the number of passages was therefore developed identifying the exact passages amount that permits to match requirements. The proper combination of controlled compaction energy and low binder content has proven to lead to the required strength and bearing capacity.

## CASE STUDY

In order to improve a feasible methodology that allows the complete achievement of international standards and regulations concerning the strength and the bearing capacity of a runway STRIP or a RESA, a full scale test section was directly set up in-situ to develop a stabilization procedure and a compaction mode that didn't provide to reach the maximum interstitial voids density and bearing capacity but would comply with the requirements expressed by ICAO standards. The

controlled compaction energy approach was therefore set up and adopted to reach the goals previously described.

The technical proposal offers remarkable advantages especially on the execution stage: valorization and reuse of on-site materials, productivity increase and consequent reduction of time for the project achievement, more independence from climate factors (above all rain) that could affects the yard works (long delay could in fact occur if rain would fill the dig left by the removed soil to be replaced), reduction of soils handling and interferences with the airport's operativity. The methodology applied is quite general and can be easily extended to other similar cases.

**Project presentation.** The project aimed at the functional requalification of the RESA at Linate International Airport of Milan following the national and international current standards mentioned above. The new RESA measures 150 x 240 m (36,000 square meters of overall surface) beyond the end of the runway head 18L. The primary project need was to arrange a final designed surface characterized by evenness and strength consistent with Italian and International regulations and, in anticipation of future, to provide a firm support for the installation of an aircraft arrestor bed. Italian Authority for Civil Aviation even required a deformation modulus $M_d$ (CNR-BU standard 146/92) equal to 100 MPa.

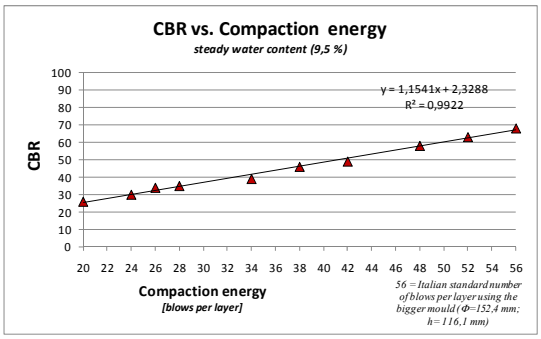
After characterizing the mineralogical composition of the on-site material, it was chosen to adopt a modification of the soil properties by stabilizing it with cement. In particular, the technical proposal included: the surface removal of the organic soil (10-15 cm thickness); landfill with suitable materials and grading; soil stabilization with cement (40 cm thickness); covering with grass (7 cm thickness) to avoid the creation of dust.

The achievement of bearing capacity requirements, in terms of strength and landing gear sinking, required a thorough analysis of the variables involved and their combinations that could affect the final outcomes. Regarding to that, a test campaign was performed both in laboratory and on-site in order to establish: the optimal mixture formula of the treated and strengthened soil; the compaction methodology (roller type, number of roller passages, etc.).

**Laboratory investigations.** Soil samples were taken out at different depths to classify and characterize the on-site material and better understand its mechanical properties and mineralogy. The investigations were carried out to choose the best-fit binder, its relative contents and the final compaction energy for the soil stabilization. Through in-situ sampling it was established the terrain stratigraphy finding the organic topsoil thickness (almost 15 cm) to be removed before the stabilization treatment took place.

Grain size analysis and soil consistency tests (liquid limit test, plastic limit test) allowed to classify the soil, designating it as A-4, according to the AASHTO classification system, with a plasticity index equal to 9. CBR index was equal to 12 % (water content = 9.3 %), below the international standards limitations. It was chosen to add Portland cement II/A 32.5 to the natural soil with a content of 1.5

percent (by weight of the dry soil) to improve its mechanical characteristics. Laboratory compaction test (modified Proctor test) allowed to determine the optimum water content (9.5 %) and the corresponding maximum dry density (1.980 g/cm$^3$) of the cement modified soil. After a day of curing and a day of imbibition, several CBR tests were executed on the cement modified samples with different water contents and various compaction energies (Khan 1998), as showed in figure 1, to establish the relationship between these parameters.




| | | | | |
|---|---|---|---|---|
| natural in situ soil | sample | 1 | 2 | 3 |
| | liquid limit | 29 | 32 | 26 |
| | plastic limit | 22 | 23 | 22 |
| | ***plasticity index*** | ***7*** | ***9*** | ***4*** |
| cement modified soil | sample | 1 | 2 | 3 |
| | liquid limit | 28 | 28 | 27 |
| | plastic limit | 23 | 24 | 23 |
| | ***plasticity index*** | ***5*** | ***4*** | ***4*** |

Figure 1: CBR index for different compaction energies *(Proctor test)* and plasticity index *(before-after stabilization treatment)*

As a first comment, it should be noticed that the stabilization with cement improved the mechanical properties of the cement-modified soil providing a bearing capacity enhancement (CBR increase) and a water susceptibility reduction (plasticity index decrease). Moreover, the CBR index of the cement-modified soil, when compacted according to usual standards, was proved to be too high for the project purpose; a controlled compaction energy should consequently be applied.

**On site full scale test area.** After the outcomes analysis from laboratory investigations, a full scale test section (5 m wide per 100 m long) was developed directly on the airfield RESA (Figure 2) to set up the soft soil stabilization with a controlled compaction energy approach.

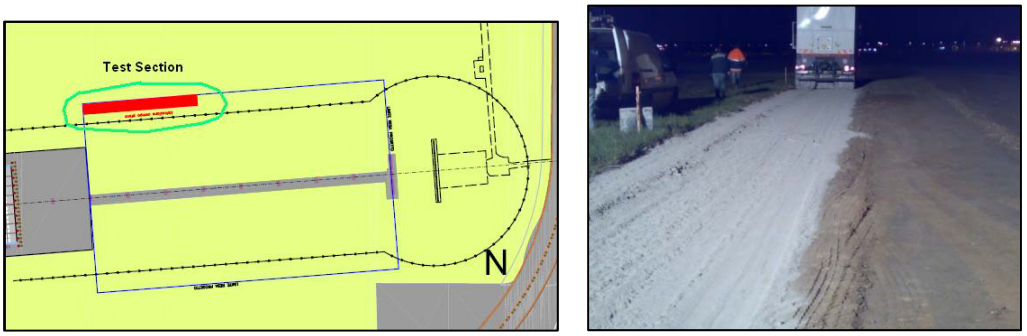



Figure 2: full scale test section

The test section was then split into four rectangular subsections in order to investigate different cement contents (cement content ranging from 1 to 3 percent by weight of the dry soil) and compaction methodologies (vibratory/pneumatic rollers combination and passages, rollers weight, frequency and amplitude on compaction effects). After removing the organic topsoil and estimating the on-site moisture content, a preliminary grading selection was performed. The cement powder was then spread over the graded surface (adding water if needed) using different binder contents, a pulvimixer afforded the mixing action. Finally, different compaction methodologies were tested and static load plate tests were continuously executed together with settlement measurements under a truck (with a pre-defined axle load).
Two different rollers were used: a single drum vibratory roller (15.8 tons) and a pneumatic roller (21 tons, ballasted).

Each sub-section was compacted using different combinations of the previous rollers changing the number of total passages (ranging from an absolute value of 6 to 14), the quantity of each roller type's passages (vibratory/pneumatic) and the amplitude of the compaction force during the dynamic action. After each passage the deformation modulus $M_d$ was evaluated such as at the end of the expected compaction sequence and after 48 hours of curing to monitor the stabilization development over time; a relationship between the number of passages, the deformation modulus and the CBR was therefore established in order to achieve the requested project strength.

Finally, the cement content and the compaction procedure were fixed as summarized below (Table 1).

| **Cement content** [optimal moisture (9.5%)] | **Compaction sequence** |
|---|---|
| 1.5 % *(by weight of the dry soil)* | ▪ Single drum vibratory roller *(≈ 15 tons, high amplitude)*: 6 passages<br>▪ Pneumatic roller *(≈ 20 tons)*: 4 passages<br>▪ Single drum vibratory roller *(static mode, ≈ 15 tons)*: 1 passage |

Table 1: final parameters for the soft stabilization and controlled compaction energy

It is important to highlight that:

- the single drum generally used to compact stabilized soil has a weight around 19-20 tons whereas it is no more than 15 tons;
- pneumatic roller has a maximum weight of around 25-27 tons, whereas it is around 20 tons;
- passages for finishing are generally at least two or four whereas it is just one.

**Real scale application.** Basing on the values obtained from both the laboratory investigations and the on-site full scale test section, the project was finally implemented working during the airport closing time at nighttime. The following structure (Figure 3) was adopted to improve RESA strength according to the international and national standards:

- Organic topsoil _ *7 cm*
- In-situ soft cement stabilization *(controlled compaction energy approach) _ 40 cm*
- Subgrade


Figure 3: RESA design structure

Construction procedures were the same already described in the previous paragraph. As a conclusive test and in order to inspect the effectiveness of the construction stage, $M_d$ measures were carried out over the RESA obtaining a mean value of 104 MPa, very close to Italian Authority requirements (100 MPa).

## CONCLUSIONS

Stabilization of soil represents a great opportunity from the environmental point of view (less materials handling, lower consumption of virgin material, etc.). This very well-known technique can express new potentialities applying approaches different from traditional, as the one here presented. Acting on the quantity of binder content (and type) and compaction energy, desired outcomes concerning strength and bearing capacity could be accomplished. The achievement, in the case study presented, of the requirements reported by Italian and international regulations was based on both such aspects and specifically on the control of the compaction energy applied after stabilization, in order to develop the required, and no more than that, bearing capacity (otherwise a too stiff surface would be obtained, not able to dissipate the aircraft energy).

At present, the treated area at Milano Linate airport, although no overrunning aircraft entered the area, does not present any type of distress after two years from the work. New measurements are planned in order to prove the effectiveness of the *controlled energy approach;* the trend over time of the deformation modulus should be analyzed in order to inspect how the curing time of the cement has affected or not the *controlled* bearing capacity achieved during the soil stabilization stage.

## REFERENCES

Baer, M.R. (1988). "Numerical Studies of Dynamic Compaction of Inert and Energetic Granular Materials", Journal of Applied Mechanics, Volume 55, Issue 1, 36 (8 pages).

Bahar, R. & al. (2004). "Performance of compacted cement-stabilized soil", Cement and Concrete Composites, Volume 26, Issue 7, Pages 811-820.

CNR-BU n°146/92. "Determinazione dei moduli di deformazione Md e Md[1] mediante prova di carico a doppio ciclo con piastra circolare", Consiglio Nazionale delle Ricerche, Norme Tecniche.

Crispino, M., Rampini, R., and Pozzi, M. (2007). "An experimental analysis of the effects of compaction on asphalt pavement macrotexture", Conference: "Advanced Characterization of Pavement and Soil Engineering Materials Conference", Athens, Greece.

Crispino, M. , Mariani, E. and Rampini, R. (2010) "Increasing asphalt pavements durability through accurate construction: a model for compaction design", Structure and Infrastructure Engineering, Taylor & Francis, DOI: 10.1080/15732471003588767.

Department Of The Army, The Navy, And The Air Force. (1994). "Soil Stabilization For Pavements", TM 5-822-14/AFJMAN 32-1019.

Floss, R. (2001). "Compaction technology in earthwork, highway and Transportation Engineering", volume I, Bomag GmbH and Co. OHG, Boppard, Germany.

Forssblad, L. (1981). "Vibratory soil and rock fill compaction", Svedala Dynapac Publication, Sweden.

Guthrie, S.W. (2001). "Selecting optimum cement contents for stabilizing aggregate base materials", Texas DoT and FHWA, Report No. FHWA/TX-05/7-4920-2.

Khan, Z.A., Al-Abdul Wahab, H.I., Asi, I., and Ramadhan, R., (1998). "Comparative study of asphalt concrete laboratory compaction methods to simulate field compaction". Construction and Building Materials, Volume 12, Issues 6-7, Pages 373-384.

I.C.A.O. (2004). Annex 14 – Volume I, "Aerodrome", Aerodrome Design and Operations, Fourth Edition.

I.C.A.O. (2006). Aerodrome Design Manual, DOC. 9157 – Part 1, Runways.

Little, D.N. & al. (2000). "Cementitious stabilization", Transportation Research Board, 00784589, 7 pages, A2J01: Committee on Cementitious Stabilization.

Ludewig, F. & al. (2006). "Compaction of granular mixture", Granular Matter, Springer, Volume 8, Number 2, pages 87-91.

National Lime Association. (2004). "Lime-Treated Soil Construction Manual", Bulletin 326.

Philippe, P. and Bideau, D., (2001). "Numerical model for granular compaction under vertical tapping". Physical Review E, Volume 63, Issue 5, Phys. Rev. E 63, 051304 (9 pages), USA: The American Physical Society.

# Case Study of the Sustainable Parking Facility at Stewart International Airport

**Dan W. Louie, P.E[1], Joseph A. Calautti, P.E.[2], Scott D. Murrell, P.E.[3]**

[1] Senior Civil Engineer, The Port Authority of NY & NJ, Two Gateway Center, 16th Fl, Newark, NJ 07102; PH (973-792-4340); dlouie@panynj.gov
[2] Principal Civil Engineer, The Port Authority of NY & NJ, Two Gateway Center, 16th Fl, Newark, NJ 07102; PH (973-792-4337); jcalautt@panynj.gov
[3]Chief Civil Engineer, The Port Authority of NY & NJ, Two Gateway Center, 16th Fl, Newark, NJ 07102; PH (973-792-4327); smurrell@panynj.gov

**Abstract**
Stewart International Airport (SWF) is located west of Newburgh, New York, in the southern Hudson Valley, and 60 miles (88.5 km) north of New York City. It was originally developed in the 1930s as a military base to provide aviation training for the cadets at the nearby United States Military Academy at West Point. Over the years, it has grown into the major passenger airport for the mid-Hudson region and continues to serve as a military airfield. Currently, SWF houses the 105th Airlift Wing of the New York Air National Guard.

In 1991, the commercial service airport went private when United Kingdom-based National Express Group was awarded a 99-year lease on the airport by its owner and operator at that time, the New York State Department of Transportation. On January 25, 2007, the Port Authority of New York and New Jersey (PANYNJ) Board of Commissioners voted to acquire the remaining years of the airport's lease, effectively ending the privatization of SWF. The PANYNJ assumed control of the airport on November 1st of that year.

Prior to taking over the airport, the PANYNJ projected that the airport's annual passenger traffic would soon be triple that of 2006. In 2007, Stewart International Airport achieved this and handled 970,000 passengers. At that time, it became apparent that the airport lacked adequate parking facilities to address the rapid increase of annual passengers and the holiday seasonal demands. Although air travel at SWF decreased following the recent economic downturn, demand for parking is expected to increase, backed by a 2010 customer survey at SWF that showed 65 percent of the airport's passengers arrive by personal car, and 28 percent use the airport's parking facilities. In addition, forecast passenger growth for the next decade projects passenger levels to ultimately exceed 2007 totals. This led the PANYNJ to embark on the expansion of the Terminal's Lot "A" public parking facility, increasing it from the current 8.7 acres to 15.1 acres.

The PANYNJ Engineering Department issued Contract No. SWF 164.007 for expanding the parking lot facility. The expansion consists of a pervious asphaltic pavement and an extensive surface and subsurface storm drainage collection system. It is designed to collect, detain and allow infiltration of the 100-year storm run-off

without any peak discharge into nearby streams and open water bodies, and also recharge the groundwater table. Best management design practices include the use of infiltration bio swales, a granular pavement sub-base with large void content, installation of "raintanks" and a sub-drainage piping system. In addition, a portion of the storm water run-off will be diverted into a cistern and utilized by an underground irrigation system with solar powered pumps for the watering of the parking lot's landscaped areas.

**Introduction**

The Parking Lot "A" expansion would be the first major parking lot for the PANYNJ to incorporate green, environmental friendly initiatives. The guiding principles of the design were to reduce the environmental impact and increase sustainable awareness within the Authority by testing new techniques, materials and methods of sustainable design.

The existing main parking lot accommodates approximately 840 parking spaces. The expansion (see Figure 1) will double its capacity by an additional 830 parking spaces. The project will also provide an improved access roadway to the airport's terminal, an additional parking lot entrance plaza and exit plaza expansion.

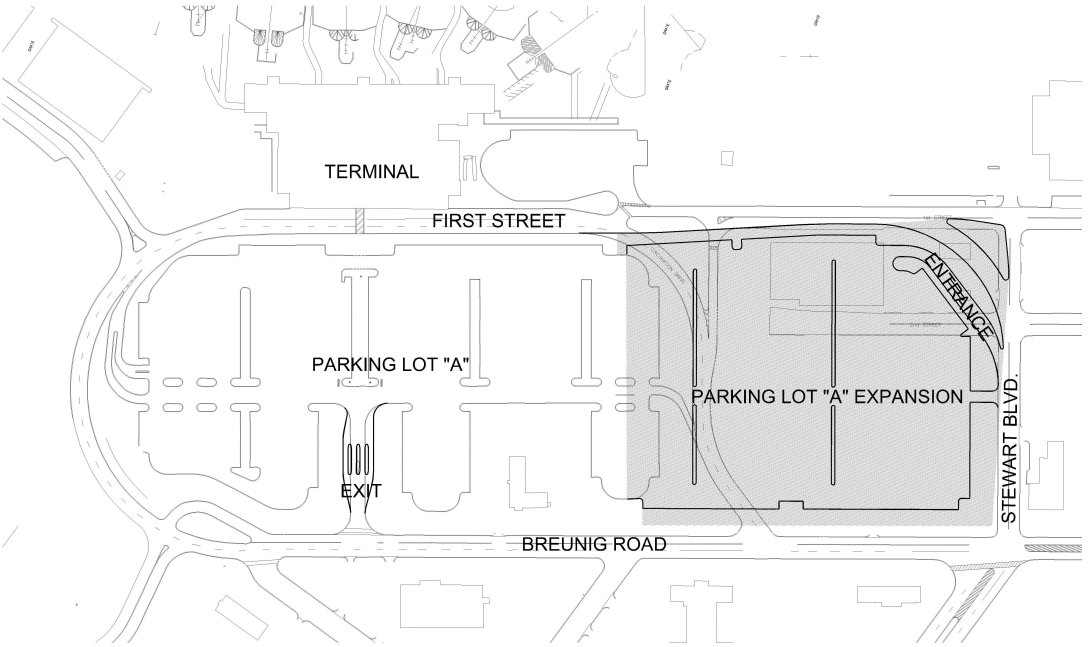



**Figure 1: Stewart International Airport - Parking Lot "A", Plan View**

**Best Management Practices (BMP)**

The requirements for selecting the Best Management Practice were dictated by space requirements, existing storm drainage utilities and FAA airport policies. Several methods of mitigation were considered. An open detention/retention basin is unsuitable due to limited space. Also, it would create standing water which would attract birds, and therefore pose a hazard for aircraft. Due to the risk of structural

failure resulting in pavement collapse, the use of an underground storage system for storm run-off was eliminated. In addition, such a system would require greater monitoring/maintenance of the underground chambers, a storm water quality control device and its maintenance, and an upgrade to the existing storm water trunk system in order to accommodate overflow from extreme storm events. A conventional storm drainage collection system connecting to the existing storm water collection system was not feasible since it required an upgrade of 3800 linear feet of the existing storm drainage trunk line, the expansion of an existing detention basin and the use of a storm water quality control device. The outflow will contribute to downstream overflow and exacerbate flood conditions.

The Best Management Practice selected was Infiltration and Bioretention. The majority of the infiltration will be handled by the pervious pavement system. This type of pavement allows rainfall and run-off to soak through the pavement surface course of open-graded asphalt concrete mix and accumulate in a porous sub-base consisting of high porosity stone base from which water would percolate naturally into the ground below. The application of pervious pavement systems is ideal for parking lots due to low traffic densities and minimal load-bearing requirements. The parking lot islands and dry swale comprise the Bioretention system. The advantages of these types of systems is that all run-off remains on-site, recharges the groundwater, and provides for the removal of Total Suspended Solids (TSS) and Total Phosphorus (TP) with simple monitoring and maintenance. The pervious pavement requires periodic visual surface inspection to check for clogging and periodic cleaning by a vacuum truck.

The Bioretention features were designed to be a supplemental and redundant system. Supplemental, if the pervious pavement system was overwhelmed by an unusually intense storm event and redundant, if the pavement became clogged. The parking lot islands and dry swales intercept, filter, treat, store and infiltrate run-off. The benefits of pervious pavement, biotention islands and dry swales are the reduction of storm water overflow pollution, the relief of flash flooding, the preservation of vegetation, the retention of natural drainage patterns and the augmentation of municipal water supplies.

For snow removal operations of a pervious pavement, equipment should be fitted with plastic plow blades and pads on the blade shoes so as to not damage or gouge the surface. Also, sanding operations of the pavement should not be performed. Deicing chemicals may be used for ice control. However, such chemicals should be environmental friendly since they would ultimately permeate into the groundwater. The natural heating of the pervious pavement by the sun, whereby melting the snow and ice, will reduce the amount of de-icing products needed (The Franklin Institute Research Laboratories, 1972).

**Design Elements**

The Pervious Pavement System (PPS) consists of suitable sub-grade, fill, geotextile filter fabric, free draining stone base reservoir, sub-drainage piping, crushed stone choker course and open-graded asphalt top course (see Figure 4). The parking lot diagonally spanned 780 feet with a change of elevation of 22 feet. The sub-grade is composed of glacial till deposits with a permeability of 1.70 inches/hour and a water table of 5 feet below the lowest stone sub-base bottom elevation. The parking lot required fill to meet the proposed grades. The borrow pit was located 1200 feet north of the project site and the soil had the same characteristics of the in-situ sub-grade. In order to accommodate the grade changes and the need for the reservoir bottom to be level, the soil infill was then shaped in terraces with berms to detain or compartmentalize the captured run-off. The fill was covered with a non-woven geotextile filter fabric to prevent the fill from contaminating the drainage stone base reservoir. The drainage stone base reservoir and sub-drainage piping was then installed. The stone base reservoir consisted of AASHTO No. 2 Uniform Graded Coarse Aggregate providing a 40% void space. The thickness of the stone base reservoir varied from 18 to 36 inches. The sub-drainage pipe consisted of 6- inch diameter perforated high-density polyethylene (HDPE) pipe with factory-installed geotextile wrap and was installed 3 inches from the bottom of the stone base reservoir to increase capacity and evenly distribute run-off within each terrace area. A crushed stone choker course was installed on top of the stone base reservoir to help stabilize sub-base for the placement and compaction of the asphalt concrete wearing course. The crushed stone choker course is comprised of clean washed AASHTO No. 57 aggregate stone. The open-graded asphalt top course mix design (see Table 1) was adapted from the NJDOT (NJDOT, 2007, Section 902.03). The NJDOT has used the mix to mitigate roadway noise. The mix called for 6% by weight of Performance Grade (PG) 76-22 modified asphalt binder with a cellulose fiber. The high percentage of asphalt binder required the addition of cellulose fiber . This serves as a stabilizer to prevent separation of the asphalt binder from the aggregate. The aggregate gradation is composed of some 3/8” stones, mainly No. 4 stone and a minimum of fines. The smaller aggregate stone size promotes a greater surface contact area with each other. The high asphalt binder content aids in preventing the aggregate from raveling under traffic.

| AGGREGATE GRADATION | | | | | |
|---|---|---|---|---|---|
| SEIVE SIZE | 1/2" | 3/8" | NO. 4 | NO. 8 | NO. 200 |
| % PASSING | 100 | 85-100 | 20-40 | 5-10 | 2-4 |
| 1) PG 76-22 ASPHALT BINDER | | | | | |
| 2) CELLULOSE FIBERS, DOSAGE RATE OF 0.3% BY WEIGHT OF TOTAL MIX | | | | | |
| 3) 6.0% MODIFIED ASPHALT BINDER CONTENT | | | | | |
| 4) 18% - 26% MARSHALL AIR VOIDS | | | | | |

**Table 1: Open-Graded Asphalt Concrete Top Course Design Mix**

The concrete block pavers used along curb edges of frontage roadway, entrance and exit plazas are also permeable. The block pavers are pre-cast concrete units that fit

together to create large gaps which to allow surface run-off to flow between blocks. Block pavers are used in lieu of grass along the roadway curbs because splashed snow salt is harmful to grass along the curb lines. The installation of permeable block pavers reduces landscaping maintenance, but still maintains infiltration. The block pavers sit on a layer of sand, geotextile fabric and 6 inches of open-graded aggregate base course. The block pavers joints are filled with crushed screenings.

The Bioretention Systems are located in the center islands of the parking lot and dry swales surround the lower edge of the parking lot. One part of the system are the weeps. The weeps (see Figure 2) are three feet wide cut outs along a concrete curb to allow run-off to pass through to the other side. On the other side of the curb there is either a grass or stone area to capture the excessive overland run-off which fails to infiltrate through the pavement surface during high intensity rainfall events. The parking lot island intercepts the sheet flow and the dry swales capture overflow from the lower elevations of the parking lot. The parking lot islands are filled with a planting soil mix of one-part compost and three-parts existing site soil that will filter both TSS and TP. Geotextile wrapped “raintanks“ were placed at a minimum depth of 36 inches below the soil surface. The “raintanks” are webbed plastic boxes that resemble plastic milk crates. The parking lot islands are only 6 feet wide from outside curb to outside curb and the “raintanks” function as a reservoir that increases storage capacity of that narrow island. The dry swales (see Figure 3) intercept, convey, filter, store and infiltrate run-off that overflow the lower surrounding edges of the parking lot during intense rainfall events. The dry swales are a minimum of 10 feet wide and are composed of the same soil mix as the parking lot islands. “Raintanks” are 18 inches under the soil surface of the dry swale and serve as a reservoir for infiltration.

**Figure 2: Parking Lot Islands, Stone Sub-base with Choker Course, Prior to Placement of Open-Graded Asphalt Concrete Top Course**

**Figure 3: Dry Swale**

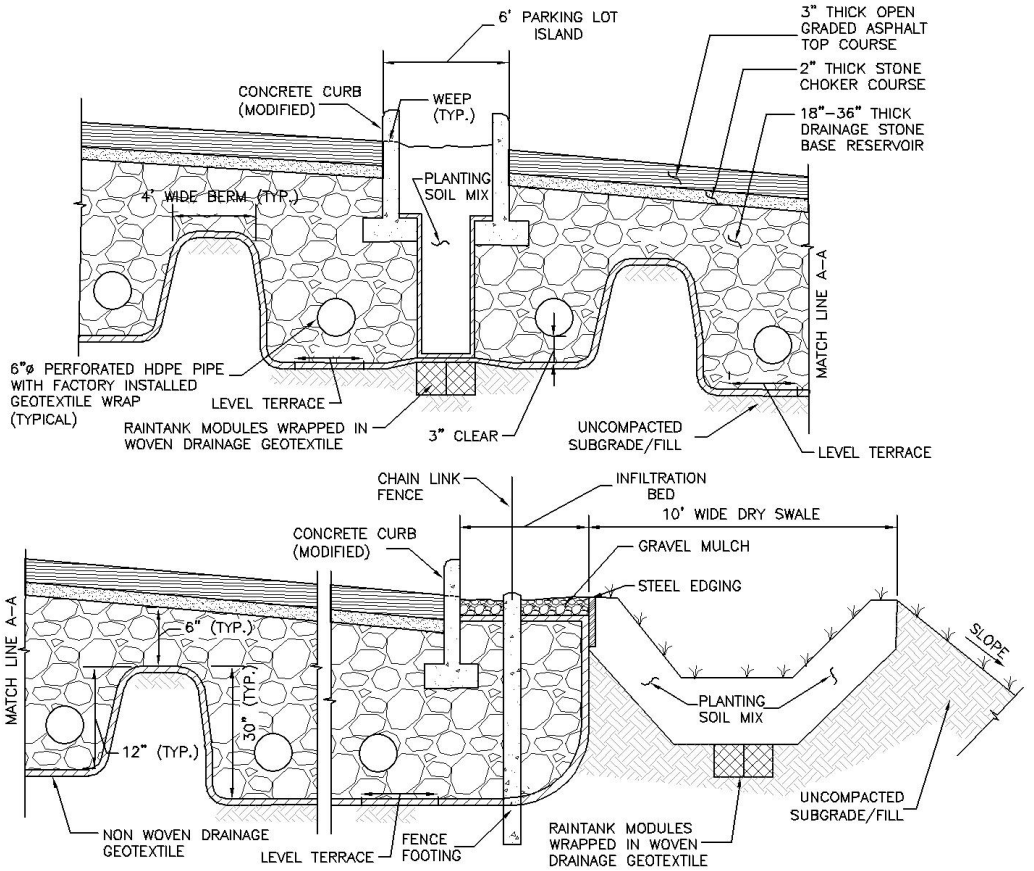



**Figure 4: Pavement, Terraces, Berms, Parking Lot Islands Dry Swales Details**

The irrigation system is carbon neutral and environmentally friendly. It reuses captured run-off and is powered by solar panels. The irrigation system waters the parking lot islands that contain trees and plants. The underground cistern is composed of a matrix of "raintanks". It is 12 feet (width) x 13.5 feet (length) x 4.3 feet (height) in dimension, and is located within the pervious pavement area below the stone base reservoir layer in the fill/sub-grade layer and holds approximately 5200 gallons. The bottom and sides are surrounded with a PVC liner to store the infiltrated run-off. The PVC liner extends beyond the cistern footprint to capture additional infiltrated run-off. The top of the cistern is surrounded with permeable geotextile and topped with sand. The cistern is connected via a 4-inch, solid cast iron pipe to a concrete sump. The sump contains the solar powered irrigation pump and related equipment. An irrigation controller adjacent to the sump connects several solar panels and precisely directs the system to deliver water when needed.

A conventional drainage system utilizing catch basins, manholes and concrete pipes was installed as a redundancy. Research from other agencies' projects suggests that there are a number of pervious pavement projects which have failed (Galli, J., 1992) and installing the conventional drainage during the initial construction would cost less than adding it in the future. The system is not connected to the existing storm drainage utilities beyond the parking lot facility, but rather drains into its own separate subdrainage piping system. The system may be used to monitor the effectiveness of the pervious pavement system. A visual inspection of the catch basins and associated pipes would indicate if there were issues with pavement surface permeability and/or sub-grade infiltration.

**Past Experience**

Prior to the hand over of the airport ownership within 2007, the PANYNJ planned for an overflow long-term Parking Lot "D". This was in response to the airport's need of additional parking for the upcoming Thanksgiving Holiday season. An existing 3-acre grassy site was paved to provide approximately 400 parking spaces. The parking lot was designed by the Engineering Department and constructed within a 2 month timeframe so it could be opened for use just prior to the Thanksgiving Holiday season. The tight time constraint for its construction, limited site space, and NYSDEC permitting were the driving factors in the design of Parking Lot "D" utilizing Infiltration and Bioretention for the BMP to address storm water run-off.

The pervious pavement of Parking Lot "D" used the same design elements as that of Parking Lot "A". However, the open-graded top course was an asphalt concrete drainage mix design typically utilized by the PANYNJ for airfield applications with low tire contact surfaces such as shoulders and erosion pavements. The Parking Lot "D" open-graded asphalt mix contained larger aggregate sizes and less asphalt binder content. After the first winter season, it was observed that the constructed Parking Lot "D" pavement surface started to ravel at the joints and slight segregation within the mat. In these areas, the stone aggregate could be pulled out by hand, demonstrating the lack of contact surface and asphalt adhesion. The beginnings and ends of several

asphalt pavement paving strips had exhibit signs of severe raveling and segregation which larger aggregate sized asphalt mixes are more prone to. In contrast, the open graded asphalt mix used for Parking Lot "A" had smaller aggregate and a tighter, more aesthetic surface finish while still providing the needed infiltration rate. Further evaluation of the Lot "D" pavement suggests that the asphalt binder either separated from the aggregate or the aggregate segregated in the hopper of the paving machine during the start-up and end paving operations. Hand raking the mix at the joints, also, may have lead to segregation. To prevent this from occurring in the Lot "A" expansion, strict material quality controls and structured paving procedures were enforced. In addition, a 12-inch wide emulsified asphalt slurry was applied on the surface of all paving joints to help mitigate raveling.

The dry swale for Parking Lot "A" was improved to store more run-off by using "raintanks" instead of the 6 inches diameter HDPE perforated pipe used in Parking Lot "D". The experiences and lessons learned from Parking Lot "D" were incorporated into the design and construction of the Parking Lot "A" expansion.

**Paving Observations**

Several interesting observations were made during the final asphalt concrete top course paving operations that are worth mentioning. Standard paving procedures and equipment were successfully used in constructing the pavement. A static roller should be used to smooth out the any ruts in the stone choker course prior to paving the top course. A portion of the stone choker course was lost in the voids of the stone base reservoir course.

The paving laborers raking and spreading the open-graded asphalt concrete by hand were asked to evaluate the material's workability as compared to conventional asphalt concrete. The laborers did not like working with the open-graded mix due to the high asphalt binder content. The asphalt binder along with the cellulose fibers would adhere to the hand tools, equipment, boots and clothing making everything sticky and somewhat messy.

The open-graded mix is viscous due to the high asphalt binder content. The open-graded asphalt top course should be allowed to cool down to a temperature of 250-260 degrees Fahrenheit before being rolled with a 10-ton static roller. Compaction efforts should be kept to a minimum three (3) passes with the 10-ton roller. More passes will not yield greater density and will result in moving the unsupported edge and thinning out the pavement. A finish roller should be used to remove roller marks.

The open-graded asphalt concrete retained heat for longer than conventional asphalt. This results in extended material working times. This phenomenon may be due to the air pockets in the open-graded asphalt acting as insulation.

**Discussion**

The site conditions posed a challenge to design a parking lot that would follow the surrounding elevations. The grade change from the existing lot to the corner of the expansion was 22 feet and did not readily lend itself to create a pervious pavement system. It required 12 terraces to capture segregate, and store the infiltrated run-off. The grading, depth of the stone base reservoir and number of terraces were designed to minimize cost and simplify construction of the pervious pavement system.

**Figure 5: Construction of the Terraces and Berms. The dark lines are the berms covered with geotextile fabric.**

Most publications suggest that the base of the stone base reservoir be placed on the in-situ sub-grade. The site was depressed and required fill to be transported from a nearby site. The airport was fortunate to have a borrow pit with the same soil characteristics as the existing sub-grade. The fill was placed with a minimum compaction effort to maintain its infiltration properties. Over time and under loading condition, this fill layer may further consolidate, and the parking lot may settle. To reduce the risk of settling, the stone bed reservoir was installed on the soil infill in late fall of 2009 and allowed to consolidate during the following winter and spring with placement of asphalt wearing surface within summer of 2010. The seasonal snowfall and rainfall aided in the initial settling. No signs of future consolidation have been observed.

The parking lot expansion opened to the public on October 18, 2010 in advance of Thanksgiving holiday travel season. The parking lot has been divided into short term, long term and employee parking. The Port Authority will provide a greater level of service to its patrons via convenience and the ability to choose parking areas with different rates close to airport's terminal area. The Port Authority will continually monitor the parking lot system's performance and any needed refinements will be incorporated in future designs and construction operations of pervious pavement systems.

## References

1. The Franklin Institute Research Laboratories (1972). “Investigation of Porous Pavement for Urban Run-off Control”, US Environmental Protection Agency, Washington, DC, p.15 para. 1

2. NJDOT (2007), “Section 902.03 Open-Graded Friction Course (OGFC)”, *Standard Specifications for Road and Bridge Construction*, New Jersey Department of Transportation, Trenton, NJ

3. Galli, .J. (1992), “Preliminary Analysis of the Performance and Longevity of Urban BMPs Installed In Prince George's County, Maryland”. Prepared for the Department of Natural Resources. Prince George's County, MD.

# The Case to Build End-around Taxiways at George Bush Intercontinental Airport for Air Quality Benefit.

**Michael T. McNerney and David Heinold**
AECOM, Fort Worth, Texas and Westford, Massachusetts

ABSTRACT

The traditional approach of reducing green house gas and air pollutant emissions is to reduce fuel consumption. However, the criteria for getting a taxiway or runway approved for FAA funding does not involve consideration of air emissions but only a reduction in delay. When the cost of reduced delay exceeds the cost of construction the project can meet a minimum threshold of benefit/cost ratio of 1.0 and, therefore, can be considered for FAA funding. The paper examines the case of the 2006 Houston George Bush Intercontinental Airport (IAH) master plan which recommended capacity enhancements to reduce delays by constructing two runways and end around taxiways.[1] However, the end around taxiway alone has been shown to substantially reduce delays and reduce fuel consumption. The FAA's Emission and Dispersion Modeling System (EDMS) was applied to characterize pollutant emissions with and without the improvements and the reduction of annual fuel consumption was used to estimate green house gases (GHG) reductions. The paper concludes that the reductions in pollutant emissions and GHG were significant and should be quantified to add to the benefit for prioritizing airport improvement projects competing for funding. The paper also suggests that in this particular case the reduction in criteria pollutants alone could be justification to begin construction of the end-around taxiway.

## INTRODUCTION

Based upon Airport Council International 2005 reported data, Houston George Bush Intercontinental Airport was the sixth busiest airport in the world in terms of aircraft operations.[2] In November 2004, George Bush Airport became only the third US airport to have three parallel simultaneous IFR landing runways. Currently, Denver International Airport is the only airport that has four parallel simultaneous and independent IFR landing runways although it does not operate in that mode as at least one parallel runway is used for departures. Dallas-Fort Worth International Airport has three parallel runways and one angled runway that can be use used for IFR landings. The results of an extensive simulation during the IAH alternative analysis and implementation planning phases of the master plan show that IAH will require a fourth simultaneous IFR landing runway when

annual operations reach 680,000, two new IFR runways when annual operations reach 780,000, and that end-around taxiways should be built as soon as possible.

An Environmental Impact Statement (EIS) was initiated in May 2009 with scoping meetings to provide public comment on the implementation of the projects that includes the two new runways. The project scoping phase of the EIS was completed and the Scoping Report was published August 2009 and is available on the project web site.[3] The current status of the EIS project is that a Draft EIS is scheduled to take two years to complete and will look at four possible alternative locations for two new runways as is shown in Figure 1.

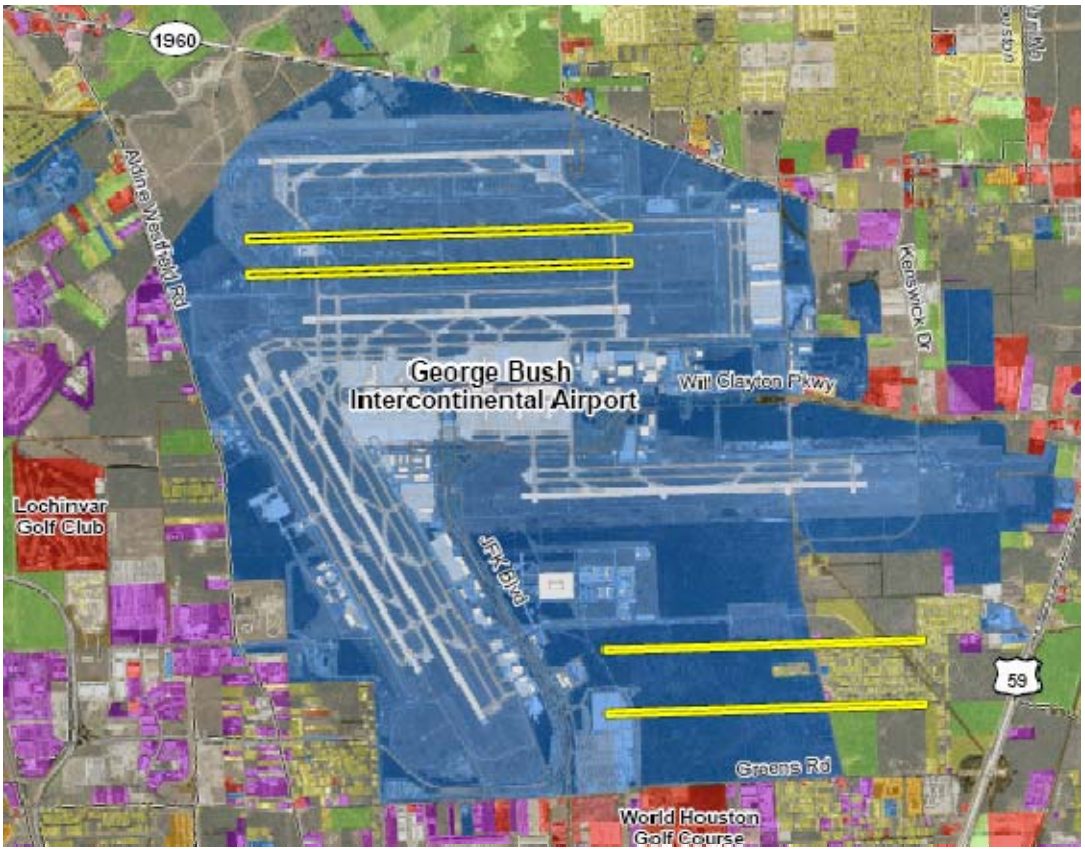


Figure 1. IAH EIS Project Scope Showing Locations of Two Proposed New Runways

Adding new air carrier runways always requires an EIS. A particularly challenging air quality issue is that the Houston metropolitan statistical area is in severe non attainment for ozone. However, if the proposed action includes new taxiways that reduce delays and associated emissions of air quality pollutants and ozone precursors these beneficial improvements may not trigger the need for an EIS from an air quality perspective and, therefore, may not need to be delayed until after the anticipated 3-4 years required for a runway environmental review.

This paper is a follow up to two previous papers. A paper was presented and published in 2007 at the 29th International Air Transportation Conference that showed the delay improvement results of the simulation modeling of capacity improvements from constructing the new runways and from constructing the end around taxiways.[4] The second paper was presented and published at the 2009 Air & Waste Management Association Annual Conference that showed the modeling of the air quality reductions of criteria pollutants and reductions in green house gases emissions that would result both from the construction of the end around taxiways alone and the construction of the new runways.[5] The first paper makes the case for the construction of the end around taxiways for capacity improvements and second paper computes the benefits gained from reductions in emission of criteria pollutants and green house gases. The premise of this paper is to bring attention to the benefit of considering funding taxiway improvements based upon the air quality improvements alone.

## CURRENT FAA POLICY FOR CAPACITY IMPROVEMENTS

The FAA has a process that has been in place for many years in order to add new runways and taxiways to improve capacity. We have been fortunate in the last decade in that forecast demand for airport capacity has not risen on the scale that was forecast at the beginning of the decade. Unfortunately, this delayed growth has been the result of three catastrophic events that have reduced demand severely. Those catastrophic events were the aftershock of the events of September 11, 2001, the fuel cost crisis of 2006, and the economic recession of 2008. Each of these three events has caused a significant turn down in the demand for commercial airline flight and for capacity improvements. However, before the events of 9-11, there was great nationwide concern for a need for capacity improvements. When the demand for air travel and runway capacity is again high, we will find ourselves with a priority to reduce delays by adding more capacity at our large and medium sized airports.

The FAA process to get a new taxiway or runway constructed with FAA funding must complete the following steps but not necessarily in this order:

1. Project must be on the FAA approved master plan
2. The airport must perform a benefit cost analysis (BCA)
3. Airport submits funding requests to FAA prioritizing projects for Airport Improvement Project (AIP) or Passenger Facility Charge (PFC) funds.
4. If required, a NEPA process must be completed
5. The project is designed
6. The project is released for bids

7. The project is managed and constructed.

The FAA policy that has been in place for many years is designed to give priority to those new runways and taxiways that have a compelling reduction in delays over projects that desired but not enough delay reduction to justify committing FAA funding. In actual practice, the sole benefit in the required benefit cost analysis to have a project move forward is to have the value of the delay reductions converted to dollars of user costs have a greater value than the cost of construction; in other words a benefit cost ratio of greater than 1.

The FAA provides guidance on the preparation of the benefit cost analysis (BCA) in a 128 page document available on their web site.[6] In the guidance, it states that the benefits that result from capacity-related airport projects and other initiatives will largely consist of cost savings to current and future airport users associated with reduced time spent in the airport system. Reduced time in system may take the form of reduced delay, more efficient processing, or reduced idle time. In addition, time and cost savings can result from the ability of a facility to accommodate more efficient aircraft.

Capacity-related projects may also contribute to the ability of an airport to process more operations and passengers, greater safety and security, reduced environmental impacts, greater comfort for travelers, and other benefits. Note that projects intended primarily to meet objectives of safety, security, design standards, and environmental objectives are not subject to BCA requirements. Airside capacity projects are intended principally to reduce airside delay, improve aircraft processing efficiency, improve predictability of landing and take-off schedules, and/or to accommodate larger, heavier, longer-range aircraft at the airport. Other benefits of airfield capacity projects may include noise mitigation, reduced aircraft emissions, and compliance with FAA standards for airport safety, security, and design.

In summary, FAA practice is that although reduction of aircraft emissions can be a substantial benefit of adding a new taxiway, it is treated as primarily as a capacity reduction tool that is factored into the standard benefit cost analysis. In practice, although the benefit of air quality emission reductions can be quantified, all projects are judged by the dollar benefits of actual delay reductions. In most cases the delay reductions are the driving reason for building new taxiways. However, in the case of the IAH end-around taxiways, the emission reductions are especially compelling because the airport is in a severe ozone non-attainment area. Delays in funding and approval will delay these beneficial air quality improvements, Thus, a process should be instituted that the end-around taxiways are immediately brought forward for construction based upon the likelihood that would materially reduce emissions. When the current BCA guidance was adopted in 1999, green house gases were not an international issue and the geographical extent of ozone nonattainment areas was much less than for the proposed range of the revised ambient ozone standard. In addition to ozone, a short-term standard has been

promulgated for nitrogen dioxide, and lower $PM_{2.5}$ and CO standards are likely. Thus, emissions both of criteria pollutants and green house gases are one of more important issues facing the aviation industry.

## PROJECT VALE

The FAA does have a program specifically designed to help fund projects at airports to help reduce airport ground emissions. Voluntary Airport Low Emissions Program (VALE) is a national program to reduce airport ground emissions at commercial service airports located in designated air quality nonattainment and maintenance areas. The program was established under the *Vision 100* Century of Aviation Reauthorization Act of 2003 (P.L. 108-176). The VALE Program allows airport sponsors to use the Airport Improvement Program (AIP) and Passenger Facility Charges (PFCs) to finance low emission vehicles, refueling and recharging stations, gate electrification, and other airport air quality improvements.

However, the program has been used primarily to add hybrid vehicles and gate improvements at airports. The program could possibly be used as vehicle to fund a new taxiway, but that is not what the program was initially envisioned to do and it has never been used to justify a project that reduces aircraft delays. The City of Houston has received funds for 32 hybrid vehicles under the VALE program. However, a program of greater scope is needed than the VALE program that can be applied to reduction of aircraft emissions not just emissions of ground vehicles.

### CASE STUDY IAH END-AROUND TAXIWAYS

Although the case study of IAH end-Around taxiways may not be entirely unique, there are some compelling reasons for beginning design and construction as soon as possible. One reason for not using the slow normal process is that these emission reductions can be taken as soon as the taxiway is constructed, and every day of delay is adding to the green house gases and precursors to ozone formation at the airport.

The circumstances and conditions at IAH that help make the case so compelling are summarized as follows:

- The airport is one of the top 10 busiest in the world
- The airport is in a severe non-attainment area for ozone
- The amount of delay is still highly significant at current levels even with a 9% reduction of operations due to the economy.

- Much of the aircraft delay can be attributed to departure delay which occurs with aircraft waiting in queue for departure.
- The end –around taxiway can significantly reduce aircraft delay for taxi-in and taxi out times by creating additional departure queuing spacing for aircraft to taxi.
- The airport is a major airline hub airport with significant domestic and international growth potential and has shown large jumps in operations in this decade after the 2001 slump.
- The City of Houston's economy is tied to petroleum industry and has been steady growing.
- Even if the new runways in the current EIS were not constructed, the end around taxiway would make a significant contribution to delay reduction.
- The amount of delay reduction for aircraft on the ground for the end-around taxiway is comparable in the first five years to the amount of delay that would be achieved from building one new runway.
- If a benefit cost analysis were prepared for the end around taxiway alone it would probably exceed a benefit cost ratio of 5:1.

## CALCULATION OF DELAY

Although, *SIMMOD*™ (FAA's Airport and Airspace Simulation Model) simulation has shown in the case of IAH, that adding two new end-around taxiways will reduce the annual aircraft arrival and departure delays on a per aircraft basis, the operation does increase the actual aircraft taxiing distance. The modeling results show that the increased taxi distance, is offset with the delay reduction and congestion reduction in the terminal apron to significantly reduce emissions.

At the start of the Master Plan in May of 2004, the airport's fifth runway was under final stages of construction, which when completed provided three parallel and independent runways that could be used for simultaneous landings in poor weather. With IAH expected to achieve triple Instrument Flight Rules (IFR) landings this five-runway capacity was a baseline for future analysis.

Figure 2 shows the 20-year forecast of annual operations developed for the master plan which topped out at 880,000 annual operations at the year 2025 and 780,000 annual operations would be reached in the year 2020. After the Master plan was completed, the more optimistic FAA 2005 forecast (TAF) for IAH indicated that 780,000 annual operations would be reached by the year 2015. After three years of over 30,000 annual operations of growth from 2003 to 2006, the airport experienced a leveling of growth for 2007. The Actual Traffic data shown in Figure 2 shows that the 2005 and 2007 FAA

forecasts did not predict the decline in traffic from the 2006 fuel crisis and the 2008 economic recession. The 2009FAA forecast is also shown in Figure 2 and predicts reaching 780,000 operations in 2023 or about 3 years after the master plan projections.

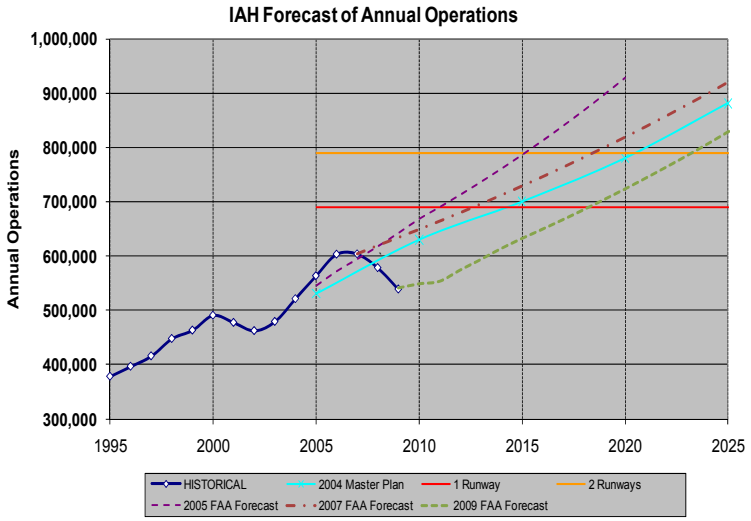
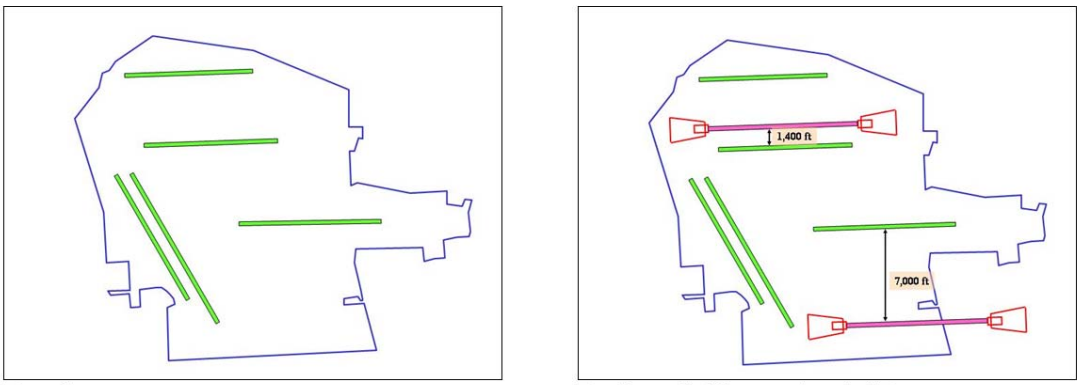



Figure 2 IAH Forecasts of Annual Operations

Regardless of which forecast is used, the conclusion is that IAH will reach 780,000 annual operations between 2020 and 2023. Because nearly all of the traffic at IAH comes from one home based airline, it is entirely conceivable that traffic can return to 2006 levels without the delay of having to purchase new aircraft once the economy rebounds in a very short time.

## RUNWAY REQUIREMENTS

A conclusion of the 2006 Master Plan was that two new runways were needed. Figure 3 shows the current baseline configuration of the airport with three east –west runways and two runways used for departures to the southeast. The preferred alternative was 1.1B with a new parallel south arrival runway and a new 12,000 foot departure runway 1,400 feet north of Runway 8R-26L. The only difference between the preferred alternative for the master plan and the alternatives in Figure 1 in the EIS is the spacing of the parallel runways.




Baseline Preferred Alternative 1.1B

**Figure 3 IAH Master Plan Runway Configurations**

The arrival runway is needed at 680,000 annual operations to keep the average arrival delay within the goal of 3 minutes per aircraft. The second runway a departure runway is needed at 780,000 annual operations to keep the average delay per aircraft under the goal of 6 minutes per aircraft.

The FAA's Airport and Airspace Simulation Model, *SIMMOD Plus*™, was used for the detailed simulation analyses. Under development by the FAA over more than 20 years, *SIMMOD*™ has been used worldwide for a variety of airfield and airspace capacity/delay studies. Output is provided for each aircraft such that hourly throughput, travel times, and delay times are easily obtained.

## AIRSIDE IMPLEMENTATION ANALYSIS

Goals for the airside simulation (performed by TransSolutions at the direction of AECOM) were established for annual average arrival delay, annual average departure delay, and annual average delay per aircraft operation. Being a hub airport, the goals for annual average delays was set for three minutes for arrival, six minutes for departure, and four minutes per average operation. Simulation runs were performed in east- and west-flow and under IFR and VFR traffic conditions.

Also simulated was the end-around taxiways designed to allow aircraft to taxi around existing Runway 8R-26L as shown as new taxiways in Figure 4. These end-around taxiways would be the primary departure routes for future departure Runway 8C-26C and the primary taxi-in route for existing Runway 8L-26R. Because of current congestion for departure queues, the terminal apron becomes clogged during departure peaks which add to the time for arriving aircraft to taxi to the gates.

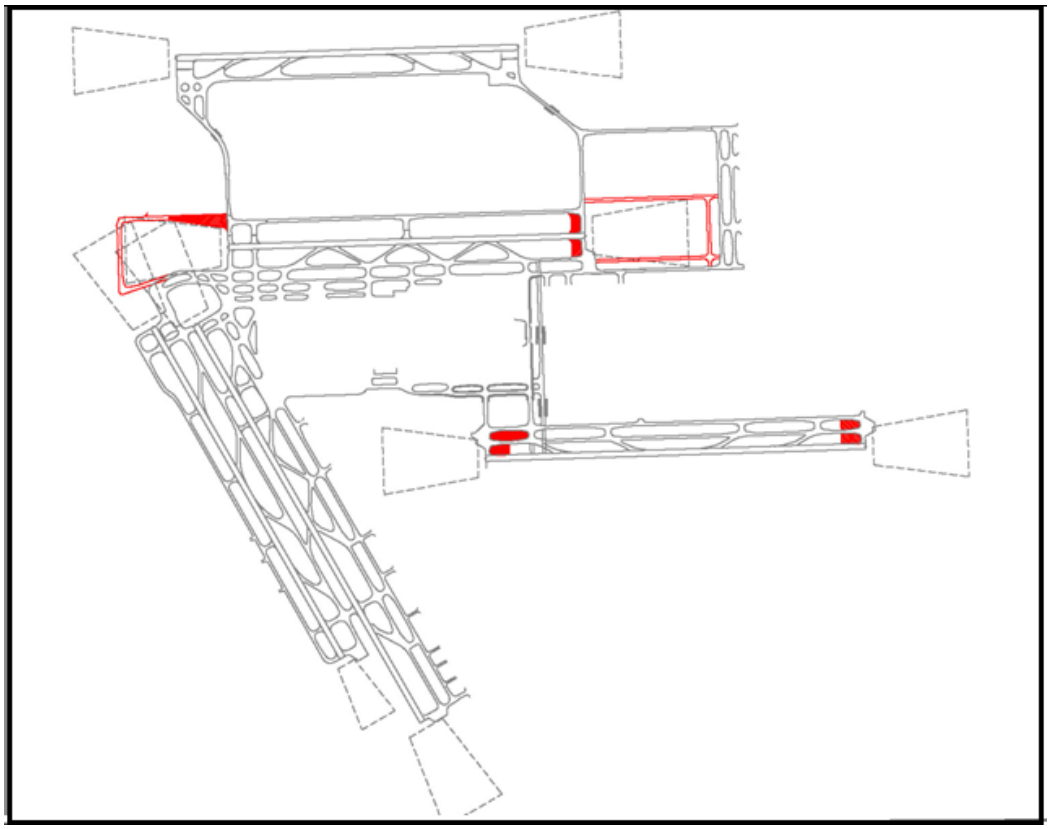


Figure 4 Master Plan Proposed End-around Taxiways around Existing Runway 8R-26L

To annualize delay, one must take into account the annual percentage of time in IFR versus VFR operations and in east- and west-flow. Currently, west-flow capacity is higher, so the FAA tower controllers prefer to operate the in west flow, even if a 5-10 knot tailwind exists. Wind analysis shows that an east-flow is the more favorable wind condition throughout the year. With the new layout, the highest capacity configuration may be different, so the annualized delays were calculated for operating in east during calm winds (up to 6 knots of tailwind) and operating in west during calm winds. Therefore, in Table 3 results are calculated for maximum east and maximum west for comparison purposes. In the overall analysis, the average of the two conditions is a good approximation.

**Table 3 Annualized Delay in Minutes for Implementation Planning**

| | 2010 | 2015 | 2020 | 2025 | 2010 | 2015 | 2020 | 2025 |
|---|---|---|---|---|---|---|---|---|
| **Average Annualized Delay Per Flight** | **Max East Configuration** | | | | **Max West Configuration** | | | |
| Baseline | 4.16 | 6.95 | 11.8 | 28.03 | 4.03 | 6.52 | 12.46 | 26.41 |
| End-around Taxiways | 3.37 | 4.23 | 6.14 | 21.63 | 3.32 | 4.28 | 5.97 | 20.89 |
| Arrival RWY | 2.23 | 3.02 | 4.6 | 8.7 | 2.18 | 3 | 4.53 | 10.13 |
| Departure RWY | 1.89 | 2.62 | 4.74 | 8.54 | 1.89 | 2.61 | 4.71 | 9.94 |
| Alternative 1.1B (two runways) | | | 2.56 | 4.12 | | | 2.95 | 4.07 |
| **Average Annualized Arrival Delay** | **Max East Configuration** | | | | **Max West Configuration** | | | |
| Baseline | 2.54 | 3.14 | 7 | 21.49 | 2.78 | 3.47 | 6.76 | 18.72 |
| End-around Taxiways | 2.43 | 2.98 | 3.91 | 6.81 | 2.46 | 2.98 | 3.97 | 7.06 |
| Arrival RWY | 1.4 | 1.7 | 1.97 | 2.56 | 1.22 | 1.46 | 1.8 | 2.35 |
| Departure RWY | 2.42 | 3.15 | 3.99 | 7.71 | 2.29 | 2.92 | 3.8 | 7.22 |
| Alternative 1.1B (two runways) | | | 1.7 | 4.08 | | | 1.7 | 3.36 |
| **Average Annualized Departure Delay** | **Max East Configuration** | | | | **Max West Configuration** | | | |
| Baseline | 5.77 | 10.76 | 16.6 | 34.56 | 5.28 | 9.56 | 18.15 | 34.11 |
| End-around Taxiways | 4.31 | 5.48 | 8.37 | 36.45 | 4.18 | 5.57 | 7.98 | 34.71 |
| Arrival RWY | 3.06 | 4.35 | 7.24 | 14.84 | 3.15 | 4.55 | 7.26 | 17.91 |
| Departure RWY | 1.37 | 2.09 | 5.5 | 9.36 | 1.48 | 2.31 | 5.62 | 12.67 |
| Alternative 1.1B (two runways) | | | 3.42 | 4.15 | | | 4.2 | 4.78 |

The most surprising result was in comparing the Baseline results to the end-around taxiways by adding only end-around taxiways but without a new runway. In the Baseline condition, the FAA tower does not favor using the new north Runway 8L-26R for departures because of neighborhood noise concerns and because of the concern that dedicating the runway to arrivals only would provide the most capacity. However, in the analysis of the condition with end-around taxiways, by adding a few departures on the north runway and using the end-around taxiways as the queuing path for aircraft departing on Runway 8R-26L, the analysis indicated a significant reduction in departure and overall delay with a small reduction in arrival delay. The reduction in delay was very significant in 2015 and 2020 but does not continue to 2025 without adding new runways.

The simulated difference in departure delay in the year 2020 (780,000 annual operations) from 17 to 8 minutes was very dramatic. Given these large reductions, the factors that contribute to the effectiveness of these minor taxiway modifications was investigated. By reviewing the animations of taxi-out patterns from the simulations it was found that without the end-around taxiways, the taxiways adjacent to the terminal are used as queuing space and aircraft arriving to the terminal actually have to cross Runway 8R-26L at midfield to minimize the congestion. With end-around taxiways, the aircraft queue on the end-around taxiways and the arriving aircraft taxi-in around Runway 8R-26L and the terminal area taxiways are clear of congestion. This explains why the end-around taxis are so effective.

## ASSOCIATED EMISSIONS REDUCTIONS

The aircraft fleet mix that was calculated for the master plan forecast of 780,000 annual operations (year 2020) was used in the emissions calculations. The end-around taxiways

provides a significant reduction in departure delay and taxi-out times that is greater in 2020 than 2015. The aircraft fleet mix of the future will change and future aircraft will be more fuel efficient, however, the only actual data for the modeling in EDMS version 5.0 is from actual engines on actual aircraft in use today. EDMS uses the landing and takeoff (LTO) cycles for calculating emissions, which is half of the number of annual operations (1 take off and 1 landing equals 2 operations).

Taxi-in time is the average time that aircraft take from exiting the runway to arriving at the gate. Taxi-in time does include arrival delay as calculated by *SIMMOD*™ as all arrival delay is airborne delay waiting for runway availability. Taxi-out time is all the time that aircraft take from departing the gate until taking the runway for takeoff. Taxi-out time does include all departure delay as well as the time aircraft are taxiing without delay. Weather delays are not included in the *SIMMOD*™ analysis. Table 4 shows the annualized average taxi-in and taxi-out times for the baseline case and the end-around taxiway case.

**Table 4 Annual Average Taxi-In and Taxi-Out Times**

| | Baseline Case | End-around Taxiway Case |
|---|---|---|
| Annualized Average Taxi-In Time | 9.54 minutes | 8.54 minutes |
| Annualized Average Taxi-Out Time | 30.94 minutes | 17.77 minutes |

EDMS is the EPA approved FAA emissions dispersion modeling system to determine the total emissions at the airport. It includes the total emissions from the aircraft during a takeoff and landing cycle. It also calculates the time of use and emissions from all ground support equipment (GSE) used to service the aircraft and auxiliary power units (APU) used to provide power to the aircraft when the main engines are not running. For this comparative analysis the EDMS default values of GSE and APU were used because they would be identical in both cases. The default values are mostly gasoline powered tugs, trucks, baggage carts and the user has to select the more modern equipment that is being used at most airports today to reduce emissions.

The only significant input difference in the two cases in calculating the emissions is the annualized average taxi-in and taxi-out times shown in Table 5. A more detailed dispersion analysis with EDMS is possible what would show the location of geographic areas of high concentrations of pollutants. However, the dispersion analysis requires additional simulation analyses to obtain the taxi paths of each aircraft in the simulation.

The results of the emissions reductions are shown in Table 6 and show that the reduced taxi-out time of the end-around taxiway results in a significant reduction in the criteria

pollutants of carbon monoxide (CO), total hydrocarbons (THC), non methane hydrocarbons (NMHC), volatile organic compounds (VOC), nitrogen oxides ($NO_x$), and sulfur oxides ($SO_x$). EDMS also calculates the particulate matter emissions of $PM_{2.5}$ and $PM_{10}$, but those are mostly during high power settings such as takeoff and they were calculated the same in EDMS 5.0 for both cases.

Carbon monoxide had the largest percentage reduction at 33.4%. Carbon monoxide is a poisonous gas which is a byproduct of incomplete combustion. Aircraft taxiing for departure are using low power settings and are more likely to have incomplete combustion; therefore, the percent reduction in CO was greater than the percent reduction in fuel consumed. Conversely, oxides of nitrogen are produced more at higher power settings and therefore had only 7.4% reduction in emissions with the end-around taxiways.

**Table 5 Emissions Calculated at 780,000 annual Operations**

| Pollutants | Unit | Baseline | End-around Taxiway | Reduction | Percent Reduction |
|---|---|---|---|---|---|
| CO | tons | 6,379 | 4,247 | 2,131 | 33.40% |
| THC | tons | 1,054 | 831 | 222 | 21.10% |
| NMHC | tons | 1,054 | 831 | 222 | 21.10% |
| VOC | tons | 1,140 | 900 | 240 | 21.10% |
| NOx | tons | 3,917 | 3,629 | 288 | 7.40% |
| SOx | tons | 516 | 416 | 99 | 19.30% |
| | | | | | |
| **Fuel Consumed** | tons | 78,198 | 63,140 | 15,058 | 19.30% |
| | | | | | |
| **Greenhouse Gases** | tons $CO_2$ | 244,874 | 197,720 | 47,154 | 19.30% |

EDMS version 5.0 does not directly calculate greenhouse gases (GHG), but it does calculate the kilograms of fuel consumed. For simplification purposes we have assumed all the fuel was kerosene-based jet fuel, standard Jet A fuel. Carbon dioxide is the largest greenhouse gas emission from jet aircraft and combustion of one gallon of jet fuel results in a release of 21.1 pounds of carbon dioxide.[7] From the fuel consumed and based on an average density of 6.74 pounds per gallon the amount of carbon dioxide was estimated and is shown in Table 5.

## DISCUSSION

It is sometimes difficult to characterize the magnitude of emission reductions in terms of direct environmental benefit or monetary value. Greenhouse gases are receiving

increasing attention at the present time especially in aviation, with the European Union having proposed mandatory emission offsets. The current administration has discussed additional taxes for green house gas emissions. In addition, the unrestricted climb of the price of Jet fuel as experienced in 2008 has given rise to interest in efficiency in operations and fuel consumption. Reducing the consumption of jet fuel provides a proportional reduction in carbon dioxide as the major greenhouse gas. Fuel savings strategies usually (depending on engine power settings) lead to reductions in criteria pollutants and hydrocarbons as modeled in EDMS.

The annual reduction of 47,000 tons of $CO_2$ is significant. The annual reduction of 2,000 tons of CO, or 200 tons of VOC or $NO_x$ may not be large in terms of the regional emissions inventory, but are substantial in terms of overall airport operations. To put the effects of the aircraft emissions reductions from constructing this taxiway into perspective, a comparison has been made with estimated emissions from ground support equipment (GSE) and auxiliary power units (APU). The EDMS version 5.0 emission estimates for default GSE and APU usage for the 780,000 annual operations of the fleet mix were modeled. Without performing a refined GSE analysis at IAH, the default values in EDMS version 5.0 assume rather large usage and mostly diesel or gasoline powered equipment.

Airlines at airports in ozone non attainment areas are spending millions of dollars to try to achieve a 75% reduction in GSE emissions through electrification of GSE. The results of this analysis show that the use of the end-around taxiway for Runway 8R-26L at IAH will provide an even greater reduction in emissions than if all GSE were converted to zero emissions equipment. Thus, the reductions achieved by the end-around taxiway are significant.

At the level of 780,000 annual operations, the savings of fuel consumed is over 5 million gallons of jet fuel in a year. At $5.00 per gallon that is over $25 million in fuel savings alone. Because the savings don't begin until the end-around taxiways are constructed, this analysis provides additional justification to begin construction of the end-around taxiways at IAH without waiting for the EIS to be completed on the two new runways.

## CONCLUSIONS

FAA generally provides funding for capacity enhancement projects to airports that are justified based upon benefit cost where the benefit is measured solely upon reduction in delay. However, this analysis demonstrates for IAH that the end-around taxiway not only reduces delays, it also provides a real benefit in reduction of emissions of criteria

pollutants as well as emissions of carbon dioxide, a greenhouse gas that contributes to global warming. The criteria pollutant and greenhouse gas emission reductions of the proposed IAH end-around taxiways are significant and should be taken under advisement as a carbon offset or reduction program. Because emission reductions occur only when the taxiway is completed, it should be completed without delay and should not be held up waiting for the EIS and record of decision to be completed. The sooner it is completed the sooner reductions in fuel consumed can be realized and emissions reduced.

## REFERENCES

[1] DMJM Aviation and RS&H, 2006 Master Plan George Bush Intercontinental Airport, Houston Airport System, July 2006

[2] Airport Council International – North America, 2006 2005 North America Final Traffic Report: Total Movements, www.aci-na.org/asp/traffic.asp?art=217, Accessed July 24, 2006.

[3] Scoping Report: George Bush Intercontinental Airport Environmental Impact Statement, FAA, August 11, 2009 available at www.iah-eis.org.

[4] McNerney, Michael T. and Belinda Hargrove, 2007 "Simulation of Quadruple IFR Arrival Runways and End-Around Taxiways at George Bush Intercontinental Airport", Aviation: A World of Growth, Proceedings of the 29th International Air Transport Conference, Irving, Texas, August 19-22, , ASCE, ISBN 13 # 978-0- 7844-0938-1, pp 108-121.

[5] Heinold, David, Michael T. McNerney, and Belinda Hargrove, "Modeling Of Air Quality Improvements from Constructing End-around Taxiways at George Bush Intercontinental Airport," Proceedings of the 102nd Air & Waste Management Association Annual Conference & Exhibition, , Detroit Michigan, June 16-19, 2009, Paper ET-1e.

[6] FAA Airport Benefit-Cost Analysis Guidance, FAA Office of Aviation Policy and Plans, December 15, 1999.

[7] University of Wisconsin, Energy Education Program accessed on July 29, 2008 at http://www.uwsp.edu/cnr/WCEE/keep/mod1/Flow/Greenhouse_Gas.htm.

# Meeting the Expanding Needs of a Metropolitan Area: the Move from General Aviation to Commercial Aviation

Saty D. Satyamurti, Ph D., P.E., Consultant[1], Antonio Massidda, B.S. Civil Engineering, Visiting Researcher[1], Stephen Mattingly, Ph.D., Associate Professor[1], Mike Shahan, M.S., Airport Director, William A. Retz, M.S., Marketing Director[2]

[1]Civil Engineering Department, University of Texas at Arlington, P O Box 19308, Arlington, TX, 76019
[2]North Texas Regional Airport, 4700 Airport Drive, Denison, TX, 75020

## Abstract

The Dallas/Fort Worth Metropolitan Statistical Area (MSA) is the nation's fourth largest urban area, with an estimated population of about 6.5 million people. This area is served by two commercial airports: DFW International (DFW) and Dallas Love Field (DAL). Those airports also serve the residents of the counties surrounding the limits of that MSA; in fact, to the North (where the population growth is concentrated), no commercial airports are available until Oklahoma City, Will Rogers World Airport (OKC). The North Texas Regional Airport (GYI) is located midway; this airport is currently undergoing major improvements, and with the increase in demand in the near future it could play an important role in the air transportation system of the region. This paper outlines the possibility that GYI can meet a portion of the air travel demand from the northern boundaries of the expanding DFW MSA, and from "Red River" counties along the Texas-Oklahoma border. For population living in this region, an expanded GYI would be the ideal airport among those in the DFW area or OKC. The authors define the GYI Primary Airport Service Region (ASR) as the territory within a 90-minutes drive time from the airport. The trip rate (TR) the number of round trip flights per region resident in the ASR is then estimated using the DFW MSA trip rate with a reduced number to reflect the economic activity in the GYI ASR; this leads to a forecast of the potential GYI passenger activity for the next twenty years. The recommended course of action to move from a general aviation to commercial aviation is indicated into three distinct phases, short, medium and long term. Phase 1, introducing cargo activities, Phase 2, begin non scheduled operations, and Phase 3, scheduled operations. The method applied to GYI study could be used to analyze similar general aviation airports that exist in other metropolitan regions.

Keywords; DFW, GYI, DAL, airport, air cargo, general aviation, commercial aviation, FAA, runway, population forecast, MapPoint, planning

## Introduction

The population and business activities in the Dallas/Fort Worth metropolitan region continue to grow and expand further away from its two commercial airports, Dallas/Fort Worth International (DFW) and Dallas Love Field (DAL). This paper analyzes the air transportation activity in the DFW metropolitan area and explores the opportunity for initiating commercial operations at the North Texas Regional Airport (GYI) to expand business opportunities for the region.

**The North Texas Regional Airport (GYI)**

*Location and history*

GYI is located in Denison, Texas, 75 miles North of Dallas and just four miles south of the Texas/Oklahoma border, as shown on Figure 1. The airport was first named "Grayson County Airport" and served as a military field (Perrin AFB) during the last world war, until the Air Force ceased operations in 1971 and the airport was returned to the Grayson County, who currently owns it.


**Figure 1: Road Map**

*Access roads*

As shown on Figure 1, access roads to GYI are:

From Dallas and Fort Worth:

- US-75; or,
- State Highway 289 (SH-289)

From the southern Oklahoma:

- US-35 and US-82 from Oklahoma City; or,
- US-69 and US-62 from Tulsa

*General Information*

GYI is classified in the Federal Aviation Administration's (FAA) National Plan of Integrated Airport Systems as a general aviation airport. The control tower was recently renovated and operates during daylight hours. The main runway, 17L-35R (9000 x 150 ft) is currently undergoing major rehabilitation, its reopening is scheduled for November 2010. Taxiway A is currently used as runway 17L-35R. Runway 13-31 (2277 x 60 ft) is open and its pavement is in fair condition and runway 17R-35L (8000 x 150 ft) is closed. Later it could be rehabilitated in order to operate with two runways, if future needs require such an upgrade. All taxiways have been recently upgraded and paved, they comply with FAA design standards category IV/V aircraft (ex. B767); their signs and pavement markings are in good condition. Instrument approach procedures include: for runway 17L, Instrument Landing System (ILS), GPS and NDB, for runway 35R, VOR/DME. On the east side of the airport, a terminal building, administrative offices and parking area are located. The airport terminal is open 24/7, providing a lounge, restrooms, a vending area and a

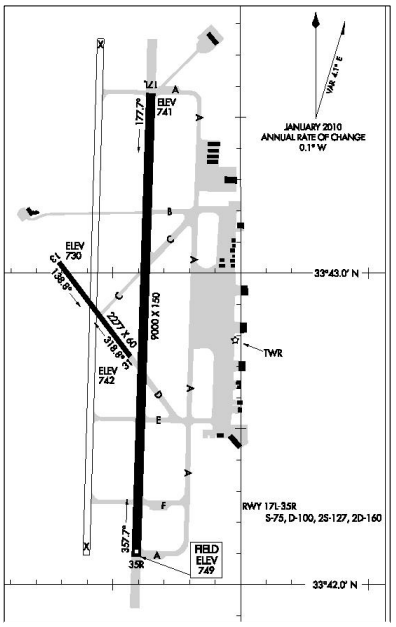



**Figure 2: GYI diagram**

flight planning board to its general aviation users. A Fixed Base Operator is located at the airfield; Avgas (100LL) for piston engine aircraft and Jet fuel are available.

As of October 2010, 180 general aviation aircraft are based at GYI, from small single and multi-engine piston to large turboprops and jets. For the twelve month period ending on April 30, 2007, there was an average of 50,000 operations at GYI. No cargo flights are currently in operation. Occasionally, commercial aircraft such as the Boeing 737, lands at the airport for periodic maintenance in the hangar facility located on the east side of the field.

*Future expansion*

GYI is already a good general aviation airport; to upgrade to commercial operations, the airside will not require major improvements. However, a passenger terminal, a fuel storage facility and adequate parking had to be built. While all these upgrades have been rigorously planned by the airport management, they are waiting for federal and state funding.

**The Airport Service Region (ASR)**

To identify the number of people that may choose GYI as their primary airport, it is useful to define an Airport Service Region (ASR). The software Microsoft MapPoint2010 allows the researchers to categorize the area within a 90-minute drive time line from GYI. Another 90-minute drive time line is drawn in reference to DFW (with software settings assuming average road and traffic conditions). The two areas intersect; therefore, the line where the drive time to GYI and DFW is the same, located by interpolation. This is defined to set the boundaries for the ASRs. Therefore, the Primary ASR of GYI is considered to be the area depicted by its 90-minute drive time line and the line of equal drive time to/from DFW, as shown in Figure 3. Thus, from each point inside the Primary ASR, the travel time to GYI is lower or equal to the drive time to DFW.


**Figure 3: DFW and GYI Airport Service Region**

In addition to the Primary ASR, this analysis also includes some population residing beyond the 90-minute boundary, because GYI is still the airport with the shortest drive time from those towns. Using a thorough analysis of employment and income statistics for the area, the study forecasts the number of potential users of this airport once it is opened to commercial traffic. The population data by County for the ASR is shown on Table 1.

The Dallas-Fort Worth-Arlington MSA is made up of eight counties in the Dallas-Plano-Irving Metropolitan Division (Collin, Dallas, Delta, Denton, Ellis, Hunt, Kaufman, and Rockwall) and four counties in the Fort Worth-Arlington Metropolitan Division (Johnson, Parker, Tarrant, and Wise).

Table 1: Population data for Primary Airport Service Region

| STATE | COUNTY | Population 2000 (US Census) | Estimated Population 2009 (US Census) | Estimated difference 2000-2009 | Forecast | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | 2010 | 2015 | 2020 | 2025 | 2030 |
| Texas | Grayson | 110,595 | 120,030 | 8.5% | 118,011 | 121,810 | 124,953 | 127,551 | 129,763 |
| | Cooke | 36,365 | 38,650 | 6.3% | 39,412 | 41,106 | 42,668 | 43,943 | 44,967 |
| | Fannin | 31,242 | 32,998 | 5.6% | 33,558 | 34,902 | 26,186 | 37,270 | 38,164 |
| | Hunt | 76,596 | 82,831 | 8.1% | 91,928 | 100,168 | 108,413 | 116,545 | 124,646 |
| | Lamar | 48,499 | 48,965 | 1.0% | 50,362 | 51,132 | 51,529 | 51,634 | 51,318 |
| | Collin* | 180,000 | 228,600 | 27.0% | 235,000 | 266,725 | 305,750 | 351,224 | 405,740 |
| Oklahoma | Marshall | 13,184 | 15,014 | 13.9% | 16,900 | 18,900 | 20,800 | 22,800 | 24,800 |
| | Atoka | 13,879 | 14,500 | 4.5% | 15,500 | 16,500 | 17,400 | 18,300 | 19,300 |
| | Bryan | 36,534 | 40,800 | 11.6% | 40,700 | 42,800 | 44,900 | 47,000 | 49,200 |
| | Pittsburg | 43,953 | 45,211 | 2.9% | 45,300 | 46,300 | 47,400 | 48,300 | 49,200 |
| | Carter | 45,619 | 48,326 | 5.9% | NA | 48,700 | 50,200 | 51,500 | 52,900 |
| | Choctaw | 15,342 | 14,872 | -3.1% | NA | NA | NA | NA | NA |
| | Coal | 6,031 | 5,856 | -2.9 | NA | NA | NA | NA | NA |
| Total | | 657,839 | 736,653 | 11.98% | 755,725 | 789,043 | 840,199 | 916,067 | 989,998 |

*Collin County only includes the residents of the following towns within the Primary ASR: Allen, Anna, Celina, Melissa, McKinney, plus 10,000 people living in smaller towns.

Texas is one of the fastest growing State in the nation. Since the year 2000, the population has increased by 18.8%, more than twice that of the nation, which has increased by 9.1%. The Oklahoma population instead has increased only by 6.9%, but some of its counties, included in the GYI ASR, have increased at a higher rate, such as Bryan and Marshall Counties. In addition, the DFW MSA is the nation's fourth largest urban area, with most of its growth concentrated on the northern side, which corresponds to the southern boundary of the GYI ASR. In fact, many large companies have already established offices and plants in the ASR, which serve as a foundation for the business operations that may increase in the future. The Texas State Data Center forecasts the population growth will continue in the upcoming years, as more businesses establish and relocate their offices, plants or headquarters into the region. The population is continued to increase because of the area's economic growth of the region and the low cost of living.

The increase in population for the ASR from 2000 to 2009 is estimated at about 12%, which is 33% more than the national average. The number of people living today in

the ASR is already over 736,000. The forecast shows that the ASR will have about 840,000 residents in 2020 and around a million in 2030. The employment/unemployment data and the median incomes of the ASR are depicted in Table 2.

Table 2: County-Level Employment status and 2008 Median Household Income for the Primary Airport Service Region

| STATE | County | Unemployment rate (%) | | | Employees | Median Household Income 2008 | % of State Median HH Income |
|---|---|---|---|---|---|---|---|
| | | 2007 | 2008 | 2009 | | | |
| | USA total | 4.6 | 5.8 | 9.3 | - | $52,029 | |
| Texas | Texas | 4.4 | 4.9 | 7.6 | - | $50,049 | 100.0% |
| | Grayson | 4.7 | 5.3 | 8.1 | 54,450 | $45,669 | 91.2% |
| | Cooke | 3.6 | 3.6 | 6.6 | 21,960 | $49,705 | 99.3% |
| | Fannin | 5.2 | 5.9 | 8.8 | 12,848 | $41,320 | 82.6% |
| | Hunt | 4.7 | 5.4 | 8.2 | 37,501 | $45,396 | 90.7% |
| | Lamar | 4.7 | 5.5 | 7.8 | 23,884 | $37,104 | 74.1% |
| | Collin* | 3.9 | 4.6 | 7.3 | 113,615 | $81,875 | 163.6% |
| Oklahoma | Oklahoma | 4.0 | 3.7 | 6.4 | - | $42,836 | 100.0% |
| | Marshall | 2.7 | 2.3 | 4.5 | 6,113 | $43,895 | 102.5% |
| | Atoka | 5.3 | 4.5 | 8.3 | 5,716 | $32,735 | 76.4% |
| | Bryan | 3.4 | 3.4 | 5.3 | 20,449 | $35,070 | 81.9% |
| | Pittsburg | 3.8 | 3.1 | 6.2 | 22,676 | $39,722 | 92.7% |
| | Carter | 3.4 | 3.2 | 5.9 | 25,338 | $39,360 | 91.9% |
| | Choctaw | 5.2 | 5.0 | 7.2 | 6,576 | $27,774 | 64.8% |
| | Coal | 6.1 | 4.8 | 8.6 | 2,375 | $32,259 | 75.3% |
| - | Total ASR employees | | | | 353,501 | - | |

Sources: US Bureau of Labor Statistics, Local Area Unemployment Statistics (LAUS) data; US Bureau of the Census, Small Area Income & Poverty Estimates Program

Despite the financial crisis affecting the US, the unemployment rate in the ASR is mostly below the national average, which is at 9.2% as of September 2010. The median household income is lower than the national median income and most are lower than the state median income, but the cost of living in all the counties in the ASR is much lower than the national average and in some cases lower than the State average. Although the majority of the ASR population resides in urban areas, a considerable number of people still live in rural areas; just 3% overall lives in rural-farm areas. Manufacturing industries located within the ASR belong to several commercial sectors, including food preparation and transformation, metal, wood, plastic and rubber. Computer/electronics and other manufacturing activities, account for a total of more than 23%. Retail, wholesale and food stores employ more than 18% of the people, which is the fastest growing employment sector for the ASR. Health care, professional services and construction-related jobs are other major employment sectors. The rest of the employment mostly belongs to the public, government or education sectors.

The increase in population, together with the growing economy of both North-Central Texas and Southern Oklahoma, will generate increasing business activity, which should enhance business air travel demand and thereby will raise the potential number of airport users, among the residents of the ASR.

**Passenger traffic at the DFW area airports – the Trip Rate (TR)**

Originating enplaned passengers are those who board on a round trip flight from a particular airport, therefore excluding the connecting passengers who originally boarded from a different airport. The trip rate is the ratio of annual originating enplaned passengers per airport region population, as shown on Table 3.

Table 3: Trip Rate in the Dallas/Fort Worth MSA

| Dallas Fort Worth Metropolitan Statistical Area | | |
|---|---|---|
| Estimated Originating Enplaned Passengers CY 2009 | Population 2009 | Trip rate |
| 12,625,026 | 6,447,615 | 1.96 |

Source: estimated from the data contained on the FAA Air Carrier Activity Information System (ACAIS) for the DFW and DAL airports in CY 2009.

Then, the total number of estimated originating passengers from the DFW MSA airports (DFW and DAL), divided by the population is defined as the DFW MSA trip rate; for the calendar year 2009, it is computed as 1.96. Therefore, each resident of the DFW MSA accounts for 1.96 round trip flights originating at the local airports.

**Estimation of GYI Originating Enplaned Passengers**

As shown on Table 1, about 755,000 people are currently living in the GYI ASR, and the number is forecast to increase up to a million within the next twenty years. However, among those people, the actual number of air travelers who currently fly from the surrounding areas airports is not available for analysis and research.

To assess the opportunity to turn GYI into a commercial airport and to encourage airlines to start operations, it is critical to estimate the number of originating enplaned passengers; to achieve this result; the authors decided to use the trip rate concept. Given the GYI ASR, the Trip Rate (TR) correlates the number of its residents with the number of Originating Enplaned Passengers (OEP) at the airport, solving the equation:

$$\frac{OEP}{ASR\ Population} = TR$$

The only known value is the population number (forecast as shown on Table 1), by estimating a value for TR per each year considered, the equation can be solved, giving the result on the estimate of OEP; the results are shown in Table 4.

Table 4: Forecast of GYI passengers from the ASR

| | 2010 | 2015 | 2020 | 2025 | 2030 |
|---|---|---|---|---|---|
| GYI ASR Population (Forecast) | 755,725 | 789,043 | 840,199 | 916,067 | 989,998 |
| ASR Residents Trip Rate (estimated) | 0.1 | 0.12 | 0.15 | 0.2 | 0.25 |
| Estimated GYI Passengers | 75,573 | 94,685 | 126,030 | 183,213 | 247,500 |

The current economic activity within the GYI ASR appears unlikely to sustain a similar trip rate as the DFW MSA because the majority of air travel is considered business related. Therefore, for 2010, a trip rate of 0.1 is assumed for the GYI ASR. This rate is in absolute terms very low; however, this conservative estimate leads to a forecast of about seventy-five thousand GYI ASR residents in the calendar year 2010 embarking on a round trip flight, with an average of two hundred and five originating enplaned passengers per day.

Regional airlines generally operate aircraft with a capacity of 100 seats or more; set a target load factor of 70% before initiating a new route. Therefore, the estimated number of GYI OEP already appears consistent with satisfying this critical airline requirement. The number of originating passengers from the GYI ASR will increase in the future, as the ASR population continues to grow and the economic activity improves.

**Reasons for air carriers to operate at GYI and marketing strategies for the airport to attract them and other investors**

With the strong support of its large number of potential users and applying a targeted marketing strategy, GYI is embarking on a program to attract companies to invest and air carriers to operate flights to and from this airport. Since the airport size and passenger traffic may not justify an air carrier operating at GYI to set up equipment and personnel for tug and tow operations, and baggage handling services, the airport could provide them at very reasonable rates. Eventually, airlines could treat GYI as a spoke in their network, allowing passengers boarding at GYI to reach their final destination through connecting flights, including international destinations. This strategy would attract more passengers from this region to use GYI for air travel.

The combined total weight of the freight landed at the DFW MSA cargo airports (DFW and Alliance Airport in Fort Worth) ranks eighth in the nation; therefore, the cargo activities with a focus on the goods moving within the ASR could be an opportunity to expand airport activities. The status of a Foreign Trade Zone (FTZ) on the airport property allows goods to be treated with special U.S. Customs rules and regulations; therefore, air cargo firms can establish a base at GYI and enjoy these benefits and competitive lease rates for the hangar facilities. In addition, the airport is located near highways free of traffic congestion (unlike the airports inside the DFW metropolitan region, where traffic congestion is a frequent occurrence), which should allow for easy movement of the freight originating or destined to sites within the GYI ASR. Other incentives such as tax abatements and tax increment financing that is offered by regional economic development agencies to encourage future development at or near GYI. This will draw other firms, institutions and manufacturers to locate their businesses in this area with lower taxation and closer to the cargo base.

Several casinos are located in Southern Oklahoma State that could benefit from the convenient location of GYI. They may consider developing partnerships with tour operators or airlines for access to charter or scheduled flights for casinos visitors/patrons.

**Reasons for Passengers to Prefer GYI rather than the DFW and Southern Oklahoma Airports**

Residents of the GYI ASR, depending on where they live, where they have to fly and what kind of airline service they need, currently have to travel to DFW, DAL, OKC or Tulsa International airport (TUL) to fly to domestic and international destinations. Since public transportation is not provided to/from this region for these airports, and the use of shared shuttle service is discouraged by its high price (ranging from $200 to $220 each way). Therefore, for ASR residents, the only way to reach the airports is

to drive a private vehicle. The use of private transportation has added costs due to the drive time, fuel expenses, the toll roads and the parking fees at the departure airport. In addition, the car parking fees typically increase with their proximity to the terminals; therefore, low price parking will be far from the terminals and connected by a shuttle bus, thus increasing the total trip time to reach the terminal. The weather and the traffic congestion on highways and freeways, in particular on those leading to the DFW area airports can further increase the total trip time and cost; therefore, GYI may attract the ASR residents because of its convenient location. It can be reached without encountering heavy traffic, and the airport management plans to offer free parking as an added incentive. Table 5 compares the access costs between DFW and GYI, when driving from several cities within the GYI ASR and the duration of the trip vary. The analysis assumes the travel cost per mile as the Standard Mileage Rate, defined by the US Internal Revenue Service; for 2010 the rate is $0.50/mile when the purpose of the trip is business and $0.165 for other purposes. Additional costs associated with tolls are considered. The value of time associated with driving to the airport and on any transit connections at the airport are not specifically considered in this comparison; the inclusion of these costs will typically enhance the benefits of accessing GYI over DFW.

Table 5: Comparison of Access Costs to DFW and GYI

| | To DFW Airport | | To GYI Airport | | DFW | | | GYI | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| From | Distance (miles) | Time | Distance (miles) | Time | 1 day business trip | 3 day weekend trip | 7 day vacation trip | 1 day business trip | Non - business trip | Max saving | Min saving |
| Sherman,TX | 69.1 | 1:08 | 7.4 | 0:13 | $95.34 | $62.04 | $81.04 | $8.14 | $2.44 | $92.90 | $53.90 |
| Denison,TX | 80.0 | 1:18 | 10.8 | 0:16 | $106.24 | $65.64 | $84.64 | $11.88 | $3.56 | $102.68 | $53.76 |
| Durant,OK | 98.8 | 1:35 | 29.6 | 0:32 | $125.04 | $71.84 | $90.84 | $32.56 | $9.77 | $115.27 | $39.28 |
| Gainsville,TX | 58.9 | 0:59 | 35.8 | 0:39 | $77.34 | $50.88 | $69.88 | $39.38 | $11.81 | $65.53 | $11.50 |
| Paris, TX | 114.7 | 1:51 | 68.0 | 1:04 | $140.94 | $77.09 | $96.09 | $74.80 | $22.44 | $118.50 | $2.29 |
| McKinney,TX | 37.9 | 0:43 | 39.4 | 0:43 | $64.14 | $51.75 | $70.75 | $43.34 | $13.00 | $57.75 | $8.41 |
| Plano, TX | 29.0 | 0:33 | 54.4 | 1:08 | $46.00 | $39.57 | $58.57 | $54.40 | $17.95 | $40.62 | - |
| Frisco, TX | 28.0 | 0:35 | 46.3 | 0:52 | $54.24 | $48.48 | $67.48 | $50.93 | $15.28 | $52.20 | - |

Comparing the specific values from Table 5, one can deduce that GYI has an access cost advantage over DFW for almost all of the cities in the GYI ASR. However, two of the selected cities, Frisco and Plano, show lower cost savings; these lower savings support the researchers' decision not to include portions of the Collin County population in the GYI ASR population estimates. McKinney, TX, also has a smaller access cost savings compared to some of the other cities. Therefore, a more detailed market analysis would be required to determine the airport that Collin County residents would prefer given a set of routes serving GYI. Nonetheless, these benefits in GYI access and use may outweigh the fact that many flights from GYI will likely require transfers before reaching the final intended destination.

Not only would the GYI ASR residents benefit from commercial operations, but business travelers appear likely to select it when headed to the ASR; in fact, they would likely choose GYI airport for its low access costs when it is the most convenient location based on their final trip destination.

## Conclusions

*Timeline for Conversion to Scheduled Commercial Operations*

This initial investigation indicates that conversion to commercial operations at GYI looks promising over the upcoming decade; the conservative forecasts indicate that GYI may have as much demand as other Texas airports, based on 2009 enplanements: as Tyler today, as Abilene by 2020 and as much as Killeen by 2030. This paper defines the path for the development of the GYI towards the start of commercial air carrier operations with a short, medium and long term plan.

In the short term (0-3 years), the airport management can attract air cargo operators, and utilize attractive marketing strategies that emphasize the advantages of its status as a FTZ. GYI can become a cargo hub that serves the GYI ASR in north central Texas and Southern Oklahoma. In addition to these advantages, the cargo airlines may also have an interest in operating from a site where the traffic congestion is not a significant factor, which can make for reduced travel time and inexpensive freight movement. Finally, the land near the airport is much more affordable than comparable locations adjacent to airports within the DFW MSA. These factors make cargo operations a likely first step towards scheduled commercial operations.

In the medium term (2-5 years), GYI is contemplating to apply to the FAA for the issuance of an Airport Operating Certificate under provisions of 14 CFR Part 139 - Class IV, in order to be able to operate only non-scheduled flights with airplanes having more than 30 seats. To meet these criteria the airport marketing team is planning to establish partnerships with the local tourist attractions and with tour operators, in particular those connected with the casinos. These tourism partnerships can bring in chartered flights from US airports to serve the tourist destinations.

In the long term (5-10 years), once the airport management and employees become familiar with large commercial airplane operations, a Class I Airport Operating Certificate will be required to operate scheduled flights. Scheduled operations will require at least one commercial airline partner. The airport management will begin exploring a variety of approaches for creating and securing a partnership.

Airports and populations that request commercial service operations are very large and this is why service is discontinued at several regional airports. As a result, commercial aviation opportunities will likely develop in exurban areas where there is a large population base but the travel time to major airports becomes too large.

*Further Investigations*

This paper provides an initial assessment of the viability of transitioning GYI from strictly general aviation airport to commercial operations airport. Beyond improving facilities and increased marketing and outreach, the airport management has little control on the pace of development and operations at the airport. Ultimately, any business that considers beginning or moving operations to GYI may decide to perform their own business plan; however, GYI management would like to undertake few more studies to improve their marketing strategies and help plan its infrastructure investments. The travel demand, both passenger and cargo, within the ASR requires a further beneficial investigation; this in depth study should identify the current conditions as well as forecast future usage. GYI should develop a collaborative

partnership with a commercial airline and undertake a vivid market and route viability study for the population living within the ASR. Further systematic evaluation of the existing facilities will also improve the airport management's ability to plan and schedule for the near future infrastructure improvements as GYI moves towards a fully functional commercial operations airport.

**References**

Coffman Associates (2003). "Grayson County Regional Airport Master Plan Update"

Federal Aviation Administration "Airport IQ 5010 – Airport Master Records and Reports" < http://www.gcr1.com/5010WEB/> (June 18, 2010)

Federal Aviation Administration "Passenger Boarding (Enplanement) and All-Cargo Data for U.S. Airports" <http://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/passenger/media/cy08_cargo.pdf> (June 18, 2010)

Jacobs Consultancy (2006). "Airport Business Plan Grayson County Airport"

Justia.com "Employment information by county" < http://www.justia.com/us-states/texas/#counties> < http://www.justia.com/us-states/oklahoma/#counties> (June 17, 2010).

Oklahoma State Data Center – Oklahoma Department of Commerce "Population Projections by City by County" < http://www.okcommerce.gov/Data-And-Research/Downloads/Population-Projections> (June 17, 2010).

Research and Innovative Technology Administration – Bureau of Transportation Statistics "Air Carriers : T-100 Domestic Segment (All Carriers)" <www.transtats.bts.gov> (June 18, 2010).

Texas State Data Center and Office of the State Demographer "2008 Population Projections – Texas Counties" <http://txsdc.utsa.edu/tpepp/2008projections/2008_txpopprj_cntytotnum.php> (June 17, 2010).

US Census Bureau "State and County quick facts" <http://quickfacts.census.gov/qfd/index.html> (June 16, 2010).

United States Department of Agriculture "County-Level Unemployment and Median Household Income", Bureau of Labor Statistics, Local Area Unemployment Statistics (LAUS) data, Bureau of the Census, Small Area Income & Poverty Estimates Program < http://www.ers.usda.gov/data/unemployment/> (June 19, 2010).

Wikipedia "The Dallas/Fort Worth Metropolitan Area" <http://en.wikipedia.org/wiki/Dallas-Fort_Worth_metropolitan_area> (June 23, 2010)

# *Ground Access for IGI Airport, Delhi: Planning and Challenges*

***Dileep Dixit**, Author*[1], and ***Phua Chai Teck**, Co Author*[2]
1 Associate General Manager, Ground Access, Delhi International Airport Limited, IGI Airport, Delhi, **PH** (+91) 9958098683; **email:** dileepdixit@yahoo.com
2 Vice president and Head Master planning, Delhi International Airport Limited, IGI Airport, Delhi, **PH** (+91) 991038301; **email**: ctphua@alum.mit.edu

**ABSTRACT**

Airside infrastructure, Terminal along with Ground access facilities, form three broad components of any airport. An airport attracts large number of passengers, meeters and greeters and staff depending upon its size, location and role. Efficient ground access system is therefore vital for any airport to serve its full potential.

Indira Gandhi International Airport (IGIA) in New Delhi is a crucial airport, which is being upgraded and modernized by Delhi International Airport Limited (DIAL) under Public Private Partnership (PPP) initiative, covering funding, designing, construction, operations and maintenance of the airport for a period of 30 years till 2036, with a further option to extend it by next 30 years. The first phase is marked by construction of new integrated terminal T3 by 2010 to handle 34 million passengers per annum. This would be followed by three more passenger terminals till 2026 in subsequent phases.

This paper presents an overview of the concept of the central spine and the connecting network within the airport boundary, design concept, other general details such as Ground access facilities, both by roadway and by metro express link, curbside configuration, multi level car parking along with parking for buses and taxi staging, express bus service to various city centers, shuttle bus service linking different parts of airport, integration of circulation plan of cargo complex and hospitality districts.

## 1.0 Introduction

Delhi, the fast emerging global city, aided by a vibrant and rapidly expanding Indian economy has experienced rapid growth and urbanisation in the past few years. The booming economy, which is set to continue, together with increased liberalisation of the air industry has had a significant impact on the demand for both domestic and international air travel at Delhi Airport. According to 2001 census the population of Delhi has been 13.8 million which makes it the third largest metropolitan area in India after Mumbai and Kolkata and one of the fastest growing metropolitan areas in Asia.

Indira Gandhi International Airport (IGIA) Delhi, being the busiest airport in south Asia, handled a total of 27 million passengers in 2009 and even the “most conservative” forecasts indicate the passenger flow to be 80 million per annum for the year 2026. The growth in cargo tonnage through the airport is forecasted to increase at a similar and possibly higher rate.

This growth over the years has magnified many deficiencies and capacity constraints confronting IGIA. In order to bridge the gap in an efficient manner and bring about a perceptible change in the service quality, Ministry of Civil Aviation decided to expand and modernise the airport through private sector participation. Besides addressing host of airside and terminal based requirements, various key land side issues to be tackled

by DIAL in planning and design of IGIA (after taking over) included severe congestion on highway access (NH8), poor accessibility, inadequate curb frontage, poor parking management, improper segregation of different modes, shortage of parking, mismanagement of taxi pick up system, minimal role and poor quality of public transport and lack of coordination between airport planners and local authorities.

In order to achieve the given mandate, an efficient master plan was prepared for phased incremental expansion of IGIA with its "Saturation" phase envisaged for 2036. New terminal T3 with a capacity to handle 34 million passengers per annum having a built up area of 5.4 million sqft was planned along with other supporting infrastructure as part of phase 1 with a target commissioning date of July 2010. Other salient features of phase 1 include - a new runway, 78 aero bridges, 9 remote parking stands, 90 walkalators, 47 escalators, 70 lifts etc. A crucial element for the successful realisation of the master plan is the integrated ground access system, supported by a multi-modal connectivity to the airport to serve air passengers, well-wishers, employees, and other airport users. An 'efficient ground transportation access' is the key to realizing the full capacity potential of any airport, its importance can't be rated less than that of airside facilities or terminal facilities.

## 2.0 Location Setting

Indira Gandhi international airport is located in in the outskirts of Delhi city close to Haryana-Delhi Border (Near Gurgaon). Airport is linked with its catchment area by National Highway 8. Airport area is mainly surrounded by residential land uses towards Delhi side and farms, scattered residential layouts, commercial and industrial locations on Gurgaon side (Figure 1).

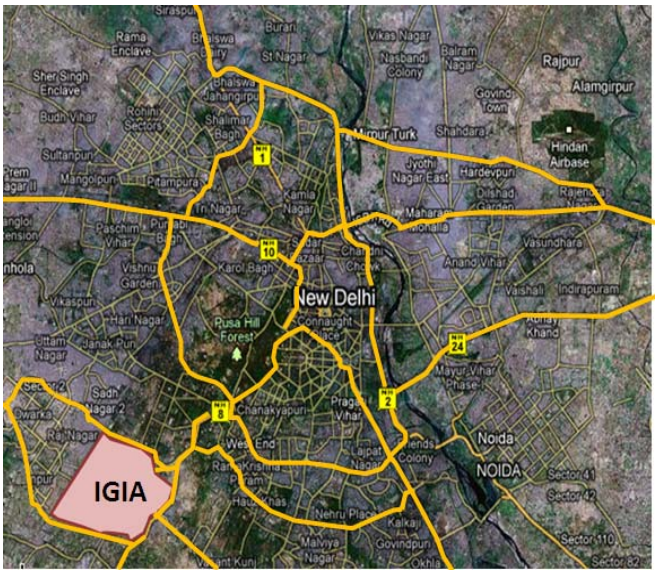



Figure 1

## 3.0 Components of Ground Access

Development of a comprehensive, multi modal and integrated airport ground access management plan for airport is based on demand estimates that are consistent with those used in master planning and terminal building design. The salient components of the airport ground access systems are:

- regional highway system,
- access roads and ramps leading into the airport,
- circulation roads at airport,
- curb loading and unloading areas,
- multi Level Car park and other parking facilities

- public transport infrastructure and intermodal provision
- dedicated airport metro express link and facilities
- taxi staging and hold up areas
- integration of circulation plan of cargo complex and hospitality district

## 4.0 Methodology

In order to develop an efficient ground access system for IGIA, a scientific approach was adopted. A thorough analyses of key traffic characteristics, travel pattern, socio – economic and behavioral traits of passenger, staff and meters and greeters was carried out based on primary surveys (conducted in the base year 2007) and through secondary data. Decisions on above ground access elements were based on projected transport demand, airline schedules, passenger and airport operating policies, physical constraints, construction deadlines and learning from international best practices from other successful airports.

### *4.1 Airport Traffic Generation - Forecasts*

Terminal specific road network and facilities have been designed to serve traffic volumes anticipated in 2016 when integrated terminal T3 attains its peak. The road network that is common to T3 and other future terminals is currently planned for 2016 demand with a provision of expansion to meet the future demand. For example the central spine is planned to be initially with a configuration of 4+4 lanes till 2016, keeping a provision of expansion upto 10+10 lanes to facilitate construction of future ramps and loops required to serve ultimate demand in the saturation phase.

According to recent survey number of PCUs using the International Terminal are 15,600/per day, while 43,400 PCUs /per day visit the Domestic Terminal. Peak Vehicular Traffic of 1,038 PCUs at international terminal takes place at 2100 -2200 Hrs. Peak traffic at domestic terminal works out to 3,099 PCUs at 1700 – 1800 Hrs. Considering only the 75% of domestic traffic would shift to integrated terminal T3 and the combined peak hour vehicular traffic for full service domestic airlines and international carriers operating from T3 would be 3118 at 2100 -2200 Hrs. Modal split analysis based on current survey indicate that Taxi at 48% and Cars at 38% account for 86% of vehicles at the airport. Share of two wheelers is 6% while public transport and Auto Rickshaws both have a nearly equal share of 4% each. The Traffic forecasts, during the peak hour at 21.00 - 22.00 hrs in 2016 works out to 5293 PCUs. The corresponding number of passengers at T3 during this period would be 3234. The total peak parking accumulation would be 4100 cars.

### *4.2 Regional highway system*

National Highway 8 (NH 8) that links the administrative capital of the country, Delhi with the country's financial capital Mumbai, offers IGIA the vital access to regional highway network. As the expansion of city of Delhi has taken place in all the directions, the airport, which is hardly 20 km from the city center (Cannaught Place) has been surrounded by dense residential and commercial settlements in the eastern and western side. In such case, the ring road system and other arterial radial roads emerging outwards from the city center form part of off-site road system for IGIA. NH 8, Rao

Tularam Marg, Mehrauli – Mahipalpur Road, Dwarka Road (near Palam) and Dwarka Express Link together with ring roads - radial roads from the city center provide the regional highway/ arterial road network to serve the airport. These road stretches require various improvement measures for example provision of grade separation, signalization in place of certain rotaries, better maintenance practices, provision of bus shelters, better enforcement etc. In order to execute these measures, efforts have been made towards achieving a good coordination between airport planners and respective external agencies/ local authorities to take note of airport requirements.

*4.3 Access roads and ramps leading into the airport*

This component is vital to ensure optimal distribution of both ingress/ egress traffic to and from the airport. Primarily four routes provide ground access to IGIA from northern side -. NH 8, Rao Tularam Marg, Mehrauli – Mahipalpur Road, Dwarka Road (near Palam). Two roads, in the south, Dwarka Express Link (which runs along the southern border of the airport) and Old Gurgaon road (near Kapashera) along with four roads in the north are key 'near-airport' roads. Any blockage in any of these roads affect the airport bound traffic flow.

The city of Delhi is expanding in the western and north western directions. There are large urban extensions, townships, freight complex, railway and inter-state bus terminals proposed and are under different stages of implementation by various agencies in the western side of the airport. Central Road Research Institute (CRRI) was engaged to study the impact of all these developments including airport modernization project by DIAL by involving all stakeholders. CRRI recommended an integrated Traffic circulation plan around IGI Airport which included several tunnels, flyovers and surface improvement projects on the eastern side, to be implemented in a phase wise manner to ensure proper dispersal of traffic on NH 8. CRRI also suggested two tunnel projects from North and South of IGIA for immediate implementation to reduce airport's dependence on NH 8 instead of a direct access from western side in order to save airport network from through traffic (non airport based traffic).

Major capacity enhancements on Delhi roads in general and particularly on the road network in the immediate proximity to IGIA, have recently been executed through reconstruction, widening and grade separation at key junctions and street scaping by city administration and federal agencies in the wake of forthcoming Commonwealth Games, to be hosted in Delhi in October 2010. NH 8 and Rao Tularam Marg have been developed as part of Delhi – Gurgaon expressway with grade separations at Palam, Mahipalpur and Radisson Junctions (specific for the airport). Improved operational management is being effected by restricting slower vehicles and "through tolling". The major gateways to access different terminals and allied facilities inside IGI airport are detailed as under (Figure 2):

1. **Airport Road interchange with NH 8 at Palam :** This Interchange provides connectivity between NH – 8, Rao Tularam Marg and approach road to Terminal – 1 (Present cluster of domestic terminals).

2. **Radisson interchange:** This existing interchange provides access from NH 8 and provides the sole entry – exit to different airport bound traffic streams..
3. **Underpass at NSG roundel** (near Palam interchange) is nearing completion, which will provide a segregated route for Dwarka bound traffic. The surface junction will dedicatedly serve the airport bound traffic.
4. **Northern Access Road** is under fast track implementation, to provide an alternate entry exit to IGIA T3, by means of a tunnel under the runway. This road project has a 3+3 lane configuration, which starts from the NSG junction (discussed above) and meets the airport road network near Hanuman Temple passing through the defence area and IGIA operational area.

   The current approach to the IGI Airport for all the traffic coming from northern parts of Delhi is via grade separated Radisson Junction on NH-8, which already experiences congestion. There is no present direct alternate access from north to the airport, which is nearly 60 to 70% of total traffic generated from the airport. Northern access road has been planned and taken up for fast track implementation in order to directly serving this traffic from north and reducing the traffic pressure on NH-8. This project involved construction of 363 m long tunnel by cut and cover method under the flight path of runway 28. Execution started from the extended CL of the runway and move towards the two ends (of the tunnel) in order to ensure work completion in the extended basic strip area within runway closure period. Land availability, approvals from airport operations team, city planning agencies, Civil Aviation and security authorities along with coordination with various agencies posed unique challenges.

5. **Northern Junction**: The new junction north of Mahipalpur junction with a provision of underpass (in future) for entering traffic from Delhi towards Airport is another crucial project that is under implementation. This junction provides vital segregation to the traffic generated at Hospitality district and acts as an alternate entry/ exit point for the airport.
6. **Future long term projects**: There are 3 other long term projects which are under feasibility and planning stage namely grade separation at Mahipalpur Junction to provide east west entry – exit from Mahipalpur - Mehrauli Road, Southern Access Road from Dwarka Express Link and Northern Underpass (at Northern Junction).

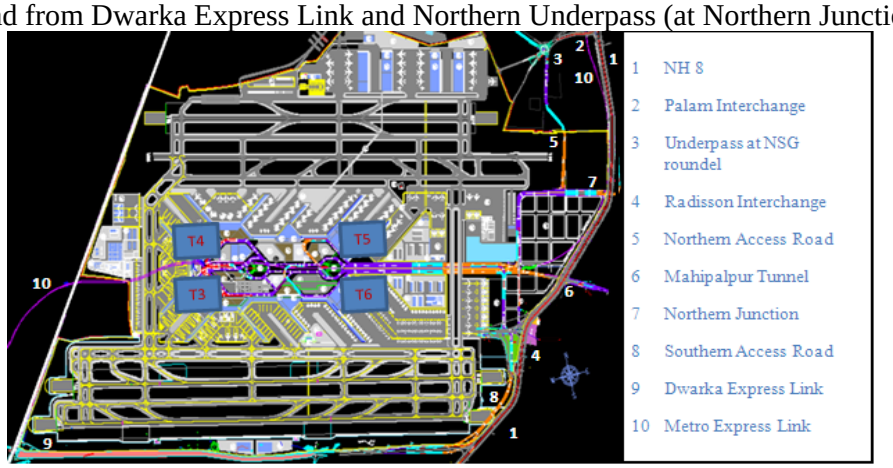



Figure 2

### *4.4 Circulation roads at airport*

The landside circulation roads at IGIA address the connectivity requirements to various current and future terminals from regional highway system and covers the Inter-terminal traffic circulation plan along with connectivity to other facilities such as Multi Level Car Park, Airport Services Building (ASB), graded taxi and bus parking, VVIP and staff access, airside access, Cargo terminal, hospitality district etc.

The master plan proposes a system of four terminals developed on either side of a central road called the 'central spine' starting from the Mahipalpur junction on NH 8 with a central reserve width of 123m. The central spine consists of two carriageways of 10 lane roads on either side with a central reserve of 6m for Metro corridor (Figure 3). The master plan facilitates independent circulation for each individual terminal by means of different level approaches for departing and arriving passengers.

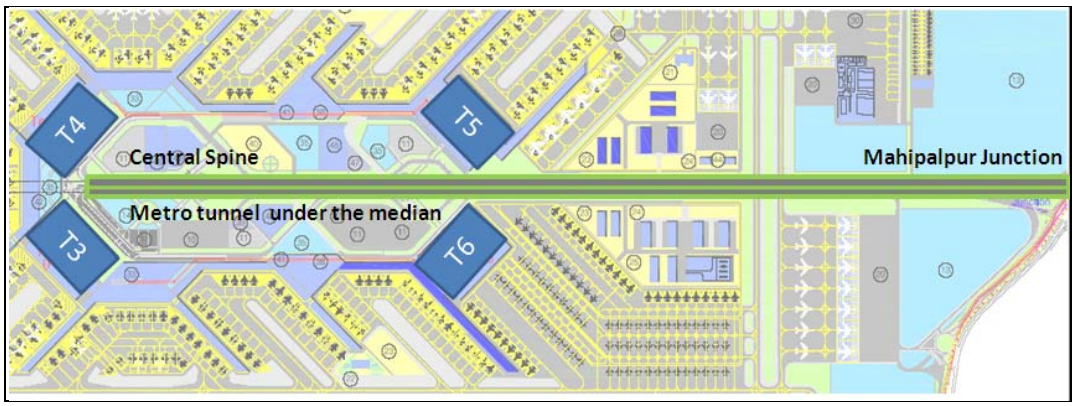



Figure 3

Currently (in Phase 1) a part of the central spine with a 4+4 lane configuration (expandable to 10+10 lanes) is linked to the existing approach road originating from Radisson junction on NH 8 as shown in Figure 4. This main approach is supported by pleasant landscapes and a network of service roads to take care of all access and circulation requirements within the airport.

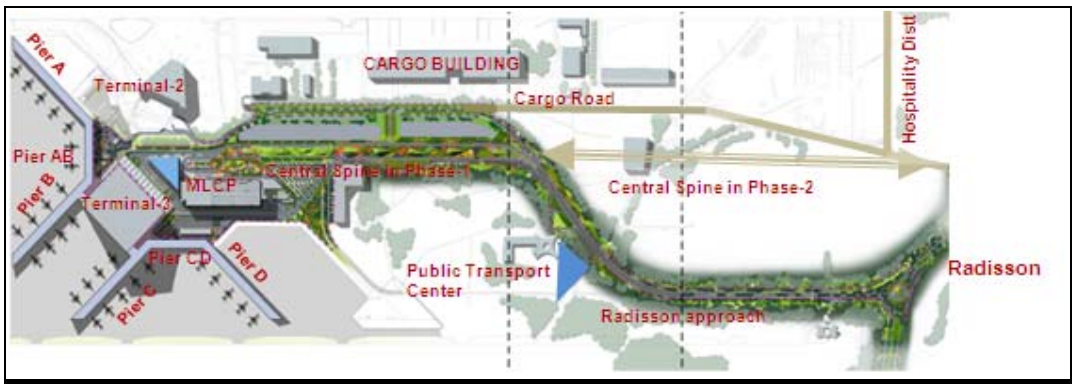



Figure 4

The central spine needs to be extended to link the northern access road with T3 network and also to provide connection to the hospitality district (a portion of which will be operational by mid 2011). Apart from extension of central spine, there is a plan to construct a separate access to the Cargo Complex through dedicated cargo road, in order to segregate cargo and passenger traffic.

### *4.5 Curb loading and unloading areas*

Departure Forecourt is on an elevated deck that will facilitate drop-off directly in front of check in concourse, while the arrival forecourt is on the surface. Both the arrivals and departure forecourts are served with multiple lane terminal frontage road (nearly 300 m in length) to distribute vehicles directly in front of terminal building. These frontage roads are separated with the help of curbs or raised islands to facilitate dropping/ picking or waiting of passengers. There are three curbs with 8 lanes in the elevated departure forecourt, while arrivals have 5 curbs with 10 lanes (Figure 5) to facilitate segregation of stopping, maneuvering and through movements of vehicles and provision of safe pick-up/ drop-off points. A cross section of the terminal curb roadway is shown in Figure 6.

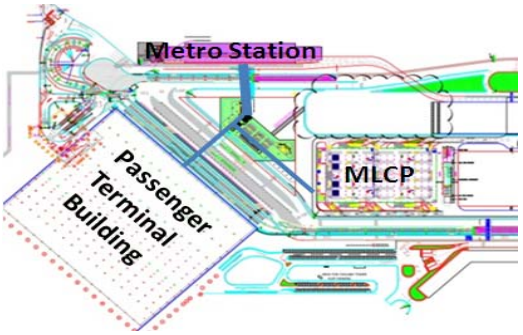



Figure 5

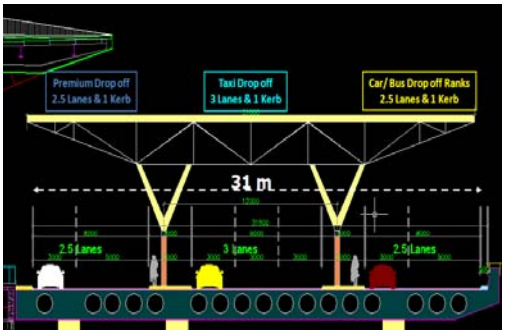



Figure 6

To segregate pedestrian flows with vehicular flows and facilitate seamless integration with metro station, Multi Level Car park (MLCP), meters greeters area, curb for Taxis and Private cars, sky walk and subways supported with elevators, escalators and ramps and walkalators have been provided (Figure 7).

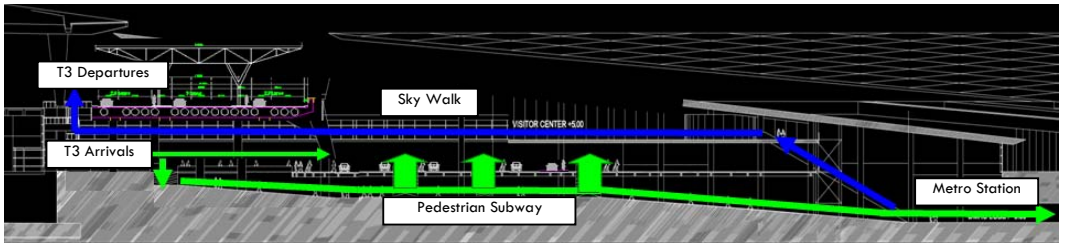



Figure 7 ← Departures – movement towards Terminal

→ Arrivals – movement towards Metro Station, MLCP, Curbside

### *4.5 Multi Level Car park and other parking facilities*

The Multilevel parking building, having a capacity of 4300 cars, is located adjacent to the New Passenger Terminal Building T-3 and is oriented from East to West parallel to central spine. Circulation of MLCP is planned in such a way that all entries are from southern road that runs parallel to the loop road leading to Arrival level of PTB, while all exits are on the northern road that merges with the central spine. These roads have barriers with toll collection system (ticket dispenser, card reader etc.).

The multilevel car park is a simple and straightforward building of 182 m by 117 m size comprising two blocks arranged side by side with two ramps in each block for up and down movement of the cars. It is a 7 level structure including ground level and terrace level (to be used for open parking). The proposed circulation system allows vehicles to move quickly through the system. If a parker enters at ground floor and parks his car at the furthest distance at terrace floor, it takes less than 5.0 minutes. While designing the parking system, an efficient geometric layout in the form of rectangular parking lots with vehicles parked on both sides of the aisle has been adopted. In addition, the parking facility is equipped with two Lifts (capable of carrying trolleys), staircase, toilets, pay booth in the entrance block, effective signage, CCTV, access control system, Space availability display system (SADS) to facilitate automated way finding, and a drop off zone in front of lift lobby on each floor.

Besides a direct connection on the ground floor, a pedestrian entry has been considered from the second floor level through covered walkway leading to the Terminal-3. Passengers/parkers coming from the Passenger Terminal Building with their luggage trolleys will be able to reach the parking building through this walk way and vice versa.

### *4.6 Public Transport and Inter-modal Provision*

Dedicated aero express buses have been planned to serve airport passengers. These buses are equipped with FIDS (flight information on display), GPS, ample luggage space and comfortable seating. Services are supported by custom designed Bus shelters at city end pick up points and enjoy direct access upto the forecourt; The tickets are issued with the help of Ticket Issuing Machines linked to Central Server at IT department; Proposals are under way to provide Web ticketing and roping in local Airline Ticketing agents for selling the Aero Express Tickets.

Other buses include regular local stage carriage buses, interstate buses, feeder buses and private coaches run by tour operators. In order to serve the airport in a more efficient manner, a Public Transport Center (PTC) has been built on the main Radisson approach road. PTC (Figure 8) has different zones for different modes/ activities namely pickup drop off zone for buses, pickup drop off zone for Auto Rickshaws and waiting lounge for passenger with separate zone for the pickup and drop off for shuttle buses. These zones are connected with the waiting lounge by means of covered walkways. The location of Public Transport Center is shown in Figure 4.

Auto Rickshaw is a popular para-transit mode prevalent on city roads though not very popular amongst airport passengers. This is also a tourist attraction amongst foreigners. Exclusive pick up and drop off arrangements for Auto Rickshaws have been made along with pre paid booth at Public Transport centre.

*4.7 Shuttle Buses*

In order to link the public transport center with T3 and different office areas within airport complex, DIAL has also planned to operate an efficient shuttle bus service. This shuttle bus service would be served by three high quality buses, initially with a frequency of 10 to 15 minutes. These buses would regularly ply round the clock along the central spine to link T3 with public transport center. Shuttle buses will and pick/ drop bus users at the bus bay abutting the waiting lounge.

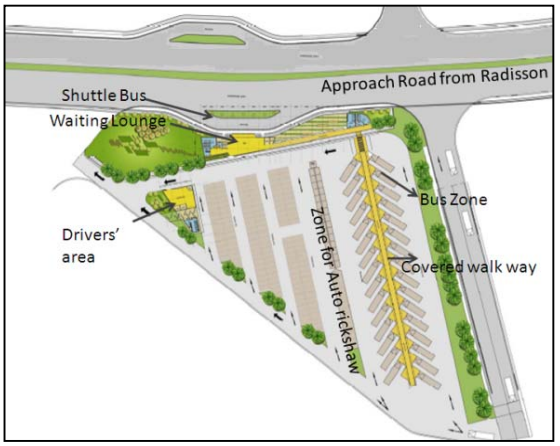



Figure 8

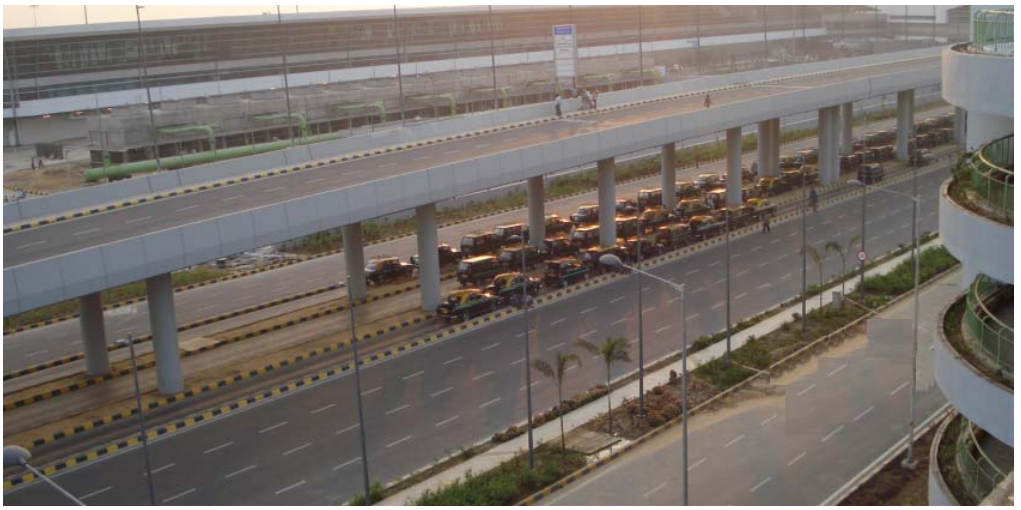
*4.8 Airport Metro Link*

There was a need to maximize the capacity of the site to handle 100 million passengers per annum at IGIA in the saturation phase. This would not have been possible by a road based system alone. In order to meet the challenge, requirement of a more sustainable transport system was felt and a dedicated Metro express link was planned for the airport to develop it as a fast, convenient, congestion free alternate mode. To make it attractive the station was planned at a comfortable walking distance from the Terminal with convenient transfer facilities. Simultaneous Construction of the huge Terminal building and the massive underground metro station along with managing an operating international terminal posed a lot of interfacing and traffic management issues. All of these were resolved in close consultation with Delhi Metro Rail Corporation (DMRC).

IGI Airport would be the first airport in India to have a dedicated metro link similar to all world class Airports today. This Metro Express Line (AMEL) will run between the New Delhi Railway Station and the Delhi-Dwarka Metro station to the south west of the airport via Shivaji Stadium (Connaught Place), Dhaula Kuan, the Aero City and the central terminal area station located between Terminals 2 and 3. The Delhi-Dwarka metro station is also connected with proposed Integrated Metropolitan Railway Terminal (IMRT) and the proposed inter-state bus terminal in the vicinity. The initial frequency of AMEL is 6 trains per hour and the planned journey time from the City centre to the airport is 18 minutes. Automated pedestrian facilities have been provided for transfer between the main terminal metro station and airport terminal forecourts and car park areas in order to ensure seamless, hassle free intermodal linkages. There is also a provision of baggage check in facility at down town metro station likely to be operational from next year. The projected ridership on the City- Airport leg is 42,000 in 2010 and 86,000 in 2021 with a similar volume on the Dwarka leg. The passenger capacity per train is 600-650 including 500 passengers seated.

*4.9 Taxi staging and hold up areas*

Taxis in India are quite economical compared to other western countries. Both premium and ordinary put together, nearly 50% of passengers depend on taxis, at the airport during peak hours. Taxis wait at large taxi holding areas, where there is a provision of parking over 1500 Taxis. Taxis are organized in queues inside these holding areas and wait for their turn to move to the forward taxi staging area under the departure forecourt ramps. Thereafter the taxis are directed to pick the passenger at the taxi curb and move out of the airport through the central spine.



*4.10 Integration of circulation plan of cargo complex and hospitality district*

The cargo terminal is located in the centre of the airport precinct, and is primarily accessed via Mahipalpur junction with the NH8, and also through Radisson approach road. There is a plan to construct a dedicated approach road for cargo traffic in order to ensure segregation of cargo traffic from the airport bound normal passenger traffic.

There is a plan to develop 256 acres area for commercial property development and it is important to assess its circulation, parking and access requirements in sync with the overall circulation plan of the airport for all phases. It is therefore important to plan commercial development in such a way that it has minimal traffic impact, and we are able to prioritize access to and from the airport.

**5.0 Capacity Analysis and Micro-Simulation Modelling**

Analyses of the internal ground access system as planned for 2016, show that the system will be adequate for the traffic volumes forecast for that year. In terms of landside traffic, the Terminal 3 is basically composed of the curbside, differentiated in Departures on level +10 and Arrivals on Ground level, and the Multi-Level-Car-Park in its immediate neighborhood. A thorough exercise of traffic assignment using Origin-Destination matrix, onto the airport road network links for the year 2016, was carried

out. Average Travel Speeds and Average Control Delays were calculated based on the simulation output using "VISSIM" Software.

During peak hour, it is observed that various direct conflicts between major traffic flows are replaced by indirect conflicts in the form of U-turns and weaving sections. During low traffic hours, minor conflicts can be tolerated without degrading the aspired level of service C.

After reviewing the calculations of the Level of Service for various stretches of road network supporting T3, it is established that the network offers a decent Level of Service C upto 2016.

## 6.0 Conclusion

It would have been inefficient and uneconomical to build at this time the number of lanes, intersection controls and parking spaces required for ultimate development. Hence an efficient road network and other support facilities coupled with a dedicated metro link have been provided, which is adequate to meet the phase 1 functional requirements of the airport upto 2016.

Integrated transport approach adopted in planning and operation of ground access infrastructure ensures convenient seamless transfers between air and ground transportation modes to reach final destination within region/ city. It helps in making metro and buses attractive and equally convenient modes of transport, thereby achieving the long term goal of modal shift from private vehicles to more environmentally and energy efficient mass transport modes.

## 7.0 References

1. The Regional Plan 2021 – National Capital Region
2. Delhi Masterplan 2021
3. Delhi Airport Traffic Forecasting Report – Final Dec 2006
4. Indira Gandhi International Airport Masterplan Report – Final 2006
5. Indira Gandhi International Airport Landside Traffic Report – Final July 2007
6. Detailed Report Airport Metro Link upto Dwarka

# Ground Access Trip Generation Models for Airport Planning

Geoffrey D. Gosling, Ph.D., M.ASCE[1]

[1]Aviation System Consulting, LLC, 805 Colusa Avenue, Berkeley, CA 94707;
PH (510) 528-8741; FAX (510) 528-8745; e-mail: gdgosling@aol.com

## Abstract

The geographical distribution of air passenger trip origins forms a critical input to studies of airport ground access and airport choice. This paper discusses the limitations of typical air passenger surveys to determine the distribution of air passenger trip origins at a required level of resolution, the resulting need for air passenger trip generation models, and considerations that arise in developing such models. The paper then presents a set of air passenger trip generation models that were developed to support the current update of the Bay Area Regional Airport System Plan Analysis and discusses their strengths and limitations. The paper concludes by discussing the need for additional analysis to improve the future utility of the models and the potential transferability of these models to other regions.

## Introduction

Studies of airport ground access and airport choice require information on the geographical distribution of air passenger trip origins. Typically this information is obtained from air passenger surveys. Commonly, the geographical distribution of ground trip origins is required at the level of resolution of the travel analysis zones used in the metropolitan transportation travel models that form the basis of the transportation service data used in the analysis. However, the sample size of most air passenger surveys is inadequate to provide a suitable level of resolution of air passenger trip origins by simply tabulating the survey results. Furthermore, there is the question of how the distribution of air passenger trip origins will change in the future in response to projected changes in regional development patterns and the composition and characteristics of households and other activities within each travel analysis zone. Clearly what is required is a trip generation model that can predict the number of air passenger trips that will originate from each travel analysis zone in some future year, as well as a generate a profile of the characteristics of those trips.

The paper discusses the limitations of typical air passenger surveys to determine the geographical distribution of air passenger trip origins at a suitable level of resolution and the issues that arise in developing air passenger trip generation models. The paper then presents a set of air passenger trip generation models that

were developed to support the current update of the Bay Area Regional Airport System Plan Analysis and discusses their strengths and limitations, as well as potential opportunities to improve the future utility of the models. The paper concludes by discussing the additional analysis needed to develop more detailed models and the potential transferability of these models to other regions.

## Determining the Regional Distribution of Air Passenger Trip Ends

Information on the regional distribution of air passenger trip ends is generally only available from air passenger surveys. These typically ask respondents to identify where they began their trip to the airport at some level of zonal detail, such as the zip code, and may also ask the nature of that location, such as a home, business, hotel, etc. Since respondents, particularly visitors to the area, may not know the zip code of their trip origin, it is becoming common practice to ask for the city and street address or a nearby street intersection or the block number of the street if respondents are reluctant to give an actual address. In the case of locations with identifiable names, such as hotels or businesses, the name is recorded and the street address is assigned later in the survey coding process. The trip origin locations can then be geocoded to latitude and longitude and subsequently assigned to any desired system of analysis zones using geographic information system software.

Since air passenger surveys are typically conducted by interviewing departing air passengers, the surveys generally obtain information on the airport ground access trip, including the trip origin. It is commonly assumed that the distribution of trip destinations for arriving air passengers is the same as the distribution of trip origins for departing passengers. While this may be true in the aggregate, there may be differences when looking at air passenger access and egress travel patterns by time of day or air party characteristics. For example, visitors on business trips arriving in the evening are likely to go directly to a hotel, whereas on their return they may leave for the airport from a business or other location they are visiting. Similarly, residents may begin their airport access trip from their workplace, but return to their home. These differences may have significant implications for ground access and egress mode use and are deserving of more study that they have been given to date.

### *Limitations of Air Passenger Survey Data*

The limited sample size of most air passenger surveys presents a challenge for using the survey results to determine the regional distribution of air passenger trip ends at the level of the travel analysis zones typically used in regional transportation planning. A typical air passenger survey may have a few thousand responses from passengers with ground origins in the region served by the airport, while most large metropolitan regions will typically be divided into more than a thousand travel analysis zones (sometimes significantly more). However, air passenger trip ends are not uniformly distributed across the travel analysis zones, so there will be many zones that do not have any responses in the survey. Of course, this does not mean that there are no air passenger trips from those zones. Furthermore, air travel parties often have more than one passenger, sometimes a large number, as in the case of tour groups or sports teams. If the number of passengers in a travel party responding to a survey is used to factor up the survey responses to annual passenger trips, this can greatly

overstate the number of annual trips from zones that happen to be the ground origin of large parties included in the survey sample.

Therefore simply factoring up the survey responses to give annual air passenger trips from each analysis zone is likely to introduce considerable distortion at the zonal level due to the limitations of the survey sample size and the effect of differing air party sizes in the sample. These distortions can be increased if air passenger trips are being estimated at a regional level from surveys performed at multiple airports, particularly at different points in time. Additional distortions can arise if the survey is not an unbiased sample of the air passenger population, due to the design of the survey sampling strategy. For example, the geographic distribution of air passenger trip origins is likely to vary by time of day, with access trips from homes predominating in the early morning and those from businesses later in the day. If the air passenger survey under-samples certain times of day this could introduce a bias into the resulting geographical distribution of air passenger trip origins. Survey sampling issues are discussed in more detail in a recent Airport Cooperative Research Program guidebook on the design and conduct of airport user surveys (Biggs et al. 2009).

Of course, the geographic distribution of air passenger trip ends given by an air passenger survey only reflects the distribution at the time of the survey. Future changes in regional development patterns will change the future trip distribution.

**Development of Air Passenger Trip Generation Models**

Air passenger trip generation models that predict the number of trips that begin or end in a given analysis zone as a function of the socioeconomic and other characteristics of the zone can address both distortions from survey sampling limitations and likely future changes in the geographical distribution of air passenger trip ends from forecast changes in regional development patterns and socioeconomic factors. Given the importance of this role, the lack of attention given to these types of air passenger trip generation models in the literature is surprising.

The topic was addressed in an early study for the National Cooperative Highway Research Program (Keefer 1966) but there has been very little work since. Keller (1966) considered the variation in air passenger trip generation rates in different counties in West Virginia, but did not correlate this with any socioeconomic factors. Picado (1994) examined trip generation rates in the Southern California region using air passenger survey data to estimate trip generation models based on population, household income, and employment at the zonal level, dividing the region into 65 analysis zones. These studies have been reviewed in more detail in a literature review performed a number of years ago for the Southern California Association of Governments (Gosling et al. 2003).

Various statewide and regional travel modeling studies have used the traditional four-step process of trip generation, trip distribution, mode choice, and travel assignment to model intercity or interregional trips that include air trips. A recent example is the modeling undertaken for the ridership forecasts for the California High-Speed Rail system (CSI et al. 2006). However, the trip generation models in these studies generally address trips by all modes and air trips are subsequently identified in the mode choice step. While in principle this process could

generate a geographic distribution of air passenger trip ends, the modeling implementation typically does not allow this to be readily done.

### *Application to the San Francisco Bay Area*

Starting in late 2008, regional planning agencies in the San Francisco Bay Area have been undertaking a study to update the Regional Airport System Plan Analysis (RASPA) and develop a Vision and Implementation Plan to meet the future airport needs of the Bay Area. Details of the current status of the study are available on the study website at http://www.regionalairportstudy.com. As part of the study, forecasts of future regional air travel demand were developed at a regional level. These forecasts distinguished between origin/destination (O/D) and connecting passengers and between domestic and international trips.

In order to allocate the forecast regional air travel demand to the airports in the region under various system development scenarios and analyze the ground access and egress travel associated with each scenario, it was necessary to develop an analysis procedure to calculate the proportion of the forecast regional air travel demand in a future year that will have trip ends within each regional travel analysis zone (TAZ). It was also desired that this procedure should reflect changes in the forecast regional distribution of population, households, and income. This required the development of trip generation models that could forecast the number of air passenger trips from a given analysis zone as a function of the zone's socioeconomic characteristics.

Separate trip generation procedures were developed for O/D air trips in domestic and international markets, for each of the following four market segments:

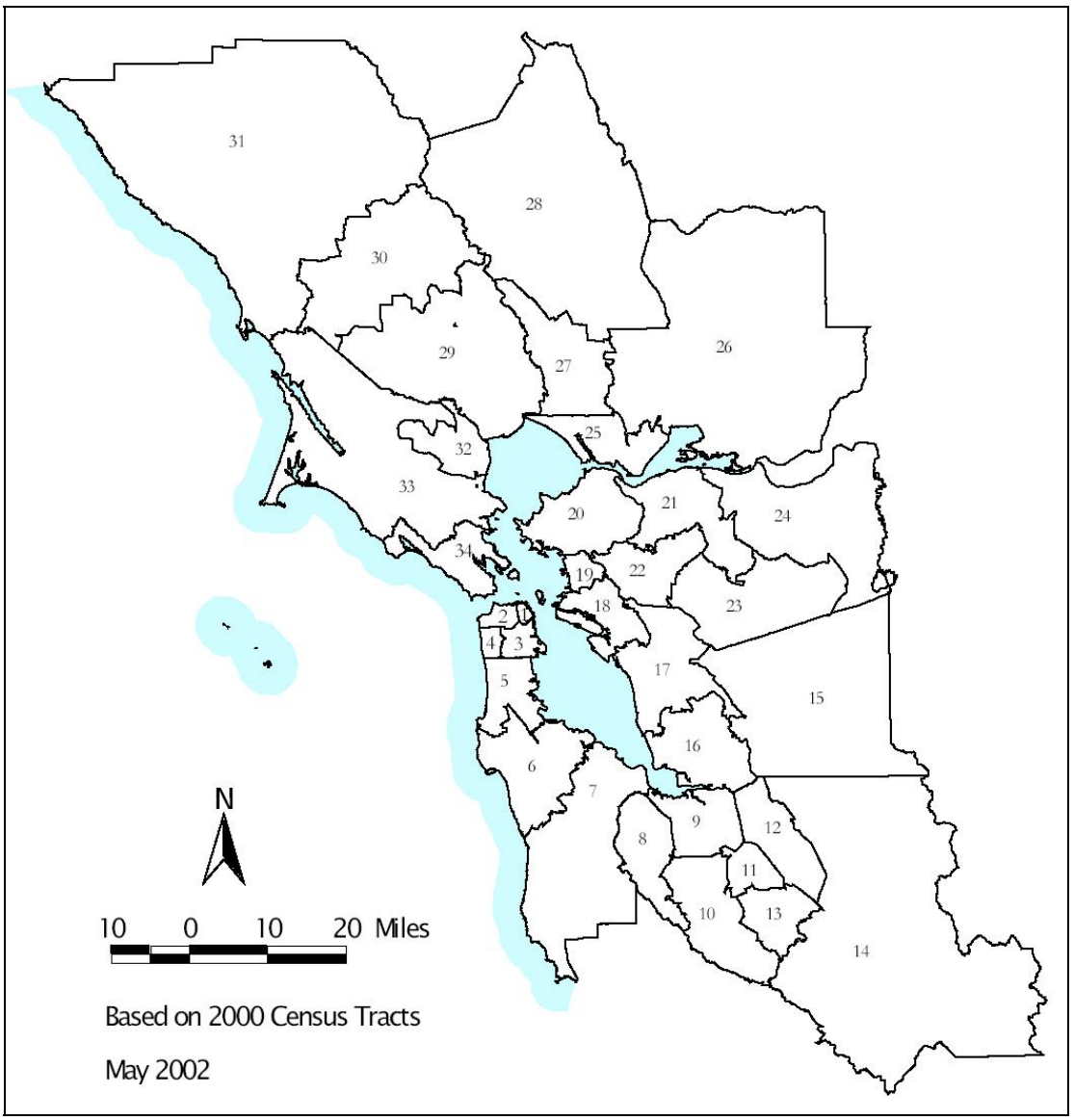
- Resident trips from home origins
- Resident trips from non-home origins
- Visitor trips from home origins
- Visitor trips from non-home origins.

However, formal trip generation models were only developed for resident trips from home origins. Visitor trips from home origins (i.e. visitors staying with family or friends) were assumed to be distributed in proportion to the resident trips from home origins. The distribution of resident and visitor trips from non-home origins, principally hotels in the case of visitors, was assumed to remain unchanged from the current distribution, as determined from air passenger survey data.

The first step in developing the trip generation procedures was to identify the regional distribution of domestic and international O/D air passenger trips by analysis zone for 2006. In the case of air passenger trips using Oakland International Airport (OAK) and San Francisco International Airport (SFO), this was obtained directly from the results of an airline passenger survey performed at the two airports by the Metropolitan Transportation Commission (MTC) in 2006. However, the third primary air carrier airport in the region, Mineta San José International Airport (SJC), was not included in the 2006 survey and the most recent air passenger survey for SJC was performed for MTC in 2001/2002. It was assumed that the geographic distribution of trip ends of air passenger trips using SJC did not change significantly from 2001/2002 to 2006, although of course the total number of such trips changed.

The results of the air passenger surveys at each airport were factored up to the total number of domestic and international O/D passengers at each airport in 2006 in each of the four market segments and summed to give the regional total of domestic and international air passenger trip ends in 2006 by market segment in each analysis zone.

The trip origins given by the survey responses were coded to the system of 1,454 TAZs used by the MTC for regional transportation modeling. However, at this level of geographical resolution, the resulting data shows too much variance from TAZ to TAZ due to the limitations of the survey sample size and air party size effects discussed above. Therefore for the purpose of model estimation, the estimated annual air passenger trips were aggregated to a system of 34 regional superdistricts shown in Figure 1.




Source: Metropolitan Transportation Commission

**Figure 1. Bay Area Travel Analysis Superdistricts**

Socioeconomic data for each of the superdistricts and the underlying TAZs for 2006 and future years were derived by MTC from regional socioeconomic forecasts prepared by the Association of Bay Area Governments (ABAG 2006, MTC 2007). The ABAG forecasts express household income in constant 2005 dollars, but MTC converted these values to constant 1989 dollars for consistency with its travel demand models. The TAZ level data were obtained from an unpublished MTC data file allocating the ABAG socioeconomic forecast data to TAZs.

Exploratory analysis was undertaken to examine the relationship between air passenger trip ends, population and average household income. Since the trip generation relationships were to be applied across analysis zones with widely varying populations, it was decided to base the analysis on trip propensity, defined as the number of annual air passenger trips (to and from the airport) per person resident in the zone. This avoids distortions from constant terms that are independent of the population of a zone. The resulting trip propensities for resident domestic trips from home origins for each of the 34 superdisticts are shown in Figure 2 in relation to the average household income for the zone.

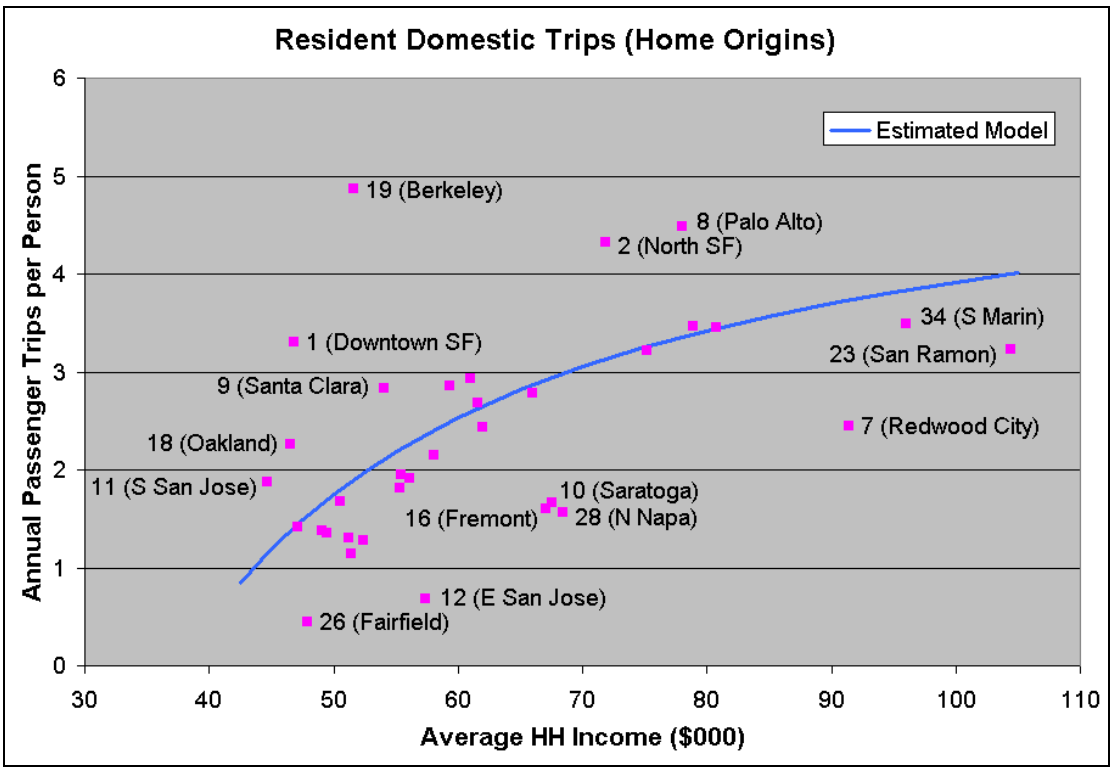



**Figure 2. Trip Propensity by Superdistrict**

It can be seen from Figure 2 that there is significant variation in trip propensity for superdistricts with similar average household income, although there is a general relationship of increasing trip propensity with increasing average household income. There also appear to be a number of distinct outliers in the data, particularly Superdistrict 19 (Berkeley) and Superdistricts 7, 10, 12, 16 and 28. Figure 2 also shows the estimated relationship between trip propensity and household income given by the trip generation model discussed below.

A detailed discussion of possible causes for the data outliers shown in Figure 2 is beyond the scope of this paper. However, it may be worth noting that the superdistrict giving the highest outlier (Berkeley) contains a major university, the University of California at Berkeley. It could be expected that university students would report a fairly low household income but would tend to make more air trips than usual for that income level. One of the other superdistricts giving a high outlier, Superdistrict 8 (Palo Alto), contains Stanford University. Other superdistricts with high outliers include downtown San Francisco and north San Francisco, both areas that are likely to have atypical household compositions, with a high proportion of single or childless adults, which could distort the relationship between household income and trip propensity. While the superdistricts with outliers well below the general pattern are predominantly residential, that is also true for many of the superdistricts that are not outliers, so this alone is not an adequate explanation for the differences in trip propensity. Furthermore, some of the differences in average household income at the superdistrict level are surprising (for example the difference between Superdistrict 7 (Redwood City) and Superdistrict 9 (Santa Clara)), so this may be more a reflection of variations in the composition of subareas within the superdistricts rather than differences between them.

A trip generation model for domestic home-origin resident trips was estimated from the observed 2006 superdistrict trip propensity and 2006 average household income, giving the following relationship:

$$\text{Pax/Pop} = 5.651 - 156.1 / (\text{AHHI} - 10)$$
$$\textit{(11.9)} \quad \textit{(-6.9)}$$
$$R^2 = 0.666 \qquad n = 26$$

where Pax = Annual air passenger trips to and from the zone
Pop = Zone population
AHHI = Average household income in zone (in thousand 1989 dollars)
t-statistics for the coefficient estimates are shown in parentheses

The functional form was chosen to reflect the decreasing marginal increase in trip propensity with increasing average household income suggested by the data shown in Figure 2 and the estimation excluded the outlier data for Superdistricts 1, 7, 10, 12, 16, 19, 26 and 28. Several different model specifications were tested, with the average household income adjusted up or down by fixed amounts, and the best fit to the data was found when the average household income was reduced by $10,000. While it seems reasonable that trip propensity is influenced by discretionary income above the amount required to meet essential household expenditures (such as housing costs and food), it is likely that the particular adjustment to the average household income found to give the best fit to the data is simply an artifact of the rather simple functional form adopted for the relationship, so it would be misleading to assign an economic interpretation to the specific amount of the adjustment.

While the estimated model coefficients were statistically significant, the data showed a wide degree of scatter about the estimated relationship, as indicated by Figure 2. Some of this scatter is undoubtedly due to limitations of the survey sample

size as well as the procedure for combining the results from the three surveys, while the remainder of the scatter may be due to factors not included in the model. The resource limitations of the study did not allow further investigation of these possible causes although this would be a useful issue to explore in future work. In order to account for this scatter in applying the model at the TAZ level or for future years, an adjustment factor was computed for each superdistrict that corrected the model results to give the trip propensity obtained from the air passenger survey results (i.e. the values of the dependent variable used in the model).

A similar process was followed for international trips. The corresponding trip generation model for international home-origin resident trips was estimated from the observed 2006 superdistrict trip propensity and the 2006 average household income, giving the following relationship:

$$\text{Pax/Pop} = 0.96 - 35.1 / \text{AHHI}$$
$$\textit{(8.2)} \quad \textit{(-5.4)}$$
$$R^2 = 0.566 \qquad n = 24$$

where the variables are defined as in the model for domestic trips and the t-statistics of the coefficient estimates are shown in parentheses. Several different model specifications were tested, with the average household income adjusted up or down by varying amounts, but the best fit to the data was found using the unadjusted average household income. Superdistrict adjustment factors were calculated as for domestic trips.

### *Estimating Trip Generation at the TAZ Level*

Because the trip generation models express the ratio of air passenger trips to zonal population in terms of the average household income in the zone, these relationships are independent of the size of the zone (at least in principle). Thus they can be used to develop projections of air passenger trips by TAZ based on the projected population and average household income for the TAZ and the relevant superdistrict adjustment factor. An additional adjustment factor was calculated to ensure that the total projected trips for the TAZs in each superdistrict summed to the superdistrict total.

## Comparison to the Results of Earlier Studies

The time that has elapsed between the earlier studies mentioned above and the current analysis makes comparison of the results of the various studies of limited value, since air travel propensity has changed over time due to factors other than household income, most notably changes in the real price of air travel. Furthermore the analysis undertaken by Keefer (1966) primarily addressed differences in air travel trip generation at the regional level between different metropolitan areas rather than differences within a given metropolitan area, although Keefer does present curves showing the change in trip generation rates with distance from the airport derived from household survey data for ten regions. These trip generation rates decline sharply with distance from the airport, but above about 8 miles from the airport they stabilize at around one weekday person trip per thousand population for trips *to* the

airport. On an annual basis this would be equivalent to about 0.7 trips to and from the airport per person. While this is significantly below the trip propensity values shown in Figure 2, Keefer's data was from surveys performed between 1958 and 1963, when air travel propensity would have been substantially less than 2006 levels.

Picado (1994) estimated a model for resident airport access trips to three of the principal Southern California airports using air passenger survey data for 1987. The model did not distinguish between home origin trips and trips from other origins and assumed a linear relationship between the annual access trips per person from each analysis zone and the per capita personal income in the zone. A dummy variable was used to increase the access trip propensity for zones close to Los Angeles International Airport (LAX) by a constant amount to reflect the higher trip generation observed for those zones in the survey data. The personal income data used in the model was obtained from the 1990 census data and thus was expressed in 1989 dollars, the same as the current analysis, although of course this reflected 1989 income levels, not 2006 income levels. Applying the estimated relationship to household income instead of personal income using the average number of persons per household for the Bay Area in 2006 (2.75) and ignoring the constant term for the proximity to LAX gave a trip propensity of about 1.0 annual trips both ways per person for a zone with an average household income of $50,000 (in constant 1989 dollars) and about 2.3 annual trips both ways per person for a zone with an average household income of $100,000. While these values are somewhat less than the trip propensity relationship shown on Figure 2, the enplanement data from which the Southern California model was estimated was for 1987, when air travel propensity would have been less than 2006 levels.

**Strengths and Weaknesses of the Current Models**

The two trip generation models for domestic and international home-origin trips by residents of the region reflect the diminishing growth in travel propensity with increasing average household income shown by the data, the model coefficients have a good statistical fit, and the overall models explain between 57 and 67 percent of the variance in the zonal trip propensity data (excluding the outlier data points). The models thus appear to provide a reasonable basis for forecasting future changes in trip generation at a superdistrict level as a result of changes in population and average household income, as well as estimating trip generation at the TAZ level.

However, the current models have a number of limitations. There is a large amount of unexplained variance in the data, which has been reflected in the superdistrict adjustment factors. It is not clear how much of this variance is due to factors that are not included in the models, how much is due to the aggregation involved in estimating the models at the superdistrict level, and how much is due to sampling errors in the air passenger survey results. To the extent that the unexplained variance is due to characteristics of the population in each superdistrict, applying the superdistrict adjustment factors is appropriate. However, if the unexplained variance is due to air passenger survey sampling error or bias in the expansion of the survey results, then it would be more accurate to apply the models without the superdistrict adjustment factors, since the unexplained variance is in fact an error around the model relationship.

Since the superdistricts are quite large, there is likely to be considerable variation in the average household incomes of subareas within each superdistrict. Because of the non-linear relationship between trip propensity and average household income, this is likely to lead to aggregation errors whereby superdistricts with a mix of higher and lower income areas will tend to have a lower average trip propensity than a more homogeneous superdistrict with similar average household income.

Another difficulty with the non-linear relationship used in the trip generation models is that at low enough average household incomes, the models can generate negative trips. While average household incomes at the superdistrict level do not reach these levels, since every superdistrict is large enough to contain some higher income areas that help raise the average household income, this is not always true at the TAZ level. The air passenger survey results show that some air passengers report very low household incomes. Thus even people with very low incomes make some air trips, and so the relationship between trip propensity and household income implied by the models may not hold true at low levels of household income. It was therefore necessary to define a lower bound on trip propensity that was applied to those TAZs with average household incomes that would result in negative or unrealistically low trip propensity levels.

## Opportunities for Future Improvements in the Models

Although the models described in this paper met the current needs of the RASPA study, there are a number of ways in which future work could improve their accuracy and utility. The most obvious is to obtain a better understanding of the reasons for the wide variation in trip propensity for a given level of average household income across the different superdistricts. In particular it would be helpful to know the extent to which this is simply a result of sampling error in the air passenger survey results and the expansion of the survey results to annual passenger trips, the extent to which this is the result of aggregating TAZs with very different household income distributions into much larger superdistricts, and the extent to which it reflects factors not included in the models. A better understanding of these issues would not only help to decide whether to apply superdistrict adjustment factors to the model results, but also lead to improved model specifications that can address factors influencing trip propensity other than household income.

Another area where additional work would be helpful would be to extend the models to predict trip generation from non-home origins. This will require the identification of appropriate independent variables, such as the number of hotel rooms or employment by different sectors of the economy, as well as assembling the relevant data at the TAZ level. Since these data are not typically available at this level, this could involve a significant amount of data processing and integration. However, the results of such an effort could not only support better air passenger trip generation models that would be responsive to projections of future changes in development patterns but could also contribute to better urban travel demand models.

## Conclusions

The air passenger trip generation models presented in this paper show an intuitively reasonable relationship between trip propensity and household income and

provide a consistent basis for assigning forecast air passenger trips from home origins to regional ground analysis zones in a way that can take account of projected future changes in population and income. The models show good statistical fit to the data on which they were estimated, although there is a large amount of unexplained variation in trip propensity at the superdistrict level and a number of the larger outliers were excluded from the model estimation.

The models described in this paper appear to provide a good starting point for future work to better understand the causes of the apparent variation in trip propensity across superdistricts with similar levels of average household income and the development of more robust trip generation models of air passenger trips from home origins. In addition, there is a need for corresponding development of trip generation models of air passenger trips from non-home origins.

### *Potential Transferability of Trip Generation Models*

Whether the air passenger trip generation models described in this paper, or future enhancements to those models, can be applied in other regions without needing to estimate new models depends on two considerations: whether the models correctly reflect the effect on trip propensity of local differences in household income and possibly other factors, and whether the relationship between trip propensity and the explanatory variables in the model applies in the other regions. Since the total number of air passenger trips in a given region can usually be determined from airport traffic statistics and airline data reported to the U.S. Department of Transportation (or similar agencies in other countries), as long as the general relationship between trip propensity and the explanatory variables is consistent, the trip propensity rates estimated from the models can always be factored to give the correct levels of trip generation. However, this represents another aspect that could benefit from future work to investigate the extent to which the underlying trip generation relationships vary from region to region.

## Acknowledgments

The development of the air passenger trip generation models for the San Francisco Bay Area described in this paper was performed as part of a subcontract from SH&E, an ICF International Company, for the Bay Area Metropolitan Transportation Commission in support of the Regional Airport System Plan Analysis. Eric Deichmann of SH&E provided valuable assistance with the analysis of the air passenger survey data used in developing the models. Several helpful suggestions for improving the paper were made by anonymous reviewers of the initial draft of the paper.

## References

Association of Bay Area Governments (ABAG). (2006). *Projections 2007: Forecasts for the San Francisco Bay Area to the Year 2035*, Oakland, Calif., Dec.

Biggs, D. C., Bol, M. A., Baker, J., Gosling, G. D., Franz, J. D., and Cripwell, J. P. (2009). *Guidebook for conducting airport user surveys*, ACRP Report 26, Airport Cooperative Research Program, Transportation Research Board, Washington, D.C.

Cambridge Systematics, Inc. (CSI), with Mark Bradley Research and Consulting. (2006). *Bay Area/California high-speed rail ridership and revenue forecasting study: interregional model system development*, Prepared for the Metropolitan Transportation Commission and the Calif. High-Speed Rail Authority, Oakland, Calif., Draft report, Aug.

Gosling, G. D., with Cambridge Systematics, Inc. and SH&E, Inc. (2003). *SCAG Regional Airport Demand Model: literature review*, Preapred for the Southern California Association of Governments, Los Angeles, Calif., Jun.

Keefer, L. E. (1966). *Urban travel patterns for airports, shopping centers, and industrial plants*, NCHRP Report 24, National Cooperative Highway Research Program, Highway Research Board, Washington, D.C.

Keller, C. R. (1966). *Commercial air travel in West Virginia: trip generation and attraction*, Master of Science thesis, West Virginia Univ., Morgantown, West Virginia.

Metropolitan Transportation Commission (MTC). (2007). *Superdistrict and county summaries of ABAG's Projections 2007: 2000-2035 – data summary*, Oakland, Calif., Aug.

Picado, R. (1994). *Air trip generation model for Southern California*, CE299 report, Submitted for partial fulfillment for the requirements for the M. Eng. degree, Dept. of Civil Engineering, Univ. of Calif., Berkeley, Calif.

# Specification and Data Issues with Airport Ground Access Mode Choice Models

Geoffrey D. Gosling, Ph.D., M.ASCE[1]

[1]Aviation System Consulting, LLC, 805 Colusa Avenue, Berkeley, CA 94707;
PH (510) 528-8741; FAX (510) 528-8745; e-mail: gdgosling@aol.com

## Abstract

Ground access mode choice models play a key role in airport ground transportation and airport system planning studies. The paper presents the findings of a recent Airport Cooperative Research Program Synthesis study that reviewed the current state of practice of airport ground access mode choice models, and discusses the range of model specifications adopted in recent studies. The paper examines the associated data issues that arise with different model specifications and discusses the implications for air passenger and other airport user surveys undertaken to support the development of such models. The paper concludes with a set of best practice recommendations for future development of airport ground access mode choice models.

## Introduction

Ground access mode choice models predict how air passengers and other airport users will choose which ground transportation services to use for their airport access trips in response to changes in the characteristics or availability of those services, as well as changes in the characteristics of the travelers using the airport. These models thus form an essential analysis tool for airport ground transportation and airport system planning studies. In addition to their use for airport ground transportation studies these models can also form an important component of airport choice models, reflecting the role of accessibility in airport choice. Because of the large number of transportation modes and services typically available at major airports, as well as the influence of a wide range of traveler characteristics, these models can be extremely complex.

In order to help airport planners and managers better understand the issues that arise in the development and use of these models, a recent Airport Cooperative Research Program (ACRP) Synthesis study (Gosling 2008) reviewed the current state of practice of airport ground access mode choice models and examined the range of model specifications adopted in recent studies. This paper summarizes some of the findings of that study, examines the associated data issues that arise with different model specifications and discusses the implications for air passenger and other airport

user surveys undertaken to support the development of such models. Based on the findings of the Synthesis study, the paper presents a set of best practice recommendations for future development of airport ground access mode choice models.

The discussion in the paper focuses primarily on air passenger mode choice models, because these are by far the most common application. However, the general principles would also apply to models of airport employee access mode choice. The discussion also refers to these models as airport access models, although airport ground transportation travel involves both access and egress trips. This reflects the prevailing practice of estimating these models on data for airport ground access travel, and assuming that the use of ground transportation modes for egress trips is symmetrical. While there are reasons to believe that this may not be entirely true, there has been very little work to date that explores this issue and no examples of airport ground egress mode choice models have been identified in the literature.

## Current State of Practice with Airport Ground Access Mode Choice Models

Given the importance of understanding air passenger airport ground access mode use it is not surprising that there have been a fairly large number of studies over the years that have developed air passenger ground access mode choice models. One of the earliest such efforts was undertaken in the early 1970s (Ellis et al. 1974). This study used a multinomial logit (MNL) model, as did several other studies that developed air passenger or intercity terminal access mode choice models over the next ten years (e.g. Sobieniak et al. 1979, Spear 1984, Harvey 1986). However, by the mid 1980s it was becoming recognized that some of the limitations of the MNL model could be addressed through the use of nested logit (NL) models (Ben-Akiva and Lerman 1985). One of the first applications of NL models to airport ground access mode choice was undertaken as part of a study of surface access to London Heathrow Airport (Howard Humphreys and Partners 1987), followed shortly thereafter by a study by Harvey (1988) that used a NL structure to develop an integrated model of airport choice and ground access mode choice for the San Francisco Bay Area. Subsequent air passenger ground access mode choice models developed for Boston Logan International Airport (Harrington et al. 1996), Portland International Airport (Portland Metro 1998), and the South East and East of England Regional Air Service (SERAS) study (Halcrow Group 2002b) used a nested structure, while other studies continued to use MNL models to represent air passenger ground access mode choice. In addition to models that have exclusively addressed airport access mode choice, a number of recent studies have used NL models to represent air passenger airport choice, with airport ground access mode choice as a lower level nest (e.g. Pels et al. 2003). However, these models generally only include a single-level nest for the airport ground access mode choice process, and thus are equivalent to MNL models from the perspective of ground access mode choice.

As part of the ACRP Synthesis study, a survey was undertaken of airport authorities, regional and state planning agencies, airport consulting firms, selected universities, and other relevant organizations to identify airport ground access mode choice models that had been developed for specific studies but not reported in the published literature. 34 respondents to the survey indicated that they had sponsored,

undertaken, or participated in studies over the previous ten years that had made use of airport ground access mode choice models, and identified 52 specific studies. Based on these responses and ground access mode choice models that had been identified in the course of prior research, detailed information was obtained on a number of recent models that had been developed for Hartsfield-Jackson Atlanta International Airport (Atlanta Regional Commission 2006), Chicago O'Hare and Midway Airports (Resource Systems Group 2004), Miami International Airport (ICF Kaiser 1995), Oakland International Airport (U.S. FTA 2002), Mineta San José International Airport (Dowling Assoc. 2002), and Toronto Pearson International Airport, Canada (Halcrow Group 2002a).

The ACRP Synthesis study included a detailed review of the structure and specifications of nine airport ground access mode choice models developed between 1995 and 2004, as shown in Table 1. Three of these models used an MNL structure, four used an NL structure, one used a combination of NL and MNL structures for different travel segments, and one used a binomial logit structure (an MNL model with only two choices).

**Table 1. Selected Recent Airport Ground Access Mode Choice Models**

| Airport/Study | Year Developed | Model Structure |
|---|---|---|
| Hartsfield-Jackson Atlanta International Airport | 2002 | Nested logit |
| Boston Logan International Airport | 1996 | Nested logit and multinomial logit |
| Chicago O'Hare International and Midway Airports | 2004 | Nested logit |
| Miami International Airport | 1995 | Nested logit |
| Oakland International Airport | 2001 | Multinomial logit |
| Portland International Airport | 1997 | Multinomial logit |
| Mineta San José International Airport | 2002 | Multinomial logit |
| Toronto Lester B. Pearson International Airport | 2002 | Binomial logit |
| United Kingdom SERAS Study | 2002 | Nested logit |

Source: ACRP Synthesis 5 (Gosling 2008)

### *Modes Included in the Models*

The modes included in the nine models are shown in Table 2, together with the number of models in which each mode was included. Private vehicle, taxi and transit were the only modes that were included in all nine models. Seven of the models distinguished between passengers dropped off by private vehicles and those using private vehicles parked for the duration of the trip, with only one model identifying the location where the vehicle was parked. Only two models included limousine as a separate mode and three combined this with taxi (the other four models may have done so by default without stating this). The representation of transit services varied across the models, with three distinguishing between rail and

bus services, three only considering rail transit, and three treating all transit services as a one mode. Seven of the models included rental car as a choice for non-resident passengers. The two models that did not include rental car assumed that the decision to rent a car was based on travel needs other than the airport access trip and therefore excluded these passengers from the mode choice decision process.

**Table 2. Modes Included in Recent Airport Access Mode Choice Models**

| Ground Access Mode | Number of Models |
|---|---|
| Private vehicle – drop-off | 7 |
| Private vehicle – park (location unspecified) | 6 |
| Private vehicle – park (short-term for passenger drop-off) | 1 |
| Private vehicle – park for trip (on airport) | 1 |
| Private vehicle – park for trip (off airport) | 1 |
| Private vehicle – drop-off and parking combined | 2 |
| Rental car | 7 |
| Taxi (including combined with limousine – 3 models) | 9 |
| Limousine (as separate mode) | 2 |
| Shared-ride van | 3 |
| Scheduled airport bus | 3 |
| Shared ride van and scheduled airport bus (combined) | 1 |
| Express bus from off-airport terminal | 4 |
| Transit (all services) | 3 |
| Transit – rail | 6 |
| Transit –bus | 2 |
| Transit –bus and intercity coach (combined) | 1 |
| Airport express train | 2 |
| Intercity rail/coach links | 1 |
| Ferry (Boston water shuttle) | 1 |
| Hotel shuttle bus/van | 4 |
| Hotel shuttle bus/van and charter coach (combined) | 1 |
| Other private modes | 1 |
| Other public modes | 1 |

Source: ACRP Synthesis 5 (Gosling 2008)

The inclusion of particular modes in a few of the models reflected the availability of those modes at the airports in question, such as the ferry service at Boston Logan International Airport. The United Kingdom SERAS study treated hotel shuttle bus and charter coach as a single mode. Although these are functionally quite different, it can be argued that the users of these two modes are essentially captive riders as a result of travel decisions separate from the airport access trip, much like rental car users.

### *Explanatory Variables*

Apart from the modes included in the model, the other major factor in the specification of logit discrete choice models is the explanatory variables included in the utility functions for each mode. There was much less commonality across the nine models, due both to differences in the way that variables were defined and to the fact that some variables were included in the utility functions for some modes but not others. The range of different variables and the number of models in which they were used is shown in Table 3. The table distinguishes between continuous variables and dummy variables (which take the value one or zero depending on whether the characteristic is present or not). It can be seen that with the exception of whether a traveler's employer is paying for the trip or the access trip begins at a residence, the dummy variables are really substitutes for continuous variables (such as the air party size) for which reasonable parameters could not be estimated. Only three of the models considered whether the travel party had a significant amount of baggage.

All models included travel time and cost variables, but the way that the travel time variables were defined varied widely across the models. In some cases separate variables were used for different modes and in other cases the same variable was used for all modes. Using a separate variable for different modes allows the estimated parameters to reflect a different perceived disutility for the same travel time component on different modes. Similarly, using separate variables for different components of the travel time, such as waiting time and in-vehicle time, allows the estimated parameters to reflect a different perceived disutility for different components of travel time. The same effect is achieved by using different weights for different travel time components, although in the first case the model estimation determines the appropriate weights from the observed choice behavior.

One model (for Boston Logan International Airport) used different cost variables for private vehicle costs and taxi and transit fares. While the disutility of a dollar should be same irrespective of the mode used, there is the difficulty of knowing what the travelers considered a private vehicle trip to cost, particularly if the vehicle was driven by someone else to drop them off at the airport. For this reason the model for Portland International Airport used a different variable for the operating cost of private vehicles used to drop off passengers. Similarly, while the taxi fare can be calculated using the taximeter rate, this depends on the distance and the taxi driver may take a longer route to save time or some taxi operators may charge a different fare for airport trips. There is also the issue of whether to include a tip in the cost.

## Data Issues with Varying Model Specifications

The ability to explore a wide range of potential model specifications in order to develop airport access mode choice models that both fit the observed data well and are responsive to planning and policy issues depends on having detailed data on the actual mode use of a large enough sample of air passengers as well as the air party characteristics of those travelers.

Since the travel times and costs involved in using fixed route modes depends on how the travelers access those modes, the mode use of each party needs to include details of those access trips, including which stations or stops were (or would be) used and the mode used to access the fixed route mode.

**Table 3. Variables Included in Recent Airport Access Mode Choice Models**

| Explanatory Variable | Number of Models |
|---|---|
| ***Continuous variables*** | |
| Off-peak highway time | 1 |
| Travel time (private vehicle) | 2 |
| In-vehicle time (all modes) | 3 |
| In-vehicle time (transit) | 1 |
| In-vehicle time (rail transit) | 2 |
| In-vehicle time (bus transit) | 2 |
| In-vehicle time plus walk time | 1 |
| Total travel time (all modes) | 1 |
| Weighted travel time (access, transfer and wait time weighted differently from in-vehicle and egress time) | 1 |
| Wait time or headway | 5 |
| Wait time plus zone terminal time | 1 |
| Waiting time plus walk time (out-of-vehicle time) | 1 |
| Walk time | 1 |
| Walk distance | 2 |
| Private vehicle access time to transit | 1 |
| Number of transfers | 1 |
| Interchange penalties (vary by type of interchange) | 1 |
| Drop-off driver time | 1 |
| Driving distance | 1 |
| Travel cost (all modes) | 8 |
| Private vehicle cost | 1 |
| Drop-off vehicle operating cost | 1 |
| Transit fare | 1 |
| Taxi fare | 1 |
| Household income | 3 |
| ***Dummy variables*** | |
| Air party size (more than a certain number) | 1 |
| One-person travel party | 1 |
| Luggage/checked bags | 3 |
| Employer pays cost of trip | 1 |
| Flights/year (more than a certain number) | 1 |
| Household income (over a certain amount) | 1 |
| Non-residence trip origin | 1 |
| Use of intermediate station | 1 |

Source: ACRP Synthesis 5 (Gosling 2008)

In the case of private vehicles parked at an on-airport or off-airport lot, it is necessary to know which lot was used since this determines the travel times and costs involved, which can be calculated if the lot used is known. Using the wrong times and costs in the model estimation will reduce the accuracy of the resulting model.

Developing model specifications that provide a fairly detailed representation of the full range of available modes and ancillary decisions (such as how to access fixed route modes and where to park) and that considers different components of the travel time involved in using those modes requires reasonably accurate data on each of those components as well as the associated costs and other level of service factors. Assembling these data for the wide range of modes and options typically available at a major airport can involve a significant amount of work. Since highway travel times and transit frequency can vary significantly by time of day, while travel times on rail services are generally more consistent, one important consideration is whether to develop different modal level of service data for different time periods. This may depend on the ability of the relevant regional transportation planning agencies to produce highway and transit network travel times for different time periods.

**Implication for Air Passenger and Airport Employee Surveys**

Much of the data on air party and airport employee characteristics and ground access mode use that are required for estimating airport ground access mode choice models can only be obtained from air passenger and airport employee surveys. Therefore the way that these surveys are designed and performed has a critical influence on the accuracy and reliability of the resulting models.

Guidance on planning and conducting airport user surveys is provided in a recent ACRP report (Biggs et al. 2009). However, beyond general issues that need to be considered in planning and performing such surveys, such as how and where to perform the survey and ensuring a representative sample of survey respondents, there are particular considerations that arise in planning surveys that will generate data for use in estimating airport access mode choice models. Unfortunately, these surveys are often planned and undertaken well before the analysts who will develop the models become involved, and thus may impose unavoidable limitations on the resulting models. Therefore when surveys are being planned that may be or will be used to develop airport access mode choice models it is highly desirable to involve the likely model developers in the design of the survey.

The first consideration is to obtain detailed information on the composition of air travel and ground access parties. It is well understood that many air passengers travel in groups, and their access mode choice depends on the number of passengers in a group, since this affects the relative costs of different modes. For modes such as taxi the cost is independent of the number of passengers in the group, while for modes such as transit the cost varies with the number in the group. However the distinction between the air travel party and the ground access party is less well understood. Although for most air passengers this is the same thing, in some cases travelers may come to the airport separately and meet at the airport to take an air trip, while others may travel to the airport together but take different flights, for example attendees at a conference who share a taxi to the airport but take different flights home. Depending on the wording of the questions it may not be possible to

distinguish these different cases. Care should be taken to distinguish between travel groups who made their mode choice decision jointly and those who made independent decisions but just happened to be riding in the same vehicle, such as different travel parties using a shared-ride van service.

An important aspect of this that arises in survey design is whether to attempt to obtain survey responses from all passengers on a sampled flight or only obtain one response from each air party. In the former case, questionnaire design needs to allow multiple responses from the same air party to be identified. In the latter case, question wording needs to allow for the fact that members of an air party may have traveled to the airport separately, or all members of the air party may not have started the ground access trip from the same location as the respondent.

The survey needs to identify the ground origin of the access trip in sufficient detail to determine the costs and travel times involved in using different modes. In general this will require ground origin locations at the level of the street address or a nearby street intersection. These locations can then be geocoded to the system of travel analysis zones used by the regional transportation planning agency to estimate highway distances and travel times and transit network service characteristics and fares. Consideration needs to be given to identifying the location of intermediate stops on the trip to the airport where a mode change occurs, such as the train station or bus stop where the travel party first boards the primary airport access mode.

It is of course necessary to ask which mode or modes the travel party used to access the airport. Since respondents may use different terminology to describe a given mode, it may be necessary to ask clarifying questions, show pictures of different types of vehicles, or ask for the name of the service used. The duration of their air trip is a critically important factor for residents of the region served by the airport because it determines the cost of parking a car at the airport during the trip.

Obviously the survey needs to ask about any air party characteristics that may be used as explanatory variable in the model, including the amount of baggage and traveler income level. Both have been shown to be important factors in air passenger access mode choice but appropriate question wording requires care. Bags come in all shapes and sizes and there is clearly a significant difference between a shoulder bag and a large suitcase in terms of the ease of use of different modes. It is common to ask how many bags were checked by members of the air party, with the assumption that checked bags tend to be larger or heavier than those carried on as cabin baggage. However, restrictions on the number of bags that can be taken through security screening has complicated this aspect, and a relatively small bag may be checked because the traveler has too many to take through security screening.

Air passenger surveys commonly ask respondents their household income by showing a card with income ranges. It has been found that respondents are generally more willing to select a particular range than give a specific number. Household income is generally sought rather than the respondent's personal income, which is typically combined with that of other members of the household in ways that could make it a poor indicator of behavior. However, in this case it is advisable to obtain information on household composition. A traveler living alone on an income of $60,000 has a very different level of disposable income from a family of four with the same household income.

## Conclusions

Developing accurate and reliable airport ground access mode choice models requires careful attention to assembling the necessary data on airport traveler characteristics, mode use, and levels of service of the various access modes. The current state of practice, as illustrated by models developed over the past 15 years for a wide variety of airports, suggests that nested logit models provide a more detailed representation of the traveler mode choice behavior than multinomial logit models. However, there has been no consistency on how explanatory variables are represented in the models, or even which variables to include.

### *Best Practice Recommendations*

Based on the review of the current state of practice with airport access mode choice models, it appears that such models should generally use a nested structure in order to be able to properly reflect such aspects as the mode of access to fixed route services and alternative airport parking options. It may also be desirable to group door-to-door services, including taxi, limousine and shared-ride van, in the same nest and fixed route public transportation services, including scheduled airport bus services, in the same nest. Modes such as rental car or charter bus, the use of which results from travel decisions that are determined by other factors than the relative level of service of airport access modes, should not be included in the mode choice model directly, but their use analyzed separately.

While there is no consistent approach to the way that explanatory variables have been defined in recent models, the experience with these models suggests that the different components of travel time need to be assigned separate weights, either by weighting the variables or, preferably, by using separate variables with their own estimated parameters. Although household income was used as an explanatory variable in only a few of the models, it is self-evident that traveler income will affect mode choice decisions, and so this variable should be incorporated in an appropriate way. Rather than including this as a separate variable, it makes more sense to express costs as a function of household income adjusted for household size. Further research is needed to determine the most appropriate functional form, but in the meantime dividing costs by the household income per person would provide a better approach than ignoring income entirely or using household income as a separate variable.

## References

Atlanta Regional Commission. (2006). *Travel demand model documentation*, Atlanta, Ga., May.

Ben-Akiva, M., and Lerman, S. R. (1985). *Discrete choice analysis: Theory and applications to travel demand*, The MIT Press, Cambridge, Mass.

Biggs, D. C., Bol, M. A., Baker, J., Gosling, G. D., Franz, J. D., and Cripwell, J. P. (2009). *Guidebook for conducting airport user surveys*, ACRP Report 26, Airport Cooperative Research Program, Transportation Research Board, Washington, D.C.

Dowling Associates, Inc. (2002). *San Jose International Airport transit connection ridership*, Final report, Prepared for San Jose International Airport, Lea+Elliott and Walker Parking, Oakland, Calif., Jun.

Ellis, R. H., Bennett, J. C., and Rassam P. R. (1974). "Approaches for improving airport access." *Transp. Eng. J.*, ASCE, 100(TE3), 661-673.

Gosling, G. D. (2008). *Airport ground access mode choice models*, ACRP Synthesis 5, Airport Cooperative Research Program, Transportation Research Board, Washington, D.C.

Halcrow Group Limited with Cansult Ltd. (2002a). *Air rail link from Lester B. Pearson International Airport to Union Station: Revenue & ridership study*, Report T8080-01-1213, Prepared for Transport Canada, Ottawa, Canada, final report, May.

Halcrow Group Ltd. (2002b). *SERAS surface access modelling*, Prepared for the U.K. Dept. of Transport, Local Government and the Regions, South East and East of England Regional Air Services Study, London, England, Jul.

Harrington, I. E., McClennen, J., Pereira, E., and Wang, C.-Y. (1996). *Summary of people mover study passenger mode choice models*, Draft memorandum, Central Transportation Planning Staff, Boston, Mass., May 17.

Harvey, G. (1986). "A study of airport access mode choice," *J. Transp. Eng.*, ASCE, 112(5), 525-545.

Harvey, G. (1988). *ACCESS: Models of airport access and airport choice for the San Francisco Bay Region – Version 1.2*, Report prepared for the Metropolitan Transportation Commission, Berkeley, Calif., Dec.

Howard Humphreys and Partners. (1987). *Heathrow surface access study*, Report prepared for the U. K. Dept. of Transport, Leatherhead, Surrey, England, Jun.

ICF Kaiser Engineers, Inc., with Gannett Fleming, Inc., and KPMG Peat Marwick. (1995). *Miami Intermodal Center: Travel demand forecast report*, Prepared for Florida Dept. of Transp. and Federal Highway Administration, Miami, Fl., Aug.

Pels, E., Nijkamp, P., and Rietveld, P. (2003). "Access to and competition between airports: A case study for the San Francisco Bay Area." *Transp. Res.*, 37A(1), 71-83.

Portland Metro. (1998). *PDX ground access study model summary*, Prepared by the Travel Forecasting Staff, Portland, Ore., May.

Resource Systems Group, Inc. (2004). *O'Hare and Midway Airport Express Train ridership forecasting study: Chicago air traveler stated preference survey report*, Prepared for Chicago Dept. of Transp., White River Junction, Vt., Jan.

Sobieniak, J., Westin, R., Rosapep, T., and Shin T. (1979). "Choice of access mode to intercity terminals," *Transp. Res. Rec.,* TRB, 728, 47-53.

Spear, B. D. (1984). *An analysis of the demand for airport bus services at Washington National and Dulles Airports*. Report DOT-TSC-FAA-84-2, U.S. Dept. of Transp., Transportation Systems Center, Cambridge, Mass., May.

U.S. Federal Transit Administration (FTA) and San Francisco Bay Area Rapid Transit District (BART). (2002). *BART – Oakland International Airport Connector, Final Environmental Impact Report/Environmental Impact Statement*, State Clearinghouse No. 99112009, Oakland, Calif., Mar.

# Assessment of Permanent Deformation Behavior of Asphalt Concrete by Improved Triaxial Cyclic Compression Testing

B. Hofko & R. Blab
*Chair of Road and Airfield Engineering, Vienna University of Technology, Vienna, Austria*

ABSTRACT: For the characterization of the permanent deformation behavior (rutting) of asphalt concrete (AC), the triaxial cyclic compression test (TCCT) is a standardized test method. The test simulates traffic loading by applying a sinusoidal compressive stress in vertical direction and a radial confining pressure, which simulates the confinement of the specimen within the pavement structure. As a result the permanent axial strain versus load cycles is obtained. In the standard test procedure the confining pressure is held constant throughout the test to simplify the test control. However, in reality the confining pressure oscillates in a sinusoidal way with a certain phase lag to the vertical loading due to the viscoelastic characteristics of AC. The phase lag of hot mix asphalt (HMA) depends on the temperature and load frequency. The paper presents an innovative approach to improve the TCCT by implementing sinusoidal confining pressure. Strain gauges are attached directly to the specimen's surface to measure the radial deformation and to obtain the phase lag between the axial loading and the radial reaction. These experiments are carried out for various mixtures at temperatures ranging from 10°C to 50°C and frequencies from 0.1 Hz to 30 Hz. Thus, radial strain and phase lag can be analyzed as a function of temperature and loading frequency. In a second step the sinusoidal confining pressure with the obtained phase lag will be implemented into the test routine and results from standard TCCTs with constant confining pressure versus improved TCCTs with oscillating pressure can be compared and discussed.

KEY WORDS: Performance based test method, high-temperature behavior of AC, triaxial cyclic compression test.

## 1 INTRODUCTION

The challenge facing designers of flexible road pavements today is the need to fully characterize the thermo-rheological properties of HMA on the one hand while on the other hand also providing a realistic simulation of the traffic- and climate-induced stresses to which pavement structures are exposed over their design lives of 20 to 30 years. Therefore a significant part of European pavement research in the last decade has been dedicated to developing and standardizing performance based test methods for HMAs on the basis of effective mechanical characteristics (Blab and Eberhardsteiner, 2006). These methods are now implemented in European Standards and used for specifying the mix properties within an advanced type testing procedure required to meet customized quality standards for materials defined in tender documents as well as for mix design.

## 2 PERFORMANCE-BASED LAB TEST

To describe the performance of HMAs entirely three indicators have to be taken into account. It is (1) the low-temperature cracking, (2) the pavement stiffness and fatigue at intermediate temperatures and (3) the permanent deformation at high temperatures (rutting).

The European Standard EN 12697-25 is responsible for cyclic compressions tests to address the high temperature behavior of HMAs. Two methods are implemented in the standard, the uniaxial cyclic compression test with restricted radial strain and the triaxial cyclic compression test (TCCT).

### 2.1 Reaction of a Flexible Pavement to Traffic Loading

A flexible road pavement that is subjected to loading by a passing wheel exhibits a reaction as shown in Figure 1. The loading $F_{ax}$ itself leads to compressive axial stresses (and strains) within the pavement structure. The axial deformation lags the axial stresses due to the viscoelastic behavior of bituminous bound materials. As each point within the structure is more or less confined in radial direction, no or very limited radial deformation occurs. Further, radial stresses result from this confinement, which lag the axial stresses as well. For cyclic loading these lags can be described as phase angles ($\varphi_{ax,ax}$, $\varphi_{ax,rad}$) which are specific viscoelastic material parameters of AC. According to the theory of viscoelasticity the phase angle depends on the temperature and frequency of loading (Findley et al., 1989).

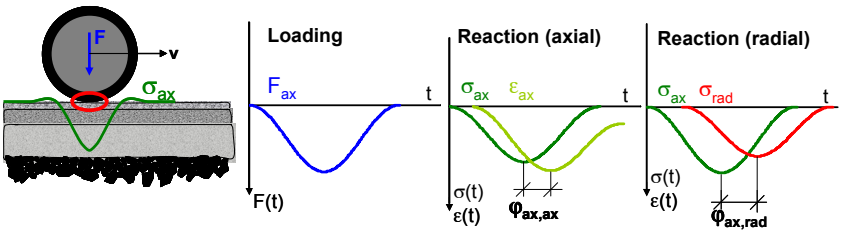



Figure 1: Loading and reaction of a flexible pavement structure due to traffic (schematic)

### 2.2 The Triaxial Cyclic Compression Test (TCCT)

To address the deformation behavior at high temperatures, the TCCT simulates the situation on the road. A cylindrical specimen (recommended minimum: h = 200 mm, d = 100 mm) is situated in a triaxial pressure cell between two load plates. The cell is filled with an incompressible medium (e.g. water) which is pressurized and creates the radial confining pressure. To protect the specimen from the medium, it is surrounded by a latex membrane. According to EN 12697-25 and as shown in Figure 2, the axial (wheel) loading is simulated by a sinusoidal compressive load $F_{ax}$:

$$F_{ax}(t) = \left[\sigma_{rad} + \sigma_a \cdot [1 + \sin(2\pi ft)]\right] \cdot A_{sp} \quad (1)$$

| | |
|---|---|
| $F_{ax}(t)$ | axial loading with respect to time |
| $\sigma_{rad}$ | radial confining pressure |
| $\sigma_a$ | amplitude of axial loading stress |
| f | frequency of loading [Hz] |
| t | time |
| $A_{sp}$ | cross-section of the specimen |

To simplify the test control a constant radial confining pressure $\sigma_{rad}$ is applied to the specimen to simulation the confinement within the pavement structure. As a reaction axial and radial strain with respective phase angles can be derived.

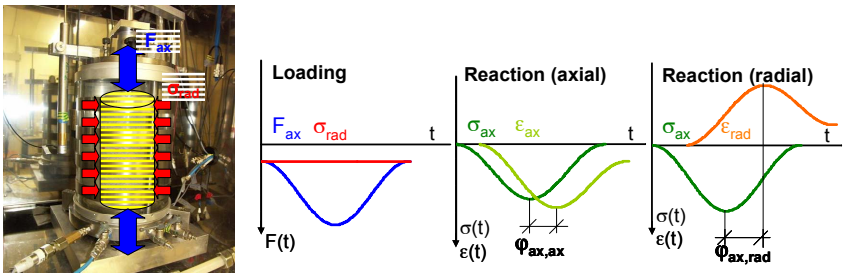



Figure 2: Loading and reaction of an HMA-specimen in the TCCT (schematic)

Depending on the type of HMA-layer (surface or base) different test parameters are given by the European Standard EN 13108-20 (see Table 1). The standard TCCT results in a relationship between the permanent axial deformation $\varepsilon_{ax,perm}$ vs. the number of load cycles N. The creep rate $f_c$ in micro strain per load cycle indicates the resistance of the HMA to permanent deformation at high temperatures. A high value of $f_c$ stands for low rutting-resistance vice verse. (EN 12697-25, 2006)

Table 1: TCCT parameters (EN 13108-20, 2006)

| | **Surface Layer** | **Base Layer** |
|---|---|---|
| **Test temperature** | 50°C | 40°C |
| **Radial confining pressure** | 150 kPa | 50 kPa |
| **Amplitude of axial loading** | 300 kPa | 200 kPa |
| **Frequency of loading** | 3,0 Hz | 3,0 Hz |

Compared to the actual stress situation in the pavement under a passing wheel with the TCCT, the constant confining pressure $\sigma_{rad}$ is a simplification which has an influence on the result itself since the 3-d state of stress in the specimen is different from that in a road pavement. By applying a cyclic confining pressure with a certain phase angle, the test procedure will simulate the state of stress in a more realistic way and thus the permanent deformation behavior in more reliable way. The challenge is therefore to (1) measure the radial deformation and (2) to determine the phase angle $\varphi_{ax,rad}$ which is not only dependent on the temperature and frequency of loading but also the binder and the mix type itself.

## 3 APPROACH

To improve the TCCT by implementing cyclic confining pressure, an extensive testing program using the simplified TCCT and uniaxial cyclic compression tests (UCCT) is carried out on different ACs. The axial loading, the radial and axial deformation are recorded and analyzed to obtain the phase angle $\varphi_{ax,rad}$. This phase angle is then used as an input value for phase lag between the axial loading and the cyclic confining pressure in the improved TCCT. Results of simplified and improved TCCT will be compared and analyzed. The approach is also depicted in Figure 3.

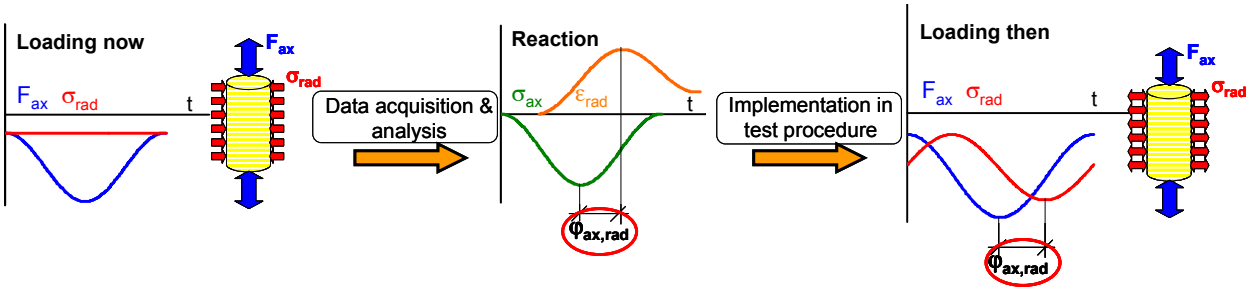



Figure 3: Approach to implement cyclic confining pressure in the TCCT

### 3.1 .Measuring Radial Deformation

To measure the radial strain of a cylindrical specimen and in consequence to assess $\varphi_{ax,rad}$, different devices based on different methods are available. For this purpose, the measuring device has to fulfill the following requirements:

- Temperature range: 0 to 60°C
- Total radial strain: > 5%
- Qualified to measure repeated loading with a frequency of up to 30 Hz
- Exact and reliable measurement of cyclic radial deformation
- Fits into the triaxial cell with its limited space

In a first step three systems were implemented and analyzed regarding their capability to fulfill above requirements (Kappl, 2007). They are depicted in Figure 4.

None of the systems was suited for the analysis of radial strain in the TCCT. The device based on extensometers fulfills the first four requirements excellently. Due to the dimensions of the device, it does not fit into the triaxial cell and could not be used. The system which measures on the basis of a rosette of steel springs with strain gauges is a compact system that fits into the triaxial cell. But the sealing which protects the strain gauges gets damaged due to the large strains that occur when attaching to the specimen. The third system based on linear variable differential transformers (LVDTs) is first of all too sluggish to record the cyclic radial deformation (see Figure 5). Still, it is a suitable device to obtain the cumulative strain for standard TCCTs, but it cannot be used for an analysis of phase lags. Furthermore, the device as the two other systems mentioned is also attached to the specimen on the outside of the latex membrane. This membrane reacts viscoelasticly itself and therefore influences measured radial the phase lag.


Figure 4: Chain-extensometer (left), steel spring device (middle), LVDT device (right) (Kappl, 2007)

### 3.2 Introducing Strain Gauges

As the mentioned devices did not lead to satisfactory results regarding radial strain analysis, another approach was needed to analyze the cyclic radial deformation successfully. As strain gauges have been employed on other inhomogeneous materials like concrete for decades, a new method was developed how to attach a strain gauge directly to an AC-specimen. As indicated in Figure 5, it records the cyclic radial deformation perfectly.

The major advantage to the three methods described in chapter 3.1 is that for the first time a measuring device is attached directly to the surface of the specimen and not outside the latex membrane. To protect the strain gauge from water, it is coated with a sealing based on silicone rubber. To guarantee reliable and repeatable measurements with strain gauges, each component of the system was optimized, including the adhesive, type of bonding, length of the strain gauge and the geometric setup.

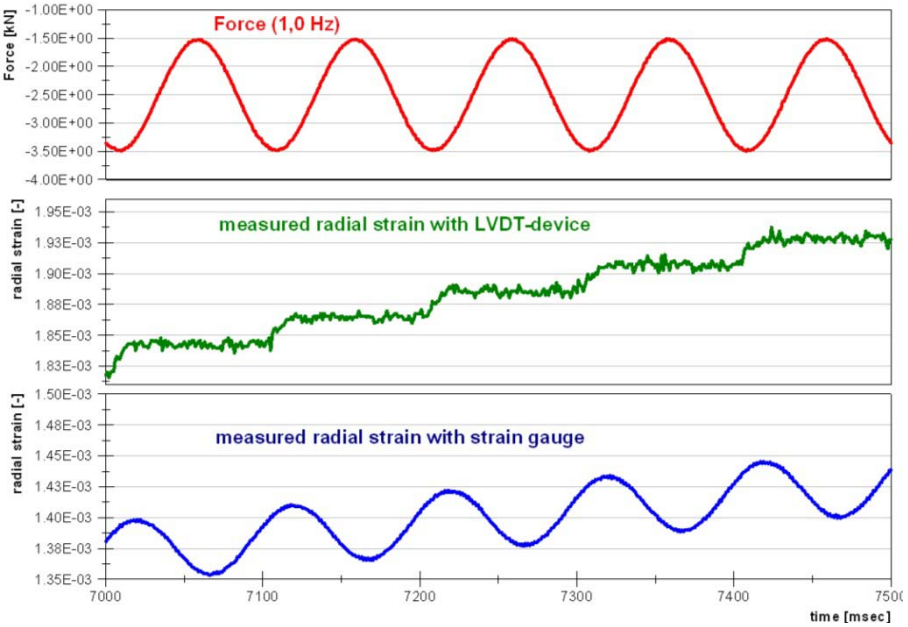



Figure 5: Measured radial deformation with LVDT-device (green) vs. strain gauge (blue)

### 3.2.1 Adhesive

The only adhesive we found suitable for a surface like AC-specimens is a two component adhesive which is also used when working with concrete. A solid acrylpolymere is mixed with a liquid monomer mixture based on methyl acryloester. The working limits regarding the temperature are -200°C and +80°C and the adhesive shows an extensibility of more than 10%. At room temperature the hardening time is about 15 minutes. The material exhibits a Young's modulus after hardening of around 13,000 N/mm² (HBM, 2009). All parameters of the adhesive are suitable for the type of tests carried out in the project. The only challenge is the high modulus of the adhesive which can be problematic at higher testing temperatures due to low stiffness of HMAs.

### 3.2.2 Type of bonding

From literature (Hoffmann, 1987) it is known, that the bonding quality of the strain gauge to the specimen is most important for correct and reliable readings. The strain gauge must be bonded force- and form-fit to the specimen. This is usually accomplished by adhering the device in a holohedral way to the surface. However, due to the high stiffness of the adhesive compared to the stiffness of HMAs, a holohedral bonding results in incorrect readings. The system of strain gauge and adhesive works like a restraint and prevents deformation.

As the strain gauges are only used for measuring strain due to tension, a different method of bonding is used. The strain gauge is attached to the specimen only punctually at both ends of the device. Figure 6 shows a specimen with an attached strain gauge. The adhesive area should be large enough, so that the force can be transmitted into the device without being too concentrated and not too large which would have a similar effect as holohedral bonding. The optimal area depends on the size of the largest aggregate used – in the present case with an AC11 (11 mm largest aggregate) a 2.0 cm by 1.5 cm area symmetrical around the end of the strain gauge showed best results.

### 3.2.3 Length of strain gauge

The minimum length of the strain gauge depends on the size of the largest aggregate fraction. Strain gauge manufacturers recommend that the length of the strain gauge should be more than 4- to 5-times the size of the largest aggregate fraction in the mix (Hoffmann, 1987). Thus it can be guaranteed that the recordings represent an average strain rather than a non-

representative strain between one specific stone and the mastix due to too short strain gauges.

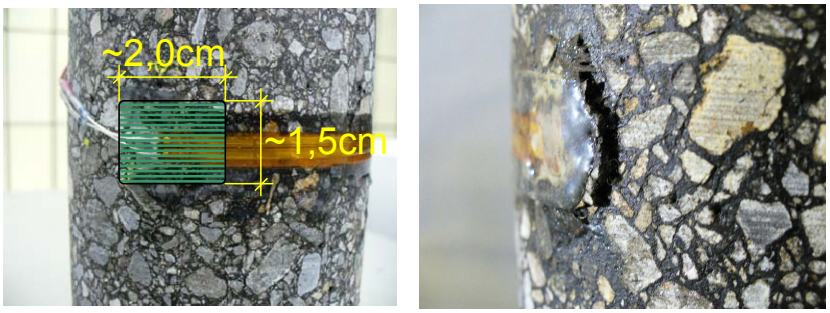



Figure 6: Adhesive area on specimen (left), typical damage due to small bonded area (right)

For the measurements strain gauges with an active, strain sensitive length of 100 mm and 150 mm respectively are used and results were compared. No significant difference was found in the results between both types, neither with regard to the recorded total radial strain nor to the calculated phase angles.

### 3.2.4 Geometric setup of strain gauges

Two setups of strain gauges were compared. In one case two strain gauges with a length of 100 mm were attached to the surface of the specimen each covering about one third of the perimeter. In the other case one 150 mm strain gauge was used covering about half the perimeter of the specimen. Again, comparative TCCTs were carried out with both setups to isolate any influence of the geometric setup. No difference was found in the results. Therefore the single strain gauge setup was used as it is less time consuming to prepare the specimen for testing. In all cases the strain gauges are attached at half height of the specimen as we expect a homogeneous state of stress in this region of the specimen.

### 3.2.5 Radial strain limits

With a thermo-rheological material like HMA, the test temperature has a significant effect on material parameters like the dynamic modulus E*. Especially at high temperatures, the binder itself exhibits more fluid-like characteristics. To prevent the adhesive area from getting separated from the rest of the specimen like shown in Figure 6, the total radial strain should be limited to 1% especially at high temperatures and the adhesive area should never be smaller than indicated in Figure 6.

## 3.3 Regression used for Analysis

To be able to analyze the phase lag analytically, the recorded data from tests must be represented by a regression functions which allows to find roots in an analytical way. For a stress-controlled test like cyclic compressions tests the advanced sinusoidal function

$$f(t) = a_1 + a_2 \cdot t + a_3 \cdot \sin(2\pi f \cdot t + a_4) + a_5 \cdot \sin(4\pi f \cdot t + a_6) \qquad (2)$$

was found to deliver best results by (Kappl, 2007). The function takes into account a linear term to describe cumulative permanent deformation, as well as two sinusoidal terms to describe the oscillating part. A detailed discussion about different regression functions can be found in (Kappl, 2007).

## 4 MATERIALS

In this first stage of the project, one mix design (AC11 pen70/100) with a variation in the binder content is studied to systematically analyze the influence of this mix parameter on the phase angle $\varphi_{ax,rad}$. In Table 2 the variation of the mix design parameters for the AC11 pen70/100 are shown.

Table 2: Variation in the mix design of AC11 pen70/100

| | | | |
|---|---|---|---|
| **binder content [M%]** | 5.0 | 5.3 | 5.6 |
| **air void content [V%]** | 3.0 | 3.0 | 3.0 |

## 5 TEST PROGRAM AND RESULTS

To identify the viscoelastic behavior of the AC under cyclic compressive loading, in a first test phase (chapter 5.1) assumptions of the viscoelastic theory regarding the independence of viscoelastic parameters to the state of stress were verified. In the second test phase (chapter 5.2) the test temperature and frequency were varied for each tested material.

### 5.1 Test Phase #1

According to the theory of viscoelasticity, viscoelastic parameters should not be influenced by the state of stress applied. To prove this assumption, TCCTs at 30°C were performed at the AC11 pen70/100 at three different deviatoric stress levels. Figure 7 provides information on the test setup and the results. "sig_m" stands for the mean axial stress, "sig_a" for the axial stress amplitude and "sig_rad" for the constant confining pressure. "sig_d" is the resulting deviatoric stress. Three specimens were tested for each deviatoric stress level. However, the highest loading setup produced too large radial strains, so only one test was performed successfully at a limited frequency range.

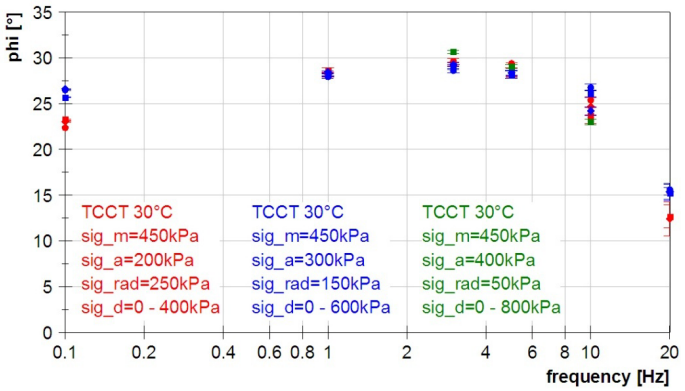



Figure 7: Phase angle between axial loading and radial deformation $\varphi_{ax,rad}$ at different deviatoric stresses for AC11 pen70/100 – frequency in logarithmic scale

#### 5.1.1 Results

In Figure 7 different colors assign different deviatoric stress levels. For each frequency, the last 20% of the oscillations were used for analysis. So it is guaranteed that the test data was taken from a steady state with constant phase angles. The markers in the diagram show the mean value of the phase angle $\varphi_{ax,\,rad}$ for each specimen at each frequency. Additionally the standard deviation is depicted as a bar. From the results it can be concluded that the state of stress has indeed no influence on the viscoelastic parameters. Thus in the test phase #2 the

10°C and 30°C tests were carried out uniaxially, since uniaxial tests are much easier and quicker to carry out. Still the 50°C tests are carried out with confining pressure according to the stress levels in EN 13108-20.

## 5.2 Test Phase #2

In the second test phase the main influences on $\varphi_{ax,rad}$ are isolated. Besides the influence of frequency and temperature, the impact of the binder content on the phase angle was investigated. Tests were carried out at 10°C, 30°C and 50°C (see Table 3) at a frequency sweep shown in Table 4. However, the 0.5 and 30.0 Hz frequency packages were only added at a later stage of the project. For each material at least two specimens per temperature were tested.

Table 3: Test set-up regarding stress-levels

| **Temperature [°C]** | 10 | 30 | 50 |
|---|---|---|---|
| **Test type** | UCCT | UCCT | TCCT |
| **sig_m [kPa]** | 600 | 250 | 450 |
| **sig_a [kPa]** | 500 | 150 | 300 |
| **sig_rad [kPa]** | 0 | 0 | 150 |

Table 4: Test set-up regarding load frequency and number of load cycles

| **Frequency [Hz]** | 0.1 | 0.5 | 1.0 | 3.0 | 5.0 | 10.0 | 20.0 | 30.0 | 0.1 | 0.5 | 1.0 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Load cycles [-]** | 25 | 50 | 200 | 600 | 1000 | 1000 | 1000 | 1000 | 25 | 25 | 100 |

### 5.2.1 Results

The results regarding the binder content are displayed in the diagrams in Figure 8. Each diagram represents a test temperature, each marker in a diagram the mean value of the phase angle $\varphi_{ax,rad}$ for one test at a certain frequency.

Again the phase angle was obtained from the last 20% of each frequency package to make sure that the test is in a steady state. The bars around each marker show the standard deviation. Each color marks one binder content (5.0, 5.3, 5.6 M%). From the results it can be concluded that the binder content has no evident influence on the phase angle. Still, as the bituminous binder is responsible for the viscous characteristics of HMAs, it is believed that the binder content does have an influence on the viscoelastic parameters. Thus we will vary the binder content in a larger range (± 0,5 M%) in the second stage of the experimental program to verify the present results.

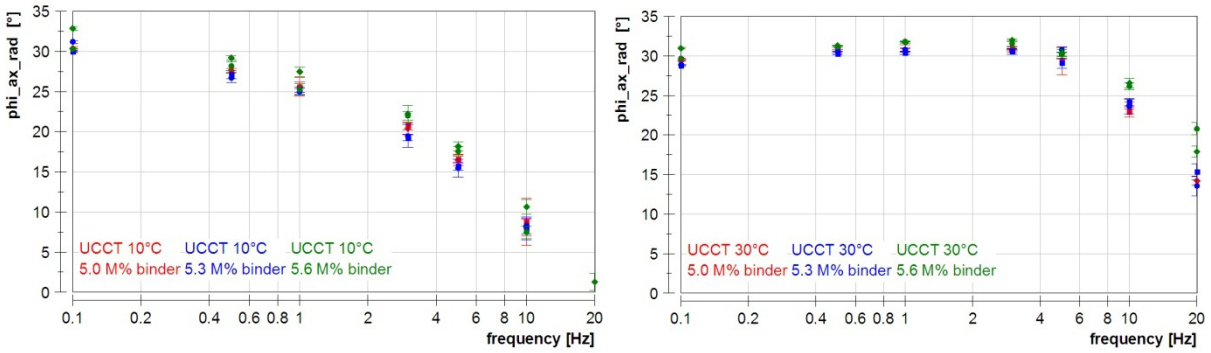



Figure 8: Impact of binder content on phase angle for 10°C (left), 30°C (right) for AC11 pen70/100 – frequency in logarithmic scale.

As expected, temperature and frequency have a crucial impact on the viscoelastic

parameters as HMAs are considered thermo-rheological materials. The left diagram in Figure 9 presents the results. The setup of the diagram is the same as for Figure 8, except that the different colors match different temperatures.

Interestingly enough, the phase angle which represents the rate of elastic vs. viscous behavior is not directly proportional to the test temperature. The highest phase angles occur at a test temperature of 30°C. This is consistent with analysis done for stiffness-tests at HMA-specimens (Blab and Eberhardsteiner, 2006). If results from these tests are presented in the black-space (norm of the dynamic modulus $|E^*|$ vs. phase angle φ), the maximum phase angles never occur at the highest temperatures.

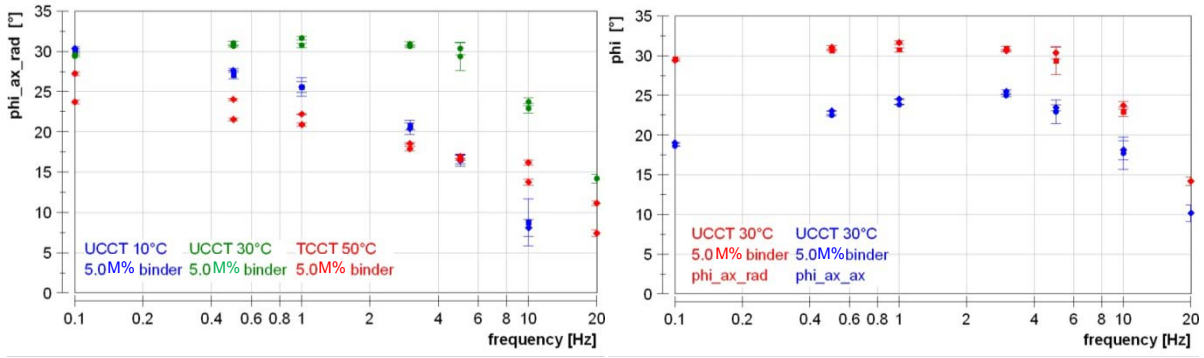



Figure 9: Impact of test temperature on the phase angle (left) and comparison of phase angle $\varphi_{ax,ax}$ and $\varphi_{ax,rad}$ (right) for AC11 pen70/100 – frequency in logarithmic scale.

By comparing the phase angle between axial loading and axial deformation $\varphi_{ax,ax}$ and the phase angle between axial loading and radial deformation $\varphi_{ax,rad}$ ,the two phase angles show a difference of 4° to 11°. $\varphi_{ax,rad}$ is always the larger one. The right diagram in Figure 9 shows an example. The present assumption is that this effect is associated with the anisotropy of the material. Right now the cylindrical specimens are drilled out of an HMA plate perpendicular to the direction of compaction. In the second stage of the project, specimens will also be drilled out in the actual direction of compaction results will be compared to validate the assumption.

Thus, for the improvement of the standard TCCT by implementing a cyclic confining pressure according to Figure 3 the phase angle $\varphi_{ax,rad}$ at 50°C and 3 Hz is needed. For the AC11 pen70/100 with 5.0 M% binder content a mean value of 18.2° with a standard deviation of 0.46° was found. These numbers were derived from tests carried out on four different specimens.

## 6 CONCLUSIONS AND NEXT STEPS

From the present stage of the project, we conclude the following:

- Strain gauge technique successfully implemented – high accuracy of measurement directly on the surface of the specimen
- Mathematical approach to derive phase angles from measured data developed
- Test setup optimized to obtain repeatable material parameters

The first test phase shows:

- No influence of deviatoric stress on viscoelastic material parameters – rheologically simple material behavior
- Radial phase lag strongly dependent on temperature and frequency; no or moderate influence of binder content

- Maximum phase lag at a temperature of 30°C – analogue to stiffness tests
- Difference in axial and radial phase lag depending on temperature and frequency between 4° and 11°.

Some issues about the test device can be optimized. The most important part is the adhesive which is very stiff and could be replaced by an adhesive with a similar stiffness like the specimen. This is necessary to prevent any influence of the adhesive on the results especially at high temperatures. There are two component adhesives which allow the user to change the stiffness by changing the mix of the components. Right now the hardening of these adhesives at room temperature takes high contact pressures and a fairly long time.

The test program will be carried on in a second stage. It will include an AC11 with SBS-modified binder at different binder contents and also different contents of air voids to analyze any influence of this mix design parameter. Functional relationship between the phase angle and mix parameters should be obtained.

The cyclic confining pressure with $\varphi_{ax,rad}$ will be implemented to improve the TCCT and results of the standard and the improved TCCT will be compared and analyzed.

In a parallel project other viscoelastic parameters, i.e. the phase angle $\varphi_{ax,ax}$, the elastic ($E_1$) and viscous part ($E_2$) of the dynamic modulus $E^*$ obtained from the TCCTs and UCCTs are also analyzed. More information on the viscoelastic behavior of HMAs at high temperatures is going to be gathered from this project. The main objective is to find functional relationships between mix parameters and the viscoelastic parameters.

## REFERENCES

Blab, R. and Eberhardsteiner, J., 2006."*Methods for the Optimization of Flexible Road Pavements*. Mitteilungen des Instituts für Straßenbau und Straßenerhaltung, Vol. 17, Vienna.

EN 12697-25, 2006. *Bituminous mixtures – Test methods for hot mix asphalt – Part 25: Cyclic compression test.* European Standard, 2006-05-01.

EN 13108-20, 2006. *Bituminous mixtures – Material specifications – Part 20: Type Testing.* European Standard, 2006-07-01.

Findley, W. N., Lai, J. S. and Onaran, K., 1989. *Creep and Relaxation of nonlinear viscoelastic Materials.* Dover Publications Inc., New York.

Kappl, K., 2007. *Assessment and modelling of permanent deformation behaviour of bituminous mixtures with triaxial cyclic compression tests.* (in German), PhD Thesis, Vienna University of Technology, Vienna.

Hoffmann, K., 1987. *Eine Einführung in die Technik des Messens mit Dehnungsmessstreifen.* Hottinger Baldwin Messtechnik GmbH, Alsbach.

HBM, 2009. *Instruction to use for the X60 2 component adhesive.* Hottinger Baldwin Messtechnik GmbH, Alsbach.

# Asphalt concrete rutting resistance assessment by the flow number test

Alex K. Apeagyei[1], Ph.D., P.E., M.ASCE
Brian K. Diefenderfer[2], Ph.D., P.E., M.ASCE

[1]*Virginia Transportation Research Council, 530 Edgemont Road, Charlottesville, VA 22903;*
*tel. 434 293-1979; fax: 434 293-1990; email:alex.apeagyei@vdot.virginia.gov*
[2]*Virginia Transportation Research Council, 530 Edgemont Road, Charlottesville, VA 22903;*
*tel. 434293-1944; fax: 434 293-1990; email:brian.diefenderfer@vdot.virginia.gov*

**ABSTRACT:** This study was conducted to evaluate the effect of recycled asphalt pavement (RAP) amounts on the rutting resistance of plant-produced SUPERPAVE asphalt concrete (AC) mixtures in the laboratory. RAP amounts ranged from 0 to 25 percent by weight of mixture. Tests conducted included the dynamic modulus test at multiple temperatures and the flow number (FN) test at 54°C to characterize the stiffness and rutting resistance of the mixtures, respectively. Significant differences in rutting resistance were observed even though the mixtures considered were designed for similar levels of traffic (between 3 million and 10 million design equivalent single-axle loads). The effect of RAP amount on rutting resistance was unexpected as FN decreased as the RAP amount increased from 10 percent to 25 percent. The practice of allowing softer binders for AC mixtures containing high RAP amounts (>20%) was suggested as a possible reason for the relatively lower rutting resistance at the 25 percent RAP level. It was concluded that the use of a high RAP amount in combination with the softer binder may have the inadvertent effect of reducing the rutting resistance of certain AC mixtures. Future work involving a larger set of AC mixtures than presented in this paper is recommended.

## INTRODUCTION

Because of the dwindling sources of good-quality aggregates, astronomical increases in the prices of bituminous materials, and need for more sustainable use of natural resources, more and more state department of transportations are specifying recycled asphalt pavement (RAP) as part of their regular asphalt concrete (AC) designs. The Virginia Department of Transportation (VDOT) recently increased the threshold of allowable RAP for SUPERPAVE mixtures from 20 to 30 percent (Maupin et al., 2008). As in Virginia, most conventional AC mixtures produced in other parts of the United States contain up to 30 percent RAP with trends toward higher RAP usage in the future. The effect of RAP on pavement performance is therefore of tremendous importance and is currently receiving much attention as evidenced by numerous recent publications on the subject (Al-Qadi et al., 2007; McDaniel and Anderson, 2001).

The effects of RAP on fatigue resistance, moisture damage, and rutting have been reported in the literature. Fatigue resistance of AC mixtures has been

reported to decrease significantly for RAP contents between 20 and 40 percent, but no such effect was reported for transitioning from 0 to the 10 percent RAP level (McDaniel and Anderson, 2001). According to the authors, using up to 40 percent RAP in well-designed AC mixtures for paving applications should, however, not have detrimental effects on performance. The authors also indicated that the fatigue life of asphalt mixtures tended to decrease with the amount of RAP unless the binder grade was adjusted to account for the amount of RAP used. McDaniel and Anderson (2001) also reported that even though the permanent shear strains (a measure of rutting resistance) from the repeated shear at constant height (RSCH) test for the mixtures decreased (indicating higher rutting resistance) with increasing RAP content, the variability in the data was high. Hajj et al. (2007) reported that AC mixtures containing 15 percent RAP showed higher resistance to moisture damage than those containing 30 percent RAP in some mixtures. In the same study, the authors showed that rutting resistance of certain polymer-modified AC mixtures containing 15 to 30 percent RAP was generally similar or lower than that of control mixtures with little or no RAP. Fatigue life was also significantly affected by RAP in these polymer-modified mixtures. Maupin et al. (2008) compared AC mixtures containing a high amount of RAP (21 to 30 percent) with mixtures containing less than 20 percent RAP and found no significance difference between the mixtures in terms of three important performance indicators: fatigue endurance limit, asphalt pavement analyzer (APA) rutting resistance, and moisture susceptibility. The authors, however, noted a higher moisture susceptibility failure rate (38 percent) in the higher-RAP mixtures compared to 7 percent in mixtures with less than 20 percent RAP (7 percent). Similar effects of RAP on moisture susceptibility using the tensile strength ratio (TSR) were reported by Li et al. (2004) for AC mixtures containing up to 40 percent RAP.

The foregoing discussion suggests the effects of RAP on AC performance are generally positive, but the need still exists to evaluate the rutting performance of high-RAP mixtures, especially with methods that can quantify the performance properties of the mixture in a more mechanistic manner. The effects of RAP on the rutting resistance of AC appear to be receiving the least attention compared to the other distresses. This is because it is generally assumed that rutting will be less in mixtures containing stiffer binders such as those found in RAP. The effect of the amount of RAP on permanent deformation or rutting of AC pavement is, therefore, the focus of this study.

Tests discussed in this paper included the dynamic modulus test at multiple temperatures and the flow number (FN) test at 54°C to characterize the stiffness and rutting susceptibility, respectively, of AC mixtures. Dynamic modulus is a material property that can accurately describe the viscoelastic behavior of AC mixtures. The dynamic modulus test has been shown in both field and laboratory studies to be highly correlated with rutting. The dynamic modulus is also a key material input parameter for the new AASHTO mechanistic-empirical pavement design guide (MEPDG). The FN test (AASHTO, 2007) was selected for this study for several important reasons. FN shows good correlation with field rutting of AC mixtures for different traffic levels (Witczak, 2007; Mohammad et al., 2006; Witczak et al. 2002). The FN test was recommended by the NCHRP 9-19 panel as one of the simple performance tests (Witczak et al., 2002) to evaluate rutting in AC. The testing temperature and testing load used during the FN test can be selected to mimic as close as possible the actual traffic

and climatic conditions at the location where the AC mixture under consideration will be placed. FN has also been used to rank asphalt mixtures in terms of rutting susceptibility in both laboratory tests and field performance (Mohammad et al., 2006). Both dynamic modulus and FN have been correlated with rutting; therefore, these tests were chosen to characterize the effect of RAP on the rutting resistance of five Virginia AC mixtures.

## PURPOSE AND SCOPE

The primary objective of this study was to compare the rutting resistance of AC mixtures designed for the same traffic levels but containing different amounts of RAP in the laboratory. All mixtures considered in the study were designed using SUPERPAVE procedures to carry the same level of highway traffic of between 3 million and 10 million ESALS over a 20-year design life.

The scope of the study was limited to five AC mixtures with nominal maximum aggregate sizes (NMAS) of 12.5 mm and 25.0 mm produced in accordance with VDOT's specifications for a Type D mixture (VDOT, 2007) using plant-mixed but laboratory-compacted specimens. The effect of RAP, gradation, and mixture stiffness are discussed**.** Tests conducted included the dynamic modulus test at multiple temperatures and the FN test at 54°C to characterize the stiffness and rutting susceptibility, respectively, of AC mixtures.

## MATERIALS AND METHODS

The mixtures were selected to be as representative as possible of the commonly used AC mixtures in Virginia. All mixtures are classified as Type D mixtures (that is, having an effective PG 70-22 binder) and are typically designed for pavements carrying medium traffic ranging from 3 million to 10 million ESALS over a 20-year design period (VDOT, 2007). VDOT provides specifications for selecting the correct binder grade for Type D mixtures depending on the amount of RAP used. For RAP amounts less than 20 percent, a PG 70-22 binder is specified; a PG 64-22 binder is specified for mixtures containing more than 20 percent RAP (VDOT, 2007). The mixtures used in this study contained RAP amounts ranging from 0 to 25 percent by weight.

All mixtures were designed using SUPERPAVE protocols. The asphalt content to achieve a design air void level of 4.0 percent at 65 gyrations was used as the design criterion. Table 1 provides design criteria details for the two mixture types considered in this study: SM-12.5D (a surface mixture with an NMAS of 12.5 mm), and BM-25.0D (a base mixture with an NMAS of 25.0 mm). The limiting values for volumetric properties for the mixtures and limiting values for the design number of gyrations are also provided in Table 1. All mixtures used in this study were obtained from contractor plants.

### Materials Sampling, Processing, and Compaction

As stated previously, samples of commonly used mixture types were collected from around Virginia. All samples were plant-mixed and sampled loose at the project site and sent to the laboratory for further processing and testing. RAP amounts ranged from 0 to 25 percent as shown in Table 2.

**Table 1.** VDOT mix design criteria for Type D mixtures

| Mix type[1] | AV[2] | VFA | VMA | F/A[3] | Number of gyrations | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | $N_d$ | $N_i$ | $N_m$ | Density (%) at $N_i$ |
| SM-12.5D | 2 - 4 | 65-83 | ≥14.0 | 0.6-1.2 | 65 | 7 | 100 | ≤89.0 |
| BM-25.0D | 2 - 4 | 64-81 | ≥12.0 | 0.6-1.3 | 65 | 7 | 100 | ≤89.1 |

[1]SM = Surface mixture; BM = Base mixture; [2]AV = Air voids; asphalt content should be selected at 4 percent air voids; VFA = Voids filled with asphalt; VMA = Voids in mineral aggregate; [3]F/A = Fines-asphalt ratio, fines-asphalt ratio is based on effective asphalt content.

A SUPERPAVE gyratory compactor was used to fabricate specimens used for the dynamic modulus and the FN tests. The mixtures for the various tests were compacted to a target air void level of 7 ± 0.5%. This was accomplished by determining the compaction height needed to obtain the required target air void content. Following compaction, a coring rig was used to create the required 150-mm-tall by 100-mm-diameter specimen from the gyratory compacted specimen for testing. At least three replicate 150 mm by 100 mm specimens were fabricated for each of the five mixtures. Table 2 shows that the average air void content of the laboratory-compacted specimens ranged from 6.95 to 7.85 percent.

The following test methods were used to determine the physical and mechanical properties of the asphalt binders in the mixtures: (i) ignition oven test (AASHTO T 308); (ii) binder extraction test (AASHTO T 186); (iii) binder recovery test (AAHSTO T 170); and binder performance grading test (AASHTO M 320).

The actual binder content of each sampled mixture was determined from the original loose mixture in the laboratory using an ignition oven test (AASHTO T-308). The asphalt content of the mixtures obtained from the ignition oven test for each mixture is shown in Table 2 and ranged from 4.43 percent to 7.05 percent. Mix A with 0 percent RAP had the highest asphalt content of 7.05; Mix D with 20 percent RAP had the lowest asphalt content of 4.43.

The recovered binder from each mixture was subjected to temperature sweep tests to characterize the mechanical properties of the binders. Because of lack of materials only three mixtures were tested. The results of the mechanical testing including SPERPAVE rutting parameter ($G^*/\sin\delta$) are summarized in Table 2.

**Table 2.** Data for laboratory-compacted asphalt mixtures

| Mix ID | RAP (%) | NMAS (mm) | $P_b$ (%) | AV (%) | $G^*/\sin\delta$ at 70°C (kPa) | Binder grade |
|---|---|---|---|---|---|---|
| Mix A | 0 | 12.5 | 7.05 | 7.50 | - | - |
| Mix B | 10 | 12.5 | 5.68 | 6.95 | 8.498 | PG 76-22 |
| Mix C | 15 | 25.0 | 4.90 | 7.85 | 4.113 | PG 70-22 |
| Mix D | 20 | 12.5 | 4.43 | 7.16 | - | - |
| Mix E | 25 | 12.5 | 5.60 | 7.03 | 2.324 | PG 70-22 |

$P_b$ = asphalt content; - = no data available.

The recovered aggregate obtained from the ignition oven test was used to conduct gradation analysis. Fig. 1 depicts the results of the gradation tests. Some differences in gradation are apparent from Fig. 1. This was expected since the mixtures had different nominal maximum aggregate sizes. Mix A containing 0 percent RAP appears to be the coarsest mixture with only about 23 percent of material passing the 4.75-mm sieve. Conversely, Mix E containing 25 percent RAP appears to be the finest with about 50 percent of material passing the 4.75-mm sieve. Similarities can be observed between Mix B and Mix E and between Mix C and Mix D. All the mixtures passed the required VDOT gradation specifications.

**Dynamic Modulus Tests**

Dynamic modulus tests were performed with an IPC Universal Testing Machine using AASHTO TP62 protocols in the uniaxial compressive mode. Tests were conducted at five temperatures ranging from -10°C to 54°C and six testing frequencies ranging from 0.1 Hz to 25 Hz. To insure against damage to the test samples, the tests were conducted starting from the coldest temperatures to the warmest temperatures. In addition, at each testing temperature, the tests were performed starting from the highest to the lowest frequency. Load levels were selected in such a way that at each temperature-frequency combination, applied strain was in the range of 75 to 125 micro-strain. This was done to ensure that testing was conducted in the linear viscoelastic range of AC.

**FN Tests**

FN tests were performed in accordance with AASHTO TP62 on specimens using the same equipment setup and specimen geometry as used for the dynamic modulus test. The FN test is designed to evaluate rutting resistance of asphalt mixtures as recommended in NCHRP Project 9-19. The test involves subjecting an AC specimen to a repeated haversine axial compressive load pulse of 0.1 sec every 1.0 sec. The test was conducted at a temperature of 54°C. This test temperature represents the 50 percent reliability maximum high pavement temperature as determined using LTPPBind software (Pavement Systems, LLC, 2005) for locations in the southeastern portion of the United States where Virginia is located. The FN tests were all performed in the unconfined mode using a deviatioric stress of 206 kPa. The tests were continued to 10,000 cycles or until a permanent strain in the AC specimen reached 5 percent, whichever came first.

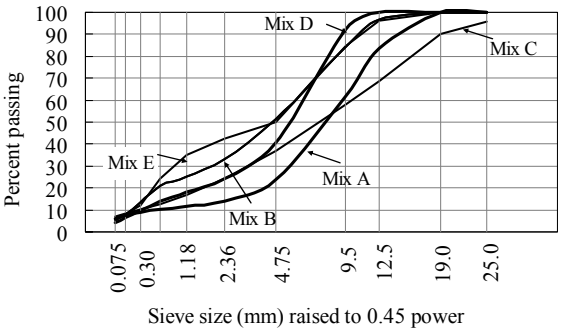



**Fig. 1.** Gradation

## RESULTS AND DISCUSSION

In this section, the effects of dynamic modulus and RAP on rutting are presented and discussed. First, dynamic modulus data at five temperatures and six testing frequencies for each of the five mixtures are used to construct dynamic modulus master curves. A master curve of dynamic modulus is developed for each mixture and plotted to illustrate the temperature and frequency dependency of the asphalt mixtures. Second, an illustration of how FN is determined numerically using standard AASHTO TP62 procedures and the results are summarized with a plot of FN versus RAP amount. Third, a brief discussion of the results is provided.

### Dynamic Modulus Tests

Dynamic modulus tests were conducted at five temperatures (-10°C, 4°C, 20°C, 38°C, and 54°C) and six loading frequencies (25 Hz, 10 Hz, 5 Hz, 1 Hz, 0.5 Hz, and 0.1 Hz). The data were used to construct dynamic modulus master curves using Equation 1. The equation is sigmoidal between dynamic modulus, E*, and the logarithm of reduced frequency at a chosen reference temperature. Non-linear optimization techniques were used to fit the experimental data to Equation 1 to obtain the model parameters and temperature shift factors. By knowing the shift factors and model parameters, dynamic modulus values at any temperature and loading frequency could be obtained. Thus the dynamic modulus master curve offers complete viscoelastic characterization of an AC mixture. A reference temperature of 20°C was used for the master curve construction.

$$\log(E^*) = \delta + \frac{\alpha}{1 + e^{(\beta + \gamma(\log f_r))}} \qquad (1)$$

where

| | | |
|---|---|---|
| $f_r$ | = | loading frequency at the reference temperature |
| $\delta$ | = | minimum value of dynamic modulus |
| $\delta+\alpha$ | = | maximum value of dynamic modulus |
| $\beta, \gamma$ | = | parameter describing the shape of the sigmoidal function |

Fig. 2 (left) shows the semi-log plot of dynamic modulus master curves for the five AC mixtures considered in this study. Several interesting observations can be made based on a close inspection of Fig. 2. For all mixtures, the amount of RAP appears to have a significant effect on dynamic modulus in both expected and unexpected ways. Mix C and Mix D containing 15 percent and 20 percent RAP, respectively, both exhibited the highest stiffness compared with the other three mixtures. As expected, Mix A containing 0 percent RAP was the least stiff at low temperatures (high reduced frequency); however, the low temperature stiffness of Mix E that contained 25 percent RAP was unexpected. It was expected that Mix E containing the highest RAP amount would be among the stiffest, but as can be seen from Fig. 2, Mix E was slightly stiffer than Mix A and Mix B containing 0 percent and 10 percent RAP respectively. It appears from the master curves that Mixes C and D were relatively stiffer than Mixes A, B, and E at the lowest test temperature.

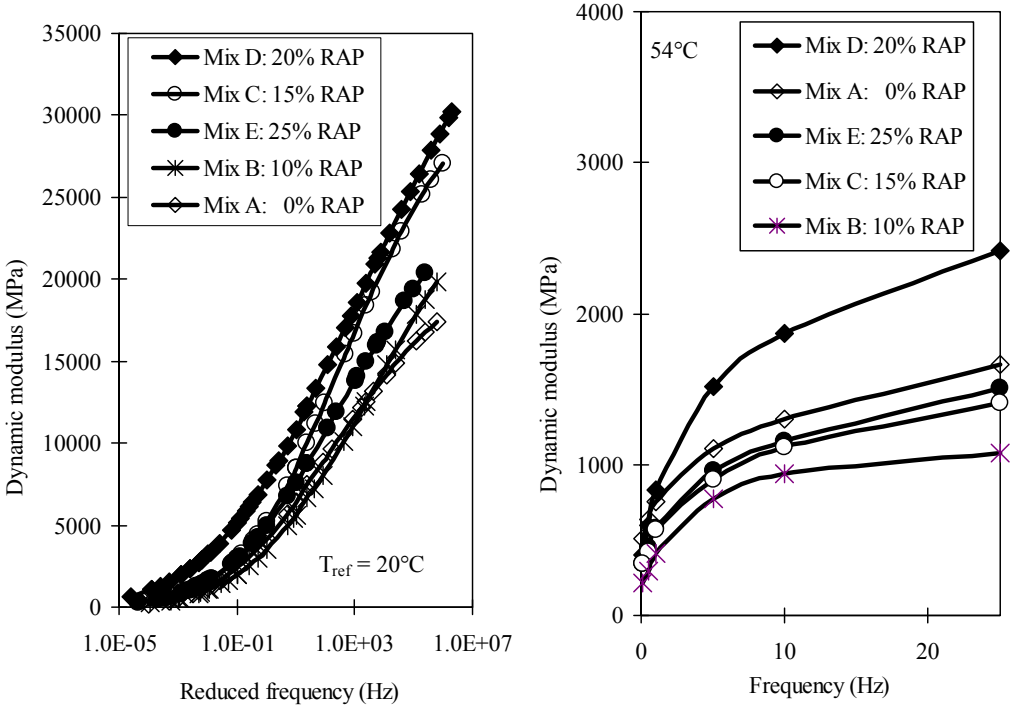



**Fig. 2.** Dynamic modulus. Left, master curve. Right, at 54°C

Dynamic modulus data obtained at 54°C are also shown in Fig. 2 (right) to illustrate the high temperature behavior of the mixtures. These high-temperature dynamic modulus data are important in understanding the effects of asphalt binders in an AC mixture as the effect of aggregate in the mixture tends to be minimal at these temperatures. In Fig. 2 (right), Mix D with 20 percent RAP was again the stiffest at the highest test temperature of 54°C followed by Mix A with 0 percent RAP. Mix E with 25 percent RAP was midway between all the mixtures as well as being softer than Mix A. The fact that Mix A without RAP exhibited stiffness that was higher than that of Mix E with 25 percent RAP was unexpected and would suggest a lower PG grade binder might have been used in Mix E as allowed under current VDOT specifications for high-RAP mixtures. The foregoing discussion of the effect of RAP on dynamic modulus suggests factors such as gradation and the base binder used may account for the behavior seen regarding the effect of RAP on dynamic modulus for the Type D mixtures used in the study.

## FN Test Results

Results of FN tests performed at 54°C are presented in this section. During the test, axial strain and cumulative number of load cycles were monitored continuously by the FN test software. From the results, plots of rate of change of strain with number of cycles could be constructed. Fig. 3 is a typical plot of accumulated strain versus number of loading cycles obtained for the mixtures showing the approximate location of FN. In practice, FN is obtained numerically as discussed next.

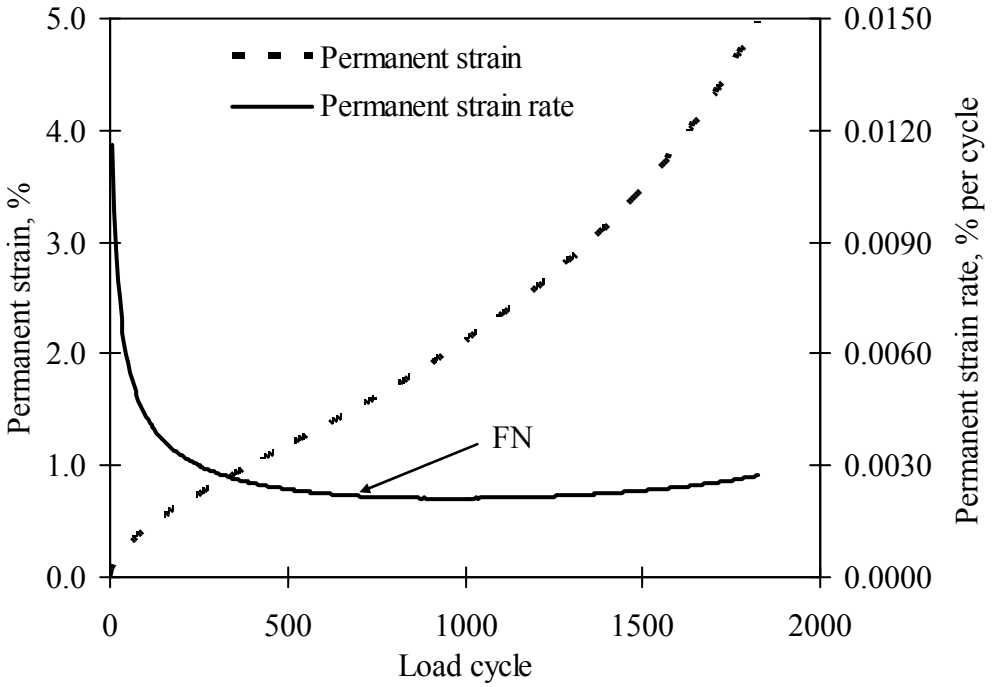



**Fig. 3.** Flow number test results for a typical asphalt mixture

FN is determined mathematically as the cycle number at which the strain rate is at a minimum. For this study, the Francken model (Biligiri et al., 2007) as implemented in the IPC software and shown in Equations 2 and 3 was used to estimate the FN. FN is defined as the point where Equation 3 changes from negative to positive.

$$\varepsilon_p(N)) = AN^B + C\left(e^{DN} - 1\right) \quad (2)$$

$$\frac{\partial^2 \varepsilon_p}{\partial N^2} = A * B * (B-1) * N^{B-2} + C * D^2 * e^{DN} \quad (3)$$

where: $\varepsilon_p$ = permanent strain from the FN test

N = number of loading cycles

A, B, C, and D = regression constants

$\frac{\partial^2 \varepsilon_p}{\partial N^2}$ = second derivative of permanent strain versus N.

Three stages are typically identified from the plot of permanent strain versus loading cycle. These stages have been classified as primary, secondary and tertiary (Witczak and Sullivan, 2002). The primary flow is characterized by a decrease in the strain rate with time. The secondary flow phase is typified by a relatively constant strain rate. The tertiary flow follows the secondary phase and begins when the strain rate begins to increase. At the tertiary stage, the specimen undergoes significant deformation with individual aggregates within the aggregate skeleton moving past each other (Kanitpong and Bahia, 2005). FN is the cycle number at the initiation of tertiary flow (Witczak, 2007).

Table 3 shows results of the FN tests obtained for each mix. The results represent the average FN for three replicate specimens for each mixture type. Also shown is the variability in test results represented by the coefficient of variation (COV), which ranged from 1.6 to 27.0 percent. This level of variability is within those previously reported by other investigators (Mohammad et al., 2006). For the mixtures considered, FN ranged from about 1100 cycles to about 7000 cycles. In general, the mixtures containing a moderate amount of RAP (10 percent and 15 percent) had higher FNs than mixtures with both high and low RAP amounts.

The effect of RAP on FN was evaluated using statistical analysis. The results of the analysis revealed several interesting effects of RAP on FN. First, no statistically significant ($p$-value = 0.19) difference in FN existed between mixtures containing 0 percent RAP (0% RAP) and mixtures containing 25 percent RAP (25% RAP). Second, the FN for mixtures containing 10 percent RAP (10% RAP) were significantly ($p$-value < 0.001) higher than for 0% RAP. Third, the FN for mixtures containing 15 percent RAP (15% RAP) and 10% RAP were statistically similar. Finally, the FN for mixtures containing 20 percent RAP (20% RAP) were significantly lower than for both 10% RAP and 15% RAP but significantly higher than for both 0% RAP and 25% RAP.

The foregoing discussion on the statistical analysis results suggests the effect of RAP on FN appears to be bilinear in nature for the mixtures considered. Fig. 4(a) illustrates such a bilinear relationship. As could be seen from Fig. 4(a), FN increased with RAP amount until about 10 percent to 15 percent RAP then decreased as RAP amount increased to 20 percent and beyond. This relationship was unexpected as it is commonly assumed that the more RAP that is added to a mixture, the higher the rutting resistance of the mixture.

One plausible reason for the unexpected effect of decreasing rutting resistance with increasing RAP amount may be an artifact of the practice of using a relatively softer binder to compensate for the stiffening effect of RAP when high RAP amounts (>20%) are used. This assertion appears to be partly supported by dynamic modulus data at 54°C, which showed that Mix A containing no RAP had a higher dynamic modulus than Mix E with 25 percent RAP.

Another possible reason for the decreasing trend of rutting resistance with RAP observed in this study could be attributed to binder stiffness. Limited binder testing results (Table 2) involving the rutting parameter G*/sinδ also appear to support the FN results. From Fig. 4(b), a clear trend of decreasing rutting G*/sinδ with increasing RAP can be seen.

**Table 3.** Flow number test results

| | | FN (cycles) | | | | |
|---|---|---|---|---|---|---|
| Mix ID | RAP (%) | Replicate 1 | Replicate 2 | Replicate 3 | Mean | COV (%) |
| Mix A | 0 | 1651 | 1180 | 2061 | 1631 | 27.0 |
| Mix B | 10 | 6781 | 6969 | 6980 | 6910 | 1.6 |
| Mix C | 15 | 6756 | 6300 | 6305 | 6454 | 4.1 |
| Mix D | 20 | 4397 | 4838 | 3919 | 4385 | 10.5 |
| Mix E | 25 | 1177 | 1076 | 1145 | 1133 | 4.6 |

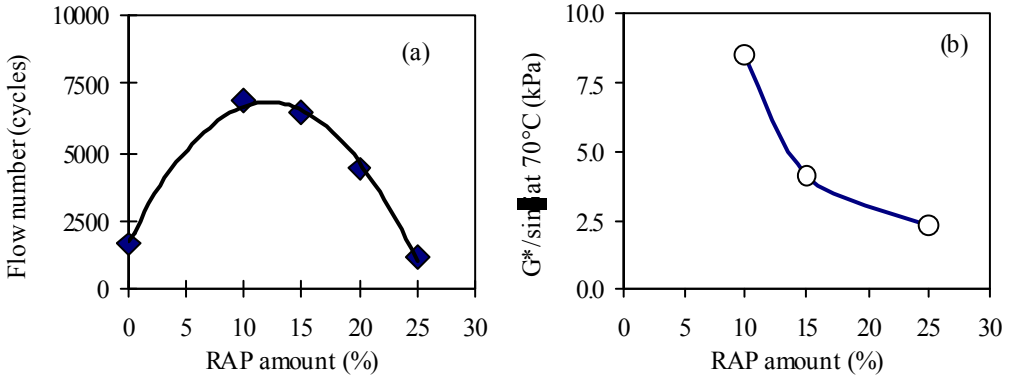



**Fig. 4.** Effect of RAP amount on rutting resistance of Type "D" mixtures. (a) FN at 54°C showing decreasing rutting resistance with RAP . (b) Decreasing rutting parameter (G*/sinδ) with RAP amount

Current VDOT specifications (VDOT, 2007), similar to many US departments of transportation, allow a PG 64-22 instead of a PG 70-22 binder to be used with higher percentages of RAP to achieve a Type D mixture. However, the results shown here suggest that this practice may have the unintended consequence of lowering the rutting resistance of these higher RAP mixtures.

As previously stated, the unexpected effect of decreasing rutting resistance with increasing RAP amount may be an artifact of the practice of using a relatively softer binder to compensate for the stiffening effect of RAP when high RAP amounts (defined as greater than 20% RAP in this study) are used. This practice has been shown by previous studies (Daniel and Lachance, 2005; Shah et al. 2007) to not always lead to mixture stiffening. Considering the performance of RAP mixtures reported in past studies, the results of the current study suggest that the current practice of allowing lower binder grades for high-RAP mixtures may need to be reviewed. With an increasing amount of RAP usage, the issue of using softer binders for high-RAP AC mixtures may become more important and will require further research. Because of the limited number of mixtures considered in this study, further studies involving additional mixtures and mixture types are warranted.

## SUMMARY AND CONCLUSIONS

This study was conducted to investigate the effect of RAP amount (0 to 25 percent) on the rutting resistance of selected VDOT's Type "D" mixtures. The following conclusions can be drawn based on the interpretation of the test results.

- The rutting resistance of mixtures containing 25 percent RAP was similar to that of mixtures containing no RAP. Both mixtures had significantly lower FN values than mixtures containing 10 percent, 15 percent, or 20 percent RAP.
- As the amount of RAP increased from 10 percent to 25 percent, FN decreased from about 7000 cycles to 1100 cycles.
- The effect of RAP amount on rutting resistance observed in this study was unexpected and contrary to the expected result that rutting resistance of AC mixtures increases with increased RAP content.

- The practice of allowing softer binders for AC mixtures containing high RAP amounts (>20% in this study) was suggested as a possible reason for the relatively lower rutting resistance at the 25 percent RAP level. This assertion was supported by limited mixture and binder stiffness data obtained in this study.
- Future work involving a larger set of AC mixtures than presented in this paper is recommended.

## ACKNOWLEDGMENTS

The authors acknowledge the assistance of the following Virginia Transportation Research Council personnel in this study: Donnie Dodds, Troy Deeds, and Aaron Dupont, for material testing, and Linda Evans, for the editorial process. Financial support for the study was provided in part with SPR funds.

## REFERENCES

Al-Qadi, I. L., Elseifi, M., and Carpenter, S. H. (2007). "Reclaimed asphalt pavement: A literature review." Research Report FHWA-ICT-07-001, Illinois Center for Transportation, Urbana.

American Association of State Highway and Transportation Officials (AASHTO). (2007). "AASHTO TP 62 standard method of test for determining dynamic modulus of hot-mix asphalt (HMA)." Washington, DC.

Biligiri, K. P., Kaloush, K. E., Mamlouk, M. S., and Witczak, M. W. (2007). "Rational modeling of tertiary flow for asphalt mixtures." In *Transportation Research Record:Journal of the Transportation Research Board,* No. 2001. Transportation Research Board of the National Academies, Washington, DC, 63-72.

Daniel, J. S., and A. Lachance. Mechanistic and Volumetric Properties of Asphalt Mixtures with Recycled Asphalt Pavement. In Transportation Research Record: Journal of the Transportation Research Board, No. 1929, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 28–36.

Hajj, E. Y., Sebaaly, P. E., and Shrestha, R. (2007). "A laboratory evaluation on the use of recycled asphalt pavements in HMA mixtures," Final Report, Regional Transportation Commission, Reno, NV.

Li, X., Clyne, T.R., and Marasteanu, M.O. (2004). "Recycled asphalt pavement (RAP) effects on binder and mixture quality." Final Report, Report No. MN/RC 2005-02, Minnesota Department of Transportation, St. Paul.

Kanitpong, K., and Bahia, H. (2005). "Relating adhesion and cohesion of asphalts to the effect of moisture on laboratory performance of asphalt mixtures." In *Transportation Research Record: Journal of the Transportation Research Board,* No. 1901. Transportation Research Board of the National Academies, Washington, DC, 33-43.

Maupin, G. W., Diefenderfer, S.D., and Gillespie, J. S. (2008). "Evaluation of Using Higher Percentages of Recycled Asphalt Pavement in Asphalt Mixes in Virginia." Report No. VTRC 08-R22, Virginia Transportation Research Council, Charlottesville.

McDaniel, R., and Anderson, R.M. (2001). "Recommended Use of Reclaimed Asphalt Pavement in SUPERPAVE Mix Design Method: Technicians

Manual." NCHRP Report 452, Transportation Research Board of the National Academies, Washington, DC.

Mohammad, L. N., Wu, Z., Obularreddy, S., Cooper, S., and Abadie, C. (2006). "Permanent deformation analysis of hot-mix asphalt mixtures with simple performance tests and 2002 mechanistic-empirical pavement design software." In *Transportation Research Record, Journal of the Transportation Research Borrad,* No. 1970. Transportation Research Board of the National Academies, Washington, DC, 133-142.

Pavement Systems, LLC. (2005). "LTPPBIND, v. 3.1 beta." Bethesda, MD.

Shah, A., R. S. McDaniel, G. A. Huber, and V. L. Gallivan. Investigation of Properties of Plant-Produced Reclaimed Asphalt Pavement Mixtures. In Transportation Research Record: Journal of the Transportation Research Board, No. 1998, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 103–111.

Virginia Department of Transportation. *Road and Bridge Specifications.* Richmond, 2007.

Witczak, M. W. (2007). "Specification Criteria for Simple Performance Tests for Rutting." NCHRP Report 580, Transportation Research Board of the National Academies, Washington, DC.

Witczack, M. W., Kaloush, K., Pellinen, T., El-Basyouny, M., and Quintus, H.V. (2002). "Simple Performance Test for SUPERPAVE Mix Design." NCHRP Report 465, Transportation Research Board of the National Academies, Washington, DC.

# A STUDY OF EFFECTIVENESS OF THIN SURFACE TREATMENTS USING HAMBURG WHEEL TRACKING DEVICE

**Md. Shaidur Rahman, EIT, S.M. ASCE [1], Mustaque Hossain, Ph.D., P.E., F. ASCE [2] and Paul Nelson, Ph.D.[3]**

## ABSTRACT

In recent years, more and more highway agencies are adopting preventive maintenance strategies or thin surface treatments to bring pavements back to appropriate serviceability for road users. This paper discusses the effectiveness of several thin surface or preventive maintenance treatments on sixteen highway test sections in Kansas. The treatments studied include thin Hot-Mix Asphalt (HMA) overlay, ultra-thin bonded asphalt surface (Nova Chip), and chip seal. These treatments were applied with three different types of surface preparation- bare surface, 1" surface recycle (hot in-place recycling), and 2" surface recycle.

Hamburg Wheel-Tracking Device (HWTD) test was conducted on the cores from the test sections with thin surface treatments under this study. The laboratory test results show that most projects exceeded the maximum rut depth limit (20 mm) specified in the study and the number of wheel passes to failure varied significantly among the projects. Cores from only three projects, two treated with Nova Chip and one with 1" HMA overlay, carried 20,000 wheel passes without exceeding maximum rut depth limit. Pair-wise comparisons or contrasts among the treatments were performed with the statistical analysis software, SAS. Air void of the HWTD test cores was found to be a significant factor affecting the performance of thin surface treatments. Since the Hamburg test samples (62 mm tall) consisted mostly of underlying layer materials, it was concluded that quality of this layer is the major determinant of performance of pavements with thin surface treatments.

## INTRODUCTION

The deterioration process of pavement starts soon after its construction due to several factors, such as, traffic loading, material degradations, and climatic effects. In general, the deterioration rate of pavements increases with use and age. Maintenance is carried out to reduce this rate (Huang, 2004). The National Cooperative Highway Research Program (NCHRP) defines preventive maintenance (PM) or thin surface treatment as "a program strategy intended to arrest light deterioration, retard progressive failures, and reduce the need for routine maintenance and service activities" (Gransberg & James, 2005). Thin surface treatments are applied to pavement to bring it back to appropriate serviceability for road users. Performance of thin surface treatment

[1]Graduate Student, Department of Civil, Construction, & Environmental Engineering, Iowa State University, 394 Town Engineering, Ames, IA 50011, PH (785) 313-4721, FAX (515) 294-8216; e-mail: msrahman@iastate.edu
[2]Professor & Associate Director, Mid-America Transportation Center, Department of Civil Engineering, Kansas State University, 2118 Fiedler Hall, Manhattan, KS 66506; PH (785) 532-1576; FAX (785) 532-7717; e-mail: mustak@ksu.edu
[3]Professor, Department of Statistics, Kansas State University, 104 Dickens Hall, Manhattan, KS 66506; PH (785) 532-0520; FAX (785) 532-7736; e-mail: nels@ksu.edu

is significantly affected by proper choice of treatment, surface preparation appropriate for existing surface conditions, and time of application. If properly designed and constructed, it can provide several benefits to the roadway surface: sealing small cracks, waterproofing the surface, improving skid resistance, offering better ride quality, and increasing pavement life. Moreover, significant savings in future rehabilitation are possible if an appropriate thin surface treatment is applied on the right pavement before rapid deterioration starts. However, thin surface treatments are not intended to increase structural capacity of pavements (Shuler, 2006; Eltahan et al., 1999).

A variety of thin surface treatments are applied to bring pavement back to appropriate serviceability for road users. The Kansas Department of Transportation (KDOT) has adopted several PM or thin surface treatments including thin overlay, ultra-thin bonded asphalt surface (Nova Chip), chip seal, micro-surfacing, and modified slurry seal. The existing pavement must be structurally sound to obtain a long performance life. Usually, the better the pavement condition when these treatments are applied, the longer the treatments will last, and the more cost effective these treatments are.

## OBJECTIVE

The main objective of this study was to evaluate the performance of three types (1" HMA, Nova Chip and chip seal) of thin surface treatments in Hamburg Wheel-Tracking Device (HWTD) test. These treatments were applied in 2007 on 16 highway sections distributed among six districts in KDOT and with three different types of surface preparation- bare surface, 1" surface recycle (1" SR or hot-in-place recycle) and 2" surface recycle (2" SR). Statistical analysis was also planned to be done to determine the significant factors that affect the performance of thin surface treatments.

## EXPERIMENTAL DESIGN AND TEST SECTIONS

Table 1 shows the experimental design in this study. For each surface preparation/treatment combination, cores from two highway test sections were collected for HWTD tests. The 1" SR/chip seal combination is not used in Kansas and thus cores were not available. The cores were collected in 2007 from 16 highway test sections in 16 counties one to two weeks after application of thin surface treatments. Cores had different depths and were sawn to desired dimensions with a radial saw. Figure 1 shows test locations and Figure 2 shows the layer distributions in the top 2.5 inches (62 mm) of pavement cross-sections since HWTD specimens were 2.5-inch (62 mm) tall.

## LABORATORY TESTS

### Gmb and Gmm Tests

The bulk specific gravity (Gmb) and maximum specific gravity (Gmm) of the specimens were determined by performing KT-15 Procedure III and KT-39 test methods, respectively. Then the percentage air voids (Va) was computed from the following equation:

$$Va(\%) = \frac{Gmm - Gmb}{Gmm} \times 100$$

**TABLE 1 Experimental Matrix**

| Surface Preparation/Treatment | 1" HMA | Nova Chip | Chip Seal |
|---|---|---|---|
| **Bare** | K-31/Anderson | K-27/Hamilton | US-160/Comanche |
| | US-159/ Atchison | US-50/Ford | K-44/Harper |
| **1" Surface Recycle** | K-27/Wallace | US-83/Decatur | NA |
| | US-36/Cheyenne | US-40/Geary | |
| **2" Surface Recycle** | US-24/Jefferson | US-24/Mitchell | K-99/Marshall |
| | K-61/Reno | US-50/Chase | US-24/Riley |

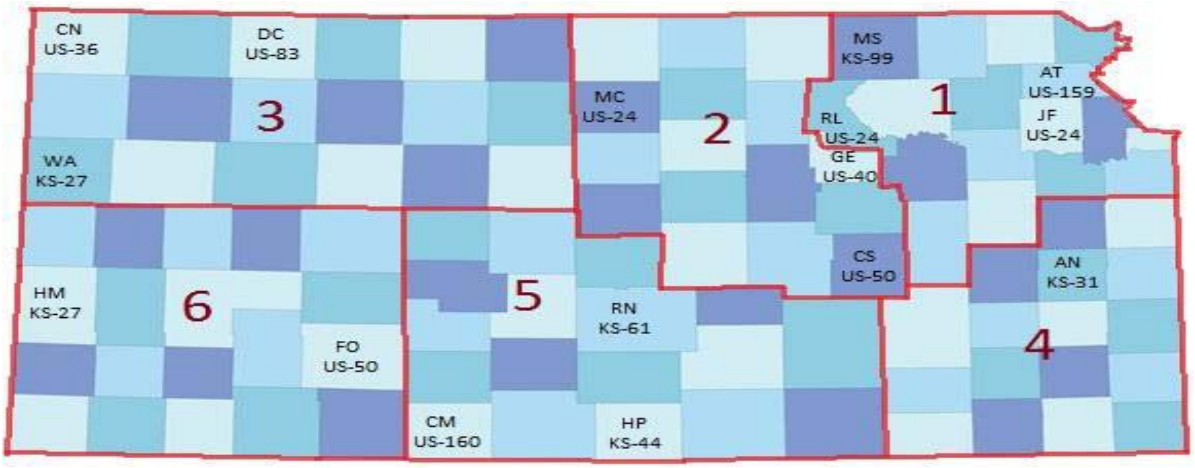



**Figure 1 Test section locations in Kansas**

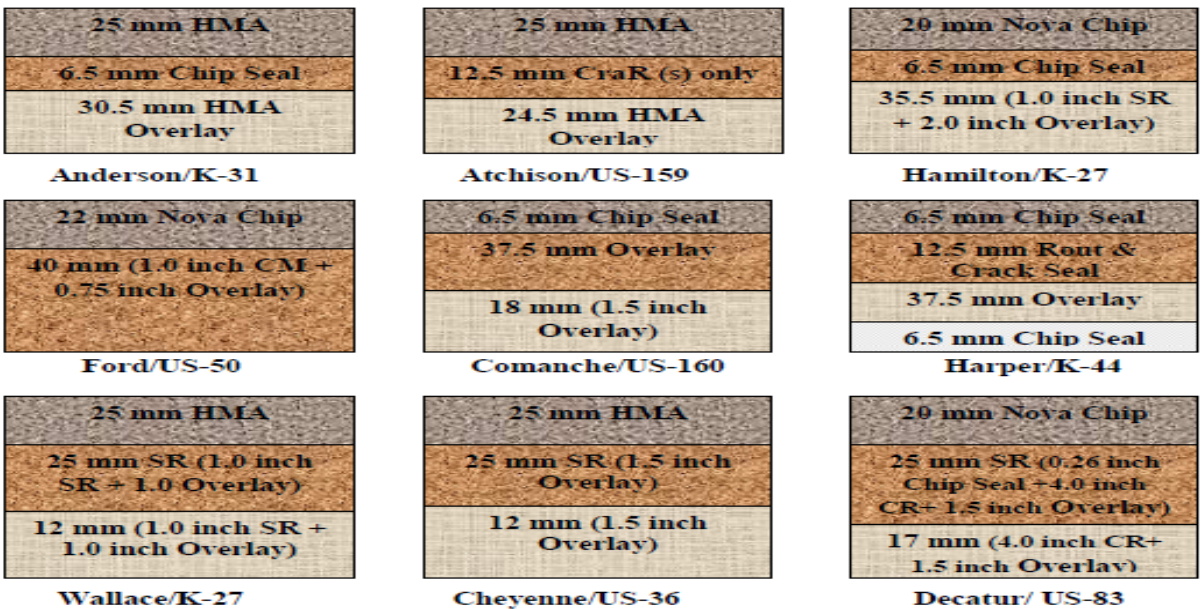



**Figure 2 Pavement cross-sections treated with thin surface treatments**

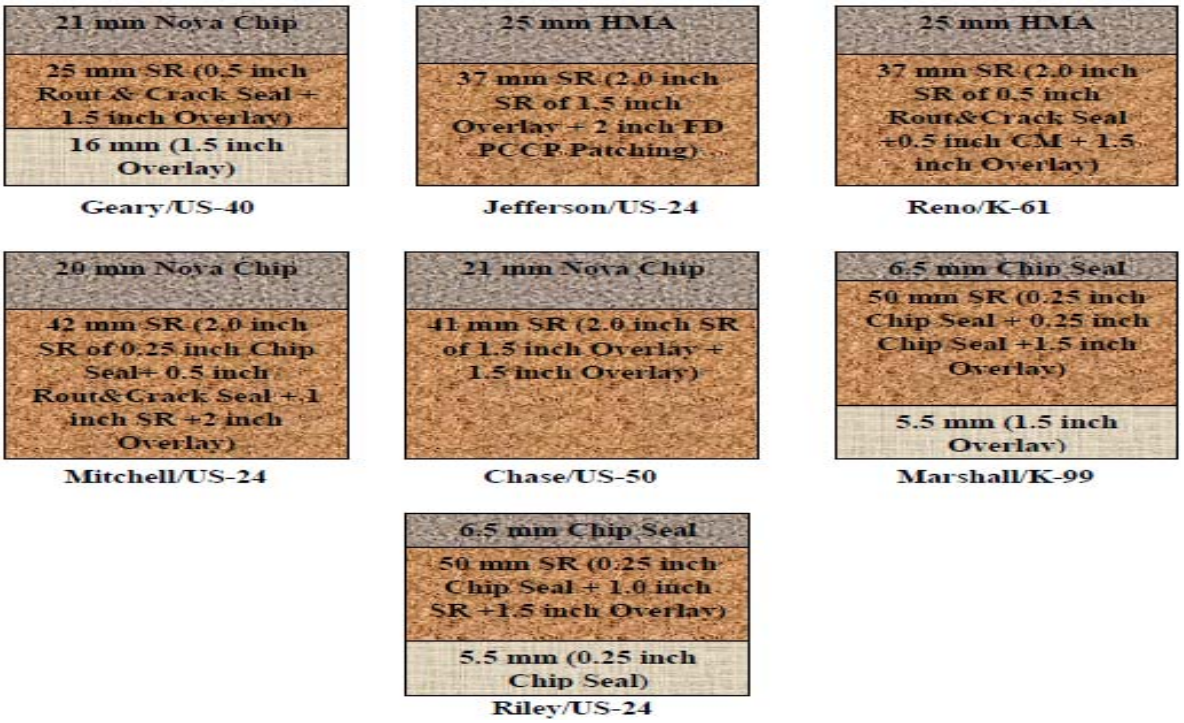


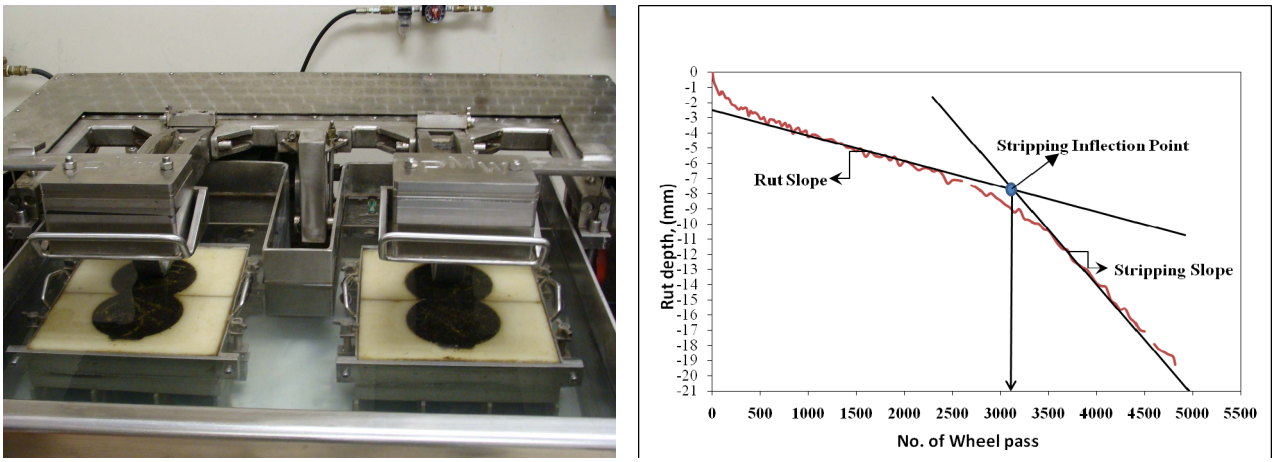

**Figure 2 Pavement cross-sections treated with thin surface treatments (contd..)**

## Hamburg Wheel-Tracking Device (HWTD) Test

HWTD tests were performed following Tex-242-F test method of the Texas Department of Transportation (TxDOT). The test is used to measure the combined effects of rutting and moisture damage by rolling a pair of steel wheels across the surface of asphalt concrete specimens immersed in hot water. Each moving steel wheel is 8 inch (204 mm) in diameter and 1.85 inch (47 mm) wide. The load applied by the wheel on the specimens is approximately 158±5 lbs (705±22 N), and the device operates at approximately 50 wheel passes per minute. The specimens used in this test should be 6 in. (150mm) in diameter and 2.5 ± 0.1 in. (62 ± 2mm) in height according to the TxDOT test method, Tex-242-F. The water temperature of the bath in which the specimens are submerged should be around 122±2 $^{o}$F (50±1°C). The rut depth measurement system consists of a linear variable differential transformer (LVDT) device. Rut depth measurement is taken at every 100 passes of the wheel.

HWTD automatically stops the test once the desired number of passes (20,000) or when the maximum allowable rut depth (20 mm) is reached. The depth of depression or rut created on the sample by HWTD test is measured and analyzed. The output from the HWTD test includes number of wheel passes to failure, and the number of passes at the rutting/creep slope, at the stripping inflection point and at the stripping slope. Figure 3 (b) shows typical output parameters obtained in the HWTD test. The rutting slopes (RS) is used to measure rutting susceptibility primarily due to mechanisms other than moisture damage, while the stripping slope (SS) measures moisture susceptibility primarily due to moisture damage. The stripping inflection point (SIP) indicates the number of wheel passes at which moisture damage starts to dominate

performance. Inverse slopes for both RS and SS are used so these slopes can be reported in terms of wheel passes along with the number of wheel passes at SIP. The higher the values of these slopes, the lesser the damage (TEX 242-F, 2009; Aschenbrener, 1995).




(a) (b)

**Figure 3 (a) Experimental set-up of HWTD test and (b) Interpretation of HWDT test results**

## RESULTS AND ANALYSIS

### HWTD Test Results Analysis

In this test, when the rut depth on the specimens reached 0.8 inch (20 mm), the sample was taken to have failed, and when the specimens carried 20,000 repetitions without exceeding the rut-depth of 0.8 inch (20 mm), the sample was assessed as a passed sample. Figure 4 shows the average number of wheel passes obtained from the HWTD tests along with the percent air voids. Laboratory test results showed that average air void of all projects treated with 1" HMA were below 10%, except the samples from US-24 in Jefferson County that had an average air void of more than 12%. All projects treated with Nova Chip showed varying average air voids, with a maximum average air void of 16.1% obtained on US-40 in Geary County and a minimum of 5.9% on US-50 in Ford County. All chip seal projects showed average air voids below 8%.

Based on the test results, most of the projects exceeded the maximum rut-depth limit (20 mm) at variable numbers of wheel passes. The core samples from K-27 in Hamilton County, US-83 in Decatur County, and K-61 in Reno County passed the test, carrying 20,000 wheel passes without exceeding the maximum rut-depth limit of 0.8 inch (20 mm). The core samples from K-27 in Wallace County carried the lowest number of wheel passes at the time of the failure. The biggest difference was also observed between the samples from the same location, such as US-24 in Riley County, which had a significant difference in left-wheel and right-wheel passes at the time of failure.

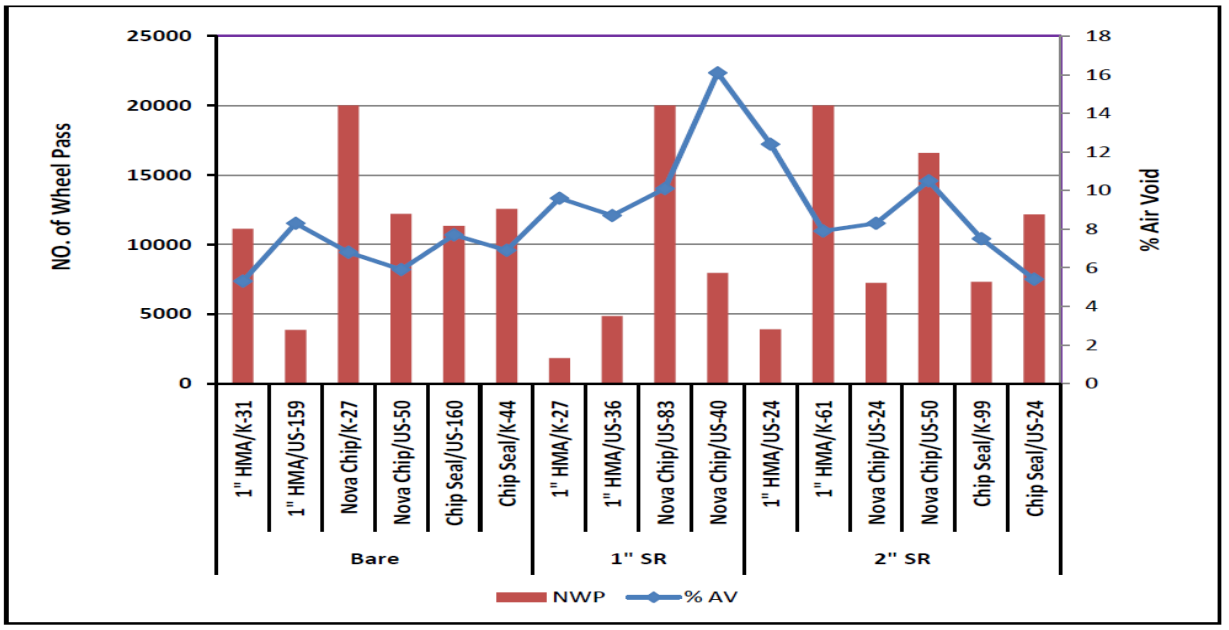



**Figure 4 Rutting performances in HWTD test along with percentage air voids**

Laboratory tests indicate that the performance of thin surface treatments varied a great deal for different projects and locations. For instance, samples from K-31 in Anderson County and US-159 in Atchison County, which were treated with 1" HMA on a bare surface, showed a significant difference in the number of wheel passes to failure, and in the number of passes at the rutting slope, stripping slope, and stripping inflection point. Similar trends were observed for the samples from different projects treated with Nova Chip, irrespective of surface types. Most of the chip seal projects, irrespective of surface types, exceeded the maximum rut depth limit (20 mm) at around 12,000 wheel passes. Nova Chip projects carried a minimum of 7,200 wheel passes from US-24 in Mitchell County, and a maximum of 20,000 wheel passes from K-27 in Hamilton County and US-83 in Decatur County, irrespective of surface types. Among projects treated with 1" HMA, irrespective of surface types, samples from K-27 in Wallace County showed a minimum of 1,825 wheel passes and a maximum of 20,000 passes was carried by samples from K-61 in Reno County.

Performance of thin surface treatments based on other parameters obtained from the HWTD tests also varied to a great extent on different projects and locations, as illustrated by Figure 5. From Figure 5 (a), it is evident that the rutting slope (averge-784 and standard deviation-242) of the highway sections treated with chip seal was consistent lower than that of 1" HMA (averge-1,099 and standard deviation-1,427) and Nova Chip (average- 2,318 and standard deviation-2,735). Among highway test sections treated with Nova Chip, cores from K-27 in Hamilton County showed a maximum rutting slope of 7,800 wheel passes per mm rut-depth, while a minimum value of 560 passes per mm rut-depth was obtained with US-24 in Mitchell County. Highway sections treated with 1" HMA had a maximum rutting slope of 3,800 wheel passes per mm rut-depth from K-61 in Reno County and a minimum of 150 wheel passes per mm rut-depth from K-27 in Wallace County, irrespective of surface types. The performance of thin surface treatments based on stripping slope, as shown in Figure 5 (b), shows almost similar trend to that of the rutting slope.

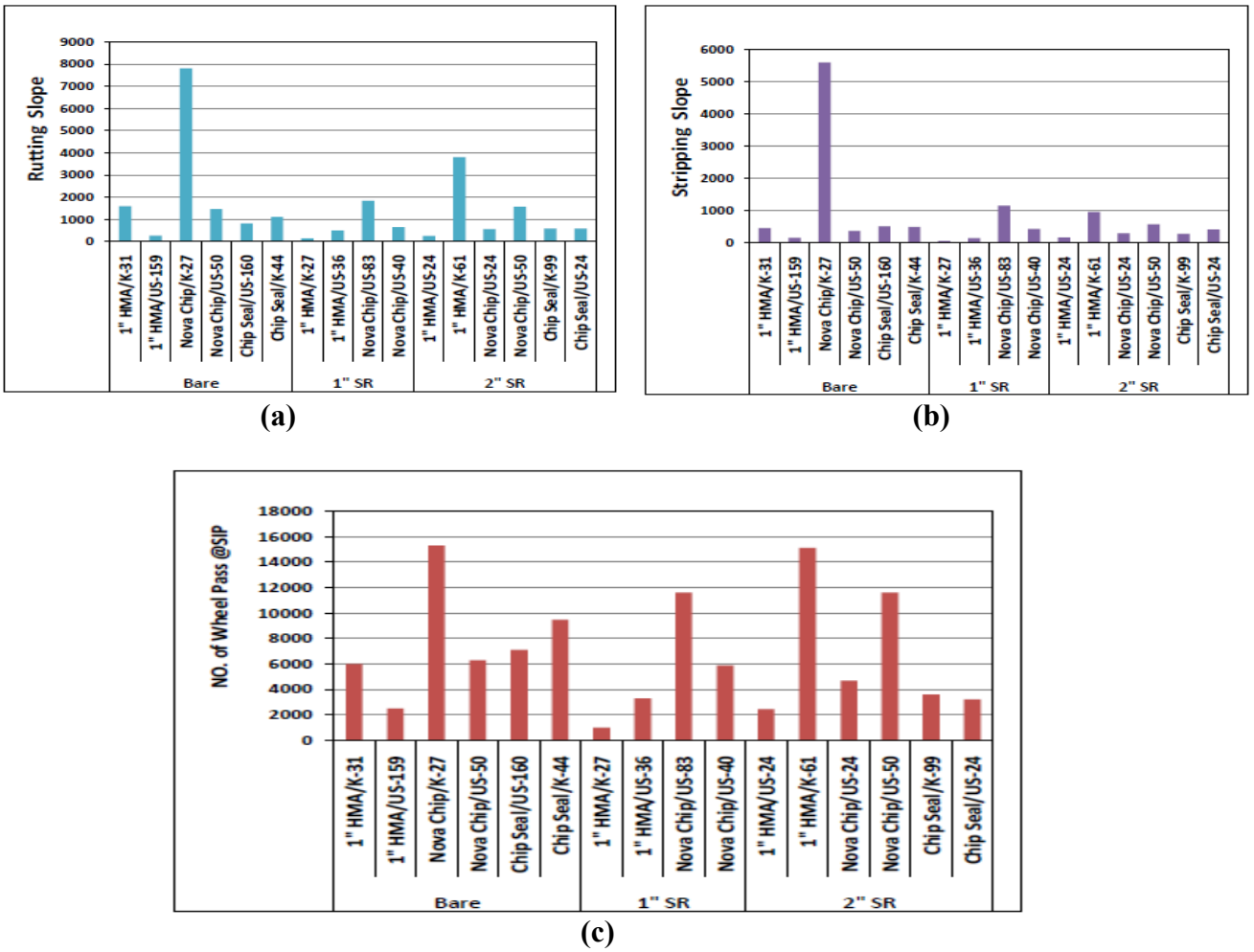



**Figure 5 Performance based on (a) rutting slope, (b) stripping slope, and (c) stripping inflection point**.

Figure 5 (c) shows the performance of thin surface treatments based on stripping inflection point (SIP). SIP is the number of wheel passes at which stripping is believed to have started and stripping dominates the performance. In general, highway sections treated with Nova Chip, irrespective of surface types, performed better in SIP compared to sections treated with 1" HMA and chip seal. An average SIP of 9,242 (standard deviation-4,207) wheel passes was obtained with the Nova Chip projects, followed by an average SIP of 5,851 (standard deviation-2,983) with chip seal, and 5,062 (standard deviation-5,201) with 1" HMA.

## Statistical Analysis

The analysis of covariance (ANCOVA) was conducted using statistical analysis software package SAS to determine the significant factors affecting the performance of thin surface treatments. ANCOVA is a more sophisticated method of analysis of variance (ANOVA) based on the inclusion of covariates into the model. The inclusion of covariate allows a reduction in observed variation between the treatments caused not by the treatment itself, but by a variation of covariate. Thus, the analysis of covariance evaluates the effect of the covariate on the response

variable and enables the comparison of treatments on a common basis relative to the values of the covariates (Kuehl, 2000). In this study, thin surface treatments with different surface preparations, as shown in Table 1, were independent variables. The dependent variable was number of wheel passes at 0.26 inch (6.5 mm) rut-depth as all chip seal projects were 0.26 inch (6.5 mm) thick. Average air voids (%) of the thin surface treatments were taken as covariates in the statistical analysis.

Pairwise comparisons or contrasts among treatments were performed in order to answer the question of whether one treatment was performing differently than the other one. To do the contrasts among treatments, null hypothesis $H_0$: LSMean (Ti) = LSMean (Tj), with alternative hypothesis Ha: LSMean (Ti) ≠LSMean (Tj), was developed based on the least squares (LS) mean values of the treatments. P-value was used to determine whether to accept or reject the null hypothesis by comparing it with a critical significance level (90% significance level).

Table 2 shows ANCOVA results using SAS. Average air voids (%) proved to be statistically significant at the 90% significance level based on the p-value (0.064). The overall treatment effect was found to be statistically insignificant (p-value=0.223) at a 90% significance level. Table 3 summarizes the results of contrasts among treatments where the null hypothesis was rejected; indicating the differences between treatments in number of wheel passes was statistically significant at the 90% significance level. From Table 3, it can be implied that the same treatment performed differently when applied on different surfaces, as did different treatments when applied on the same surface.

**TABLE 2 Results of ANCOVA**

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square | F | Pr >F |
|---|---|---|---|---|---|
| Treatment | 7 | 137742792.7 | 19677541.8 | 1.82 | 0.2235 |
| Air voids (%) | 1 | 51920809.6 | 51920809.6 | 4.81 | 0.0644 |
| Error | 7 | 75606449.9 | 10800921.4 | | |

**TABLE 3 Significant Contrasts between Treatments**

| Significant Contrasts between Treatments at 90% Significance Level |
|---|
| 1” HMA (bare surface) vs. 1” HMA (2” SR surface) |
| Nova Chip (bare surface) vs. 1” HMA (2” SR surface) |
| Chip Seal(bare surface) vs. 1” HMA (2” SR surface) |
| 1” HMA (1” SR surface) vs. 1” HMA (2” SR surface) |
| 1” HMA (2” SR surface) vs. Chip Seal (2” SR surface) |

## CONCLUSIONS

Based on this study, the following conclusions can be made:

- Pavements treated with thin surface treatments showed a high variability in the number of wheel passes to failure in the HWTD tests. Even cores from the same location had very different values of the number of wheel passes. While most of the projects exceeded the maximum rut depth (20 mm) at a variable number of wheel passes; three projects, two treated with Nova Chip and one with 1" HMA, were observed to pass the criterion of maximum 20,000 wheel passes without exceeding the maximum rut depth of 0.8 in. (20 mm).

- Stripping started earlier in the HWTD tests for highway test sections treated with 1" HMA and chip seal compared to the Nova Chip projects.

- Quality of pavement layer materials underneath thin surface treatments is the major determinant of performance of pavements with thin surface treatments.

- Air void was found to be a significant factor affecting the performance of thin surface treatments. Statistical analysis results also revealed that the type of treatment and surface preparation resulted in contrasting performance of thin surface treatments.

## ACKNOWLEDGEMENT

The authors wish to acknowledge the financial support provided by the University Transportation Center (UTC) at Kansas State University for this study.

## REFERENCES

Aschenbrener, T. *Evaluation of the Hamburg Wheel-Tracking Device to Predict Moisture Damage in Hot-Mix Asphalt.* Transportation Research Record 1492, Transportation Research Board, Washington, DC, 1995, pp. 193-201.

Eltahan, A., J. Daleiden, and A. Simpson. *Effectiveness of maintenance treatments of flexible pavements.* Transportation Research Record, n 1680, National Research Council, Oct. 1999: 18-25.

Gransberg, D. and D. M. B. James. *Chip Seal Best Practices: A Synthesis of Highway Practice.* NCHRP Synthesis 342, Transportation Research Board, Washington, D.C., 2005.

Kuehl, R. O. *Design of Experiments: Statistical Principles of Research Design and Analysis***,** Second Edition, 2000.

*State Testing Procedures, TEX 242-F Draft.* http://www.pmw-wheeltracker.com/test_procedures/Tex-242-F%20DRAFT.pdf (Accessed October 2009)

Shuler, S. *Evaluation of the Performance, Cost-Effectiveness, and Timing of Various Preventive Maintenance*s: Interim Report. Report No. CDOT-DTR-R-2006-6. May 2006.

Yang, H. *Pavement Analysis and Design*, Second Edition, 2004.

# Optimized Design for 4.75-mm NMAS Superpave Mix Thin Overlay

Farhana Rahman, Ph.D., EIT[1], Mustaque Hossain, Ph.D., P.E., F.ASCE[2], Stefan A. Romanoschi, Ph.D., P.E., M.ASCE[3], and Cliff Hobson, P.E.[4]

[1]Kansas State University, Department of Civil Engineering, Manhattan, KS 66506, Tel. No. (701) 212-0360, Fax No. (785) 532-7717, e-mail: farhana@ksu.edu
[2]Professor, Kansas State University, Department of Civil Engineering, Manhattan, KS 66506, Tel. No. (785) 532-1576, Fax No. (785) 532-7717, e-mail: mustak@ksu.edu
[3]Associate Professor, University of Texas at Arlington, Department of Civil and Environmental Engineering, UTA Box 19308, Arlington, TX 76019-0308, Tel. No. (817) 272-0340, Fax No. (817) 272 2630, e-mail: romanoschi@uta.edu
[4]Kansas Department of Transportation, Materials & Research Center, 2300 Van Buren, Topeka, KS 66611, Tel. No. (785) 291-3843, Fax No. (785) 296-2526, e-mail: cliff@ksdot.org

## ABSTRACT

Transportation infrastructure facilities such as highways consume large quantities of materials in initial construction and periodic maintenance and rehabilitation. A Superpave asphalt mixture with 4.75-mm nominal maximum aggregate size (NMAS) is a promising, low-cost pavement preservation treatment that can be applied in thin lifts for low- to medium-volume facilities. This study evaluates the performance of 4.75-mm NMAS Superpave asphalt mixture and proposes an optimized mixture design for a long-lasting ultra thin hot-mix asphalt (HMA) overlay.

Twelve different 4.75-mm NMAS mix designs were developed in the laboratory using two binder grades and three different percentages of natural (river) sand in combination with finer quarry materials. Test results from two laboratory performance tests show that rutting and moisture damage of such mixes during Hamburg Wheel Tracking Device and Lottman tests are aggregate-type and source specific. Anti-stripping agent affects moisture sensitivity test results irrespective of natural sand content and binder grade. Finally, optimum combinations of aggregates and binder were developed for 4.75-mm NMAS mixture based on the statistical analysis of performance data.

## INTRODUCTION

The common distresses on asphalt pavements can be partly addressed by proper selection of construction materials and by developing suitable mix design. Hot-mix asphalt (HMA) mixes with smaller aggregate size are commonly used in pavement preservation and maintenance projects. In 2002, the National Center for Asphalt Technology (NCAT) developed Superpave mix design criteria for the 4.75-mm nominal maximum aggregate size (NMAS) (Cooley et al. 2002). In a corrective maintenance program, a fine-graded mixture is well accepted in leveling and shimming of the existing pavement before overlay application. For preventive maintenance, thin-lift application of the fine mix is an excellent alternative to stretch

the maintenance budget if the pavement does not experience any major distresses. This application primarily improves ride quality, reduces permeability, and sometimes leads to minor crack healing.

## PROBLEM STATEMENT

With the advent of Superpave, many state agencies recommended coarse-graded mixtures and some agencies have begun to use stone mastic asphalt (SMA) mixes, which in turn, reduce the use of fine materials (West et. al. 2006). Ultra thin-lift application of HMA with 4.75-mm NMAS mixture is a fairly new concept in pavement preservation. The implementation of 4.75-mm NMAS Superpave mix will reduce the accumulated screening stockpiles and hence, provide a use for materials that could become a "by-product" of the HMA industry. A recent survey suggested that coarse-graded Superpave mixes result in higher permeability compared to the dense-graded mixes at the same air void and material selection plays a significant role in reducing the problem (Mallick et al. 2003). However, potential limitations of fine mixtures include concerns with rutting, moisture susceptibility, gouging, scuffing, and skid resistance. In 2002, the 4.75-mm NMAS designation and criteria were added to the AASHTO Superpave specifications to fit the need for small aggregate size mixtures. However, these criteria were based on a combination of experience, limited laboratory testing and engineering judgment.

## OBJECTIVE

The main objective of this research study was to evaluate performance of 4.75-mm NMAS Superpave mixes with varying river sand content and binder grade. The performance was assessed in terms of rutting and moisture sensitivity. Statistical analysis was performed to develop regression equations to characterize rutting potential and moisture sensitivity of the 4.75-mm NMAS mixture. Finally, the study introduced an optimized design procedure for a long-lasting ultra thin HMA overlay based on the developed performance equations.

## FINE MIX SPECIFICATION IN KANSAS

Currently 4.75-mm NMAS Superpave mixture is designated as SM-4.75A in Kansas. Gradation of this mixture is selected to pass over the maximum density line on a 0.45-power chart in sand sizes and thus, the mixture is considered fine. This gradation is essentially controlled by the materials retained on 1.18-mm (No. 16) and 0.075-mm (No. 200) sieves. Current KDOT specifications also allow up to 35% natural sand provided the fine aggregate angularity (FAA) of the blend meets the required criteria (FAA $\geq$ 42). Higher dust-to-effective binder ratios are allowed in these mixtures. The required mixture design criteria are shown in Table 1.

**Laboratory Mix Design of 4.75-mm NMAS Superpave Mixture.** Two rehabilitation projects were done in Kansas in 2007 that used 4.75-mm NMAS Superpave mixes. Laboratory mix designs were developed in this study using

materials from these projects. All mixtures would have to have 4% air voids with $N_{des}$ level at 75 gyrations. This compaction effort was selected as 4.75-mm NMAS mix is normally used for low-volume traffic conditions (ESALs less that 3 millions). Variations of these mix designs were planned by changing the binder grade and also by varying natural sand content in the combined mix for two different aggregate sources in Kansas. The baseline 4.75-mm NMAS mixture designs were obtained from the US-160 and K-25 rehabilitation projects. Twelve different mix designs were developed considering two aggregate sources, two binder grades and three different natural sand contents. Anti-stripping agent was used in the mixes for the US-160 project since the baseline mixture also had an anti-stripping agent. For each mix combination, tests were done for rutting and moisture sensitivity.

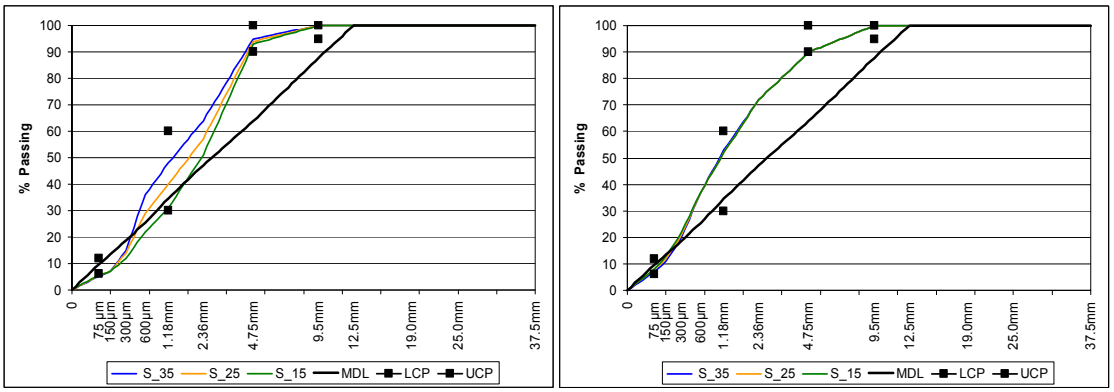
**Table 1. Mixture Design Criteria for Kansas 4.75-mm NMAS Superpave Mix (Hossain et al. 2010).**

| Criteria | Specifications | Comments |
|---|---|---|
| | **Compaction Effort** | |
| $N_{ini}$<br>$N_{des}$ & $N_{max}$ | Function of 20-year design ESALs | Similar to all other Superpave mixes |
| | **Volumetric Properties** | |
| Air voids | 4% ± 2% at $N_{des}$ | Similar to all other Superpave mixes |
| VMA | 16% min. for reconstruction/major modification project | may be reduced by 1% for 1-R jobs |
| VFA | 65-78 | Function of 20-year design ESALs |
| $\%G_{mm}$ @ $N_{ini}$ | 90.5 | Function of 20-year design ESALs and layer depth |
| $\%G_{mm}$ @ $N_{max}$ | 98.0 | Similar to all other Superpave mixes |
| Dust-to-binder ratio | 0.9 to 2.0 | 0.6-1.2 or 0.6-1.8 |
| Tensile Strength Ratio, min. (%) | 80 | 80 |

The AASHTO standard practice (R 35-4), *Superpave Volumetric Design for Hot-Mix Asphalt (HMA)*, was followed during the mix design phase of this study (AASHTO. 2004). This standard was used to evaluate the 4.75-mm mixture properties following KDOT volumetric specifications for SM-4.75A mix. At first, gradation analysis was performed on all materials following AASHTO T27, *Sieve Analysis of Coarse and Fine Aggregate,* and AASHTO T11, *Materials Finer than 75 µm (No. 200) Sieve in Mineral Aggregate by Washing.* Four trial aggregate blends were developed for 35%, 25% and 15% natural sand that satisfied KDOT gradations for SM-4.75A mixture. The fine aggregate angularity (FAA) of the combined gradation was determined by the KT-50 test procedure for each combination. Control points for the 4.75-mm sieve (100-90% passing) were strictly observed in the

blending process to maintain a true 4.75-mm NMAS Superpave mixture. Figure 1 shows the gradations developed in this study.

For statistical and experimental design purposes, aggregates from each source were again subdivided into three major categories. Based on aggregate particle-size distribution and percent fines retained on the 0.075-mm (No. 200) sieve, the subsets were defined as coarse material (among groups), screening material, and river sand (Table 2). After selecting aggregate blends for 35%, 25%, and 15% river sand content, design asphalt content for each gradation was determined considering two different binder grades (PG 64-22 and PG 70-22). Detailed individual aggregate gradations can be found somewhere else (Rahman 2010).




**(a)** **(b)**

**Figure 1. 0.45 power charts for 4.75-mm NMAS Superpave laboratory mixture (a) US-160 and (b) K-25.**

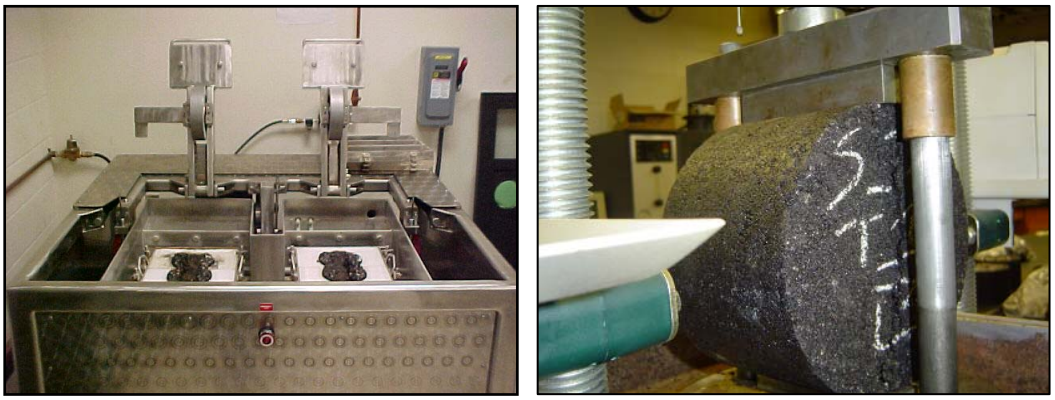
**Table 2. Aggregate Subsets on US-160 and K-25.**

| Source | Aggregate Subsets, (%) | | | | | |
|---|---|---|---|---|---|---|
| | Coarse Material | | Screening Material | | River Sand | |
| | Max. | Min. | Max. | Min. | Max. | Min. |
| **US-160** | 45 | 32 | 33 | 26 | 35 | 15 |
| **K-25** | 40 | 30 | 43 | 33 | 35 | 15 |

## PERFORMANCE TESTS

**Hamburg Wheel Tracking Device Rutting Evaluation (TEX 242-F Draft 2009).** Rutting or permanent deformation of the designed mixtures was evaluated using the Hamburg Wheel Tracking Device (HWTD) and following Tex-242-F test method of the Texas Department of Transportation. This wheel tracking equipment is operated under the mechanism that a pair of wheels apply moving loads to the specimen in order to simulate rutting in an accelerated manner. The depth of depression or rut created on the sample is measured and analyzed. Tex-242-F evaluates the premature failure susceptibility of bituminous mixture due to weakness in the aggregate skeleton, moisture damage and inadequate binder stiffness. The test measures the depression and number of wheel passes to failure (Figure 2a). Each moving steel wheel of HWTD is 203.6 mm (8 inches) in diameter and 47 mm (1.85 inches) wide. The load applied by the wheel is approximately $705 \pm 22$ N ($158 \pm 5$ lbs) and the

wheel passes over the test specimen approximately 50 times per minute. The water control system of HWTD is capable of controlling the test temperature from 25 to $70^0$C (77 to $158^0$ F) with a precision of $\pm 2^0$C ($4^0$F). The rut depth measurement system consists of a linear variable differential transformer (LVDT) device. Rut depth measurement is taken after every 100 wheel passes.



**(a)** **(b)**

**Figure 2. Experimental set up and tested samples in (a) Hamburg Wheel Tester, and (b) TSR load frame.**

**Moisture Susceptibility Test (KT-56) (Hossain et. al. 2010).** This test is used to measure the change in tensile strength resulting from the effects of saturation and accelerated water conditioning of the compacted bituminous mixture in the laboratory. It evaluates the ability of the compacted bituminous mix to withstand long-term stripping action and also assesses the liquid anti-stripping additives used in the asphalt mix. Kansas test procedure KT-56, *Resistance of Compacted Bituminous Mixture to Moisture Induced Damage*, a slightly modified version of AASHTO T283, was followed in this study. The test specimens are prepared using the Superpave gyratory compactor. Vacuum container with a minimum diameter of 200 mm (8 inches) and the inside height capable of holding a minimum of 25 mm (1 inch) of water above the specimen is selected to achieve percent saturation (65-85%). A vacuum pump that applies a 30-mm of Hg absolute pressure is also attached to the vacuum container. The test compares the indirect tensile strength test results of a dry sample and a sample exposed to water and freezing/thawing cycle. The specimens are conditioned at $60 \pm 1^0$C ($140 \pm 2^0$F) as well as at $-18 \pm 3^0$C ($0 \pm 5^0$F).

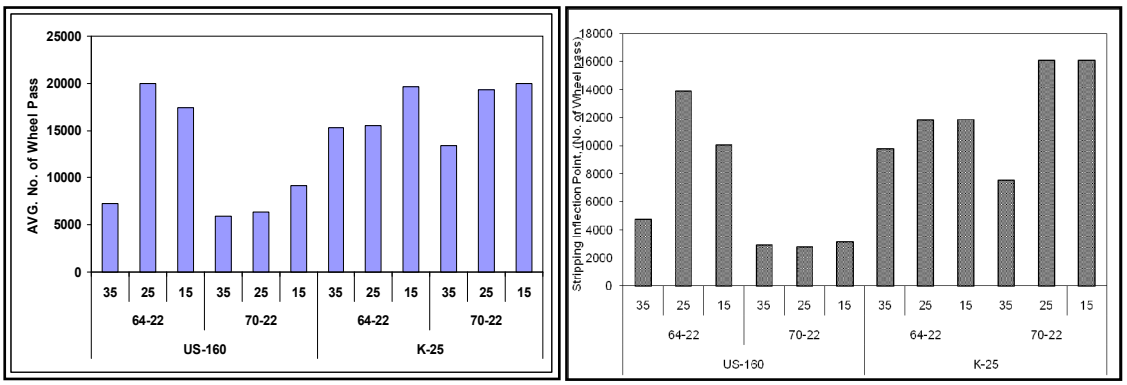
At least six specimens with air voids of $7\% \pm 0.5\%$ are required. The specimens are 150 mm (6 inch) in diameter and $95 \pm 5$ mm ($4 \pm 0.2$ inch) thick. The results are reported as the tensile strength ratio (TSR) of average conditioned sample tensile strength to average dry sample tensile strength. KDOT specification requires a minimum TSR of 80% for the HMA mix not to be potentially moisture sensitive (Figure 2b).

## RESULTS AND ANALYSIS

**Rutting Performance.** The Hamburg Wheel Tracking Device (HWTD) was used to evaluate rutting and stripping performance of all 12 mixes. Three replicates were

produced for a particular mix design to obtain unbiased test results. The specimens had air voids of 7±1% and were tested at $50^0$ C. Figure 3 shows mix performance with respect to the HWTD test output parameters such as the average number of wheel passes and stripping inflection point (number of wheel pass before stripping).

Figure 3 shows that natural sand content was an important factor affecting rutting performance of laboratory mixes. In general, the number of wheel passes increased with decreasing natural sand content. Also, mix performance was aggregate source specific. In most cases, there is no significant difference between performance of mixes with 25% and 15% natural sand. Binder grade did not appear to affect the mixture performance appreciably (Figure 3a). Figure 3b illustrates that stripping inflection point for a particular mix was highly aggregate-source specific. Better aggregate structure (granite aggregate structure from K-25 aggregate source) may help defer the onset of stripping. Most of the K-25 mixes experienced delayed stripping distress compared to the US-160 mixes. Among all laboratory mixes, the K-25 mix with PG 70-22 binder and 15% to 25% sand content carried the highest number of wheel passes till stripping. The average number of wheel passes increased more than 50% compared to the mix with 35% river sand.




**(a)** **(b)**

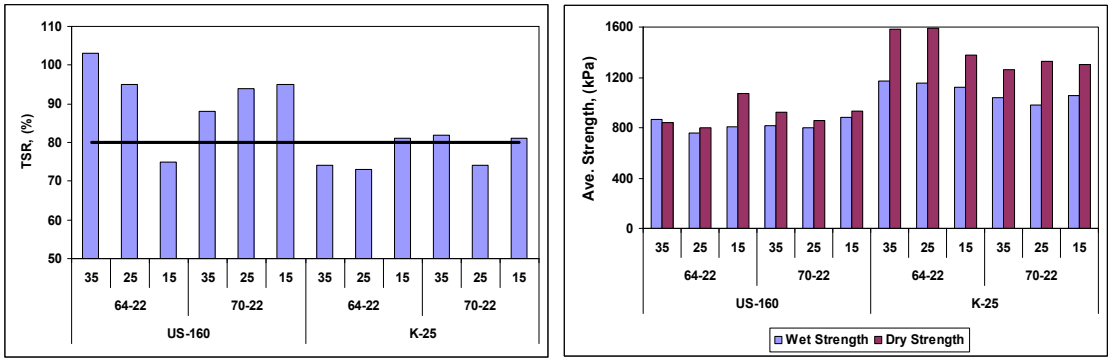
**Figure 3. HWTD laboratory mix performance (a) average no. of wheel pass and (b) stripping inflection point at different sand content and binder grade.**

However, the performance of mix with PG 70-22 binder grade on US-160 was notably different than that of the mix with PG 64-22 binder. The number of wheel passes was significantly lower during HWTD testing. This may indicate that anti-stripping agent might degrade a PG binder (Rahman 2010, Rahman et al. 2010).

**Tensile Strength Ratio:** For all twelve mixes designed in the laboratory, tensile strength ratio (TSR) was determined as per KT-56. For this test, specimens were compacted at 7±0.5% air voids. Six samples were compacted for a particular mix design: three samples were conditioned (freeze/thaw) and three were unconditioned. All six were tested for tensile strength in indirect tension mode. The ratio of the average tensile strength of the conditioned to that of the unconditioned samples was considered as the performance measure during testing.

Figure 4 shows a plot of TSR and comparison of dry and wet tensile strengths for all twelve mixes. In general, mixes with anti-stripping agent (as on US-160) had

higher TSR values compared to mixes with no anti-stripping agent (as on K-25). All mixes on US-160 passed the minimum TSR requirements specified by KDOT with the exception of the mix with 15% natural sand and PG 64-22 binder. The average TSR for mixes with PG 64-22 binder on US-160 was 91% while an average of 92% was achieved for mixes with PG 70-22. This implies that the effectiveness of anti-stripping agent depends on binder grade and aggregate source. Fifty percent of the design mixes on K-25 failed to meet the required TSR criteria. The average TSR on K-25 ranged from a minimum of 73% to a maximum of 81% for the mixes with PG 64-22 binder and a minimum of 74% to a maximum of 82% for the mixes with PG 70-22 binder. Although the dry and wet strength of mixes on K-25 were significantly higher than that of US-160, their ratio failed to meet the minimum TSR requirement.




**(a)** **(b)**

**Figure. 4 (a) Tensile strength ratios (b) dry and wet strength comparison.**

## REGRESSION ANALYSIS OF 4.75-mm NMAS MIX PERFORMANCE

Regression analysis is a statistical technique of modeling dependency of a response variable on one or more independent variables (Weisberg 2005). The goal of regression analysis in the current study is to develop the regression functions/equations for laboratory-designed mixtures based on the performance test results. The aggregate source, binder grade, and natural sand content were considered as independent variables while the number of wheels passes and tensile strength ratio, respectively obtained from the HWTD and moisture sensitivity tests, were considered as the response variables.

**Rutting Prediction Equation.** The rutting regression function was developed based on HWTD test data (Figure 3a). As mentioned earlier, the number of wheel passes (NWP) is always considered as the response variable while aggregate subsets, binder grades and river sand content were considered as the independent variables in the regression equation. Two steps were followed to develop the rutting prediction equations. In the initial phase, forward selection method was used to identify the influential design factors to be considered in the regression equation. Independent variables were selected based on the coefficient of determination ($R^2$), overall F-statistics, and p-values. For a specific mix design, the independent variables, such as, coarser material (CA1), screening material (CA2), and river sand content (NSC) are somewhat correlated. Hence, in a particular regression function, either CA1 or CA2

or NSC and binder grade (PG) were considered during the goodness test. The $R^2$ (0.71) and p-value (<0.0001) of the US-160 mix proved that both PG and CA1 could be the best design variables to fit the design function. On the other hand, binder grade did not have any potential influence on the K-25 mixes. Based on $R^2$ (0.31) and p-value (0.018), screening material (CA2) was selected as the best design variable to develop the rutting prediction model. In the second phase, the following two multiple linear regression equations with interaction terms were developed to predict the rutting performance of the mix. The independent variables of the regression equations such as coarse materials (CA1) and screening materials (CA2) are in percentages by weight of total aggregate. The binder grade is either 0 (PG 64-22) or 1 (PG 70-22).

$$NWP(US-160) = -18516 + 16639 \times PG + 856.4 \times CA1 - 624.66 \times PG \times CA1 \quad (1)$$
$$R^2 = 0.80 \qquad p-value < 0.0001$$
$$NWP(K-25) = -3791.03 + 547.26 \times CA2 \quad (2)$$
$$R^2 = 0.30 \qquad p-value = 0.0180$$

**Moisture Sensitivity Prediction Equation.** The prediction equation to assess moisture sensitivity was developed based on wet-to- dry strength ratio (TSR) data obtained from the Lottman tests. Fifty percent of the K-25 laboratory mixes failed to meet the TSR criterion (Figure 4). Therefore, regression analysis was performed only for US-160 mixes. The aggregate subsets, binder grades and river sand content were considered as the independent variables while tensile strength ratio (TSR) was taken as the response variable. During variable selection phase, forward selection method discarded the binder grade (PG) at a cutoff point of 0.5 for all aggregate combinations. The F-statistic (1.63) and p-value (0.2712) of the CA2 subset described the design equation were better compared to CA1 and NSC. Since binder grade was not selected to be included in the regression equation, an interaction term between CA2 and PG was introduced into the regression equation to check for improved $R^2$ and p-value. The best-fit moisture damage prediction model selection criteria were set based on the coefficient of determination ($R^2$) of the overall model and p-values of estimated parameters. Addition of interaction variables (between PG and CA2) in the multiple linear regression equations significantly improved the coefficient of determination ($R^2$ =0.98) and p-values of individual estimated parameters. Equation (3) represents the moisture damage prediction model in-terms of tensile strength ratio (TSR, %)). The independent variables such as screening materials (CA2) are in percent by weight of total aggregates and binder grade (PG) is considered either 0 (PG 64-22) or 1 (PG 70-22).

$$TSR = 207 - 139.205 \times PG - 4.0 \times CA2 + 4.85 \times PG \times CA2 \quad (3)$$
$$R^2 = 0.98 \qquad p-value < 0.0001$$

## OPTIMIZATION OF 4.75-MM NMAS MIX PERFORMANCE

The main objective of the optimized design procedure was to select the best combination of aggregate blending from some set of available aggregate subsets. The

selection technique was based on maximizing or minimizing the real functions, called objective functions, by sequentially choosing the values of real or integer variables from available subsets. In the present study, the multi-objective optimization technique was used to identify the significant aggregate subsets and binder grade combinations, which can address the common distresses on HMA pavements. The objective functions for the optimization process were the regression equations (eq.1, 2, and 3) described in the previous sections. The goals and constraints of the objective functions considered during the optimization process are presented in Table 3.

**Table 3. Constraints and Goals for Objective Functions (Rahman 2010).**

| Projects | Constraint Functions | Goals |
|---|---|---|
| **US-160** | $0 \le PG \le 1$<br>$32 \le CA1 \le 45$<br>$26 \le CA2 \le 33$<br>$15 \le NSC \le 35$<br>$CA1 + CA2 + NSC = 93$ | $NWP \ge 20{,}000$<br>$TSR \ge 80\%$ |
| **K-25** | $0 \le PG \le 1$<br>$30 \le CA1 \le 40$<br>$33 \le CA2 \le 43$<br>$15 \le NSC \le 35$<br>$CA1 + CA2 + NSC = 98$ | $NWP \ge 20{,}000$ |

The goals of objective functions for rutting and moisture damage were to minimize these distresses by maximizing the number of wheel passes (NWP) and tensile strength ratio (TSR) higher or equal to 20,000 repetitions and 80%, respectively. The constraint functions were based on the upper and lower limits of the individual aggregate subsets. The binder grade PG was considered to be continuous within 0 to 1 limit. The summation of coarse, screening, and river sand materials within the aggregate blend were 93% and 98% for US-160 and K-25 mixes, respectively. During this optimization problem, the feasible solutions generated unique values of binder grade, percentage of coarser material, screening material, and river sand content considering the limits and constraints.

The multi-objective optimization process proposed feasible aggregate and binder combinations to address all three major distresses on the asphalt pavement at each location. For US-160 mixes, binder grade PG 64-22 proved to be effective over PG 70-22. Seventeen to 22 percent river sand in the designed aggregate blend was sufficient to produce optimized design combinations instead of using 35% natural sand (allowable state practice). On the other hand, K-25 mixes with 15 to 20 percent natural sand content were found to produce the optimized mix design. Higher binder grade PG 70-22 was more effective for the K-25 aggregate source compared to PG 64-22.

## CONCLUSIONS

Optimized mix design developed for 4.75-mm NMAS Superpave mixture primarily focused on aggregate/binder combination and completely addressed the common

distresses found on asphalt pavements. However, the optimum aggregate-binder combinations are different from the in-place mixes used on two projects. Investigation on RAP was outside the scope of this research study. Moreover, because of larger aggregates in RAP, it is doubtful whether RAP can be used in 4.75 mm NMAS mixture without some kind of fractionation of RAP. Again the process would add to the cost. Based on this study, the following conclusions can be made:

- Rutting performance during the Hamburg wheel tracking device tests was aggregate source specific. Effect of higher binder grade on rutting performance is inconclusive.
- The anti-stripping agent affected the moisture sensitivity test results, irrespective of natural sand content, binder grade, and aggregate source. Mixes without anti-stripping agent failed to meet the Tensile Strength Ratio (TSR) criteria specified by the Kansas Department of Transportation.
- Optimized design combinations suggested limiting the river sand content in between 15% and 20% rather than 35% (current practice) for the Kansas 4.75-mm NMAS Superpave mixture.

## REFERENCES

AASHTO. (2004). "Specifications. Standard Specifications for Transportation Materials and Methods of Sampling and Testing." *24th Edition Part B*, 444 North Capitol Street, NW, Washington, D.C.

Cooley, L. A., James, R. S., and Buchanan, M. S. (2002). "Development of Mix Design Criteria for 4.75 mm Superpave Mixes-Final Report." *NCAT Report 02-04*. National Center for Asphalt Technology, Auburn University, Alabama.

Hossain, M., Maag, R. G., and Fager, G. (2010). "Handbook of Superpave Volumetric Asphalt Mixture Design and Analysis." *Superpave Certification Training Manual*, Kansas State University, Manhattan, Kansas.

Mallick, R. B., Cooley, L. A., Bradbury, R. L., and Peabody, D. (2003). "An Evaluation of Factors Affecting Permeability of Superpave Designed Pavements." *NCAT Report 03-02.* National Center for Asphalt Technology, Auburn University, Auburn, Alabama.

Rahman, F. (2010). "Performance Evaluation of 4.75-mm NMAS Superpave Mixture." *Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy*, Department of Civil Engineering, Kansas State University, Manhattan, Kansas.

Rahman, F., Hossain, M., Romanoschi, S. A., and Hobson, C. (2010). "Evaluation of 4.75-mm Superpave Mixture." *TRB Preprint CD-ROM Paper #10-2608,* 89th Annual Meeting of the Transportation Research Board, Washington, D.C.

TEX 242-F Draft. http://www.pmw-wheeltracker.com/test_procedures/Tex-242-F%20DRAFT.pdf. Accessed July 15, 2009.

Weisberg, S. (2005). "Applied Linear Regression." *3rd Edition*, John Wiley & Sons Inc. Publication, 111 River Street, Hoboken, New Jersey, USA.

West, R. C., and Rausch, D. M. (2006). "Laboratory Refinement of 4.75 mm Superpave Designed Asphalt Mixture." *NCAT Phase I Draft Report.* National Center for Asphalt Technology, Auburn University, Auburn, Alabama.

# Mechanical Response of Modified Asphalt Pavements

S. Anjan kumar[1], P. Alagappan[2], J. Murali Krishnan[3], and A. Veeraragavan[4]

[1]Ph.D. Research Scholar, Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India. PH (91) 044-22575292; e-mail: anjaankumar@yahoo.co.in.
[2]M.Tech. Student, Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India, PH (91) 044-22575292, e-mail: alagappan.ce@gmail.com
[3]Associate Professor, Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India, PH (91) 044-22574284, e-mail: jmk@iitm.ac.in (contact author)
[4]Professor of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India, PH (91) 044-22574282, e-mail: av@iitm.ac.in

## Abstract

In this work, pavement cross-section as stipulated by the Indian Roads Congress code of practice for pavement design was used in the stress-analysis. Typically, two types of asphalt layers are laid over three to four layers of granular materials for pavements constructed for heavy traffic volume in India. Asphalt mixtures pertaining to these top two layers were fabricated with polymer modified and unmodified asphalts in the laboratory. These mixes were tested in the simple performance test equipment at temperature of 60 $^{o}$C under a wide range of frequencies. A four parameter Burgers' model was used for fitting the experimental data for both the asphalt layers. A two-dimensional finite element model of the pavement structure was used within the ABAQUS computing environment. The granular base, subbase and subgrade granular layers were assumed to be elastic. The pavement model was subjected to cycles of periodic loading and the stresses and strains were monitored at all the critical locations. This was used to quantify the influence of modifiers on the mechanical response of the pavement and parametric analysis was carried out to find out enhanced service life of the asphalt layers that could be achieved due to the use of high-performance materials.

**Keywords**: Viscoelasticity; Finite element analysis; Modified asphalt; Pavement engineering

## Introduction

Stress-strain analysis of an asphalt pavement structure subjected to random traffic load and varying environmental conditions is a complex task. One important factor which plays the critical role is the use of appropriate material constitutive relation in the structural model. Since the asphalt mixtures exhibit viscoelastic stress-strain behavior, identification of tractable constitutive model and using them within an appropriate computational framework becomes a critical issue. This becomes all the more important when one is interested in quantifying the beneficial effects of asphalt layers constructed with modified asphalts. Finite element modeling of pavements, if validated can be extremely useful, because it can be used directly to estimate primary response parameters without resorting to potentially costly field experiments (Helwany et al. 1998).

Different types of modeling attempts are reported in the literature as far as characterization of viscoelastic property of asphalt mixtures are concerned. For instance one can see the use of Boltzman's superpostition principle with time-dependent function (Papagiannakis et al. 1996), time dependent shear modulus functions (White et al. 1997), generalized Kelvin model at intermediate and high temperature (Elseifi et al. 2006), generalized Maxwell model (Mulungye et al. 2007) and elasto-visco-plastic model (Kettil et al. 2007) to mention a few.

Due to the complex relationship between loading type and rate on the response of asphalt pavement, issues related to quantifying the damage based on static load or dynamic load needs considerable clarification. For instance Papagiannakis et al. (1996) found the relative damage to range from 1.2 to 2.8 times depending on whether a static or dynamic load was applied. It was found that axle configurations (single axle and dual tandem) substantially influence the primary response of pavement structure and longitudinal strains are highly sensitive to tire pressure both at the top and bottom of viscoelastic layer (Helwany et al. 1998). The stress state resulting from a radial tire was found to be larger in magnitude and focused near the surface than those obtained from the traditional vertical loading conditions. These high shear stresses near surface could be an explanation for rutting failure in asphalt mixes as discussed by Novak et al. (2003). It was also found that cyclic loading on non-linear granular layers results in large deflection on subgrade than the static loading (Hadi and Bodhinayake, 2003).

Perkins and Edens (2002) introduced a parameter called traffic benefit ratio (TBR) and defined it as the ratio of vertical compressive strain on the subgrade with and without geo-synthetic reinforcement using 12.5 mm rutting as the failure criteria. This ratio was incorporated in an empirical pavement damage model and was used to explain the enhancement in pavement life.

Modification to the base asphalt is carried out to improve the rutting and fatigue cracking resistance of the material. It was found that by using styrene-butadiene-styrene polymer (SBS) modified asphalt in the top layer alone, the service life could be enhanced by 1.295 times and thickness reduction of 30 % could be achieved (Hadidy and Tan, 2009). Also use of polypropylene modified asphalt in stone matrix asphalt (SMA) was found to enhance the service life by 1.48 times and a total thickness reduction of 40 % could be achieved in SMA layer alone (Hadidy and Tan, 2009). It is also reported in the literature that decrease in thickness can also result in

increased strain at the top of the layers leading to top-down cracking (Kim et al. 2009).

As seen from the above discussion, a wide variety of investigations have been carried out to characterize the stress-strain response of pavements using finite element methods. Different constitutive models for asphalt layers ranging from linear elastic to viscoelastic to visco-plastic have been used. Granular layers have been modeled either as elastic-plastic or linear elastic. The influence of different loading conditions and tire configurations have also been investigated and compared with field data. However, one of the important issues missing in most of these studies is the quantification of the benefits of using modified asphalt in all the asphalt layers. While it is well known that the use of modified asphalt for the surface course can lead to substantial structural advantage, the benefits of using modified asphalt for all the asphalt layers has not been quantified. It will also be interesting to investigate the role of asphalt layer thicknesses on the final performance of the pavement system. This work reported here is aimed at finding answers to such questions.

## Experimental Investigations

### *Materials*

SBS polymer modified asphalt (PMB70) and viscosity grade 30 (VG30) asphalt confirming to Bureau of Indian Standards, India (IS:15462, 2004 and IS:73, 2006) were used. Aggregate gradations recommended by Indian specification (MoRT&H, 2004) for the surface and binder course layers were adopted and are shown in Table 1.

Table 1.Aggregate gradations

| Sieve size, mm | Bituminous concrete (BC)-Grade 2 | | Dense bituminous macadam (DBM)-Grade 2 | |
|---|---|---|---|---|
| | Specification limits, % passing | Adopted, % passing | Specification limits, % passing | Adopted, % passing |
| 37.5 | | | 100 | 100 |
| 26.5 | | | 90-100 | 95 |
| 19 | 100 | 100 | 71-95 | 83 |
| 13.2 | 79-100 | 89.5 | 56-80 | 68 |
| 9.5 | 70-88 | 79 | - | |
| 4.75 | 53-71 | 62 | 38-54 | 46 |
| 2.36 | 42-58 | 50 | 28-42 | 35 |
| 1.18 | 34-48 | 41 | - | |
| 0.6 | 26-38 | 32 | - | |
| 0.3 | 18-28 | 23 | 7-21 | 14 |
| 0.15 | 12-20 | 16 | - | |
| 0.075 | 4-10 | 7 | 2-8 | 5 |

### *Material characterization*

Bituminous concrete and dense bituminous macadam mixes used in the surface and binder course layers were prepared at constant asphalt content of 5.25 and 4.5 %

respectively. Dynamic modulus test on asphalt mixes with modified and unmodified asphalts was carried out as per AASHTO TP62, 2007 using simple performance test equipment. Controlled haversine compressive loading, with frequency varying from 25 to 0.01 Hz was applied. The test results reported here correspond to 60 $^{o}$C. All specimens were tested under unconfined condition within the linear viscoelastic range by keeping the strain in the range of 75 to 125 micro-strains. Variation in the storage and loss modulus of asphalt mixes with frequency for modified and unmodified asphalt is shown in Figure 1 and 2.

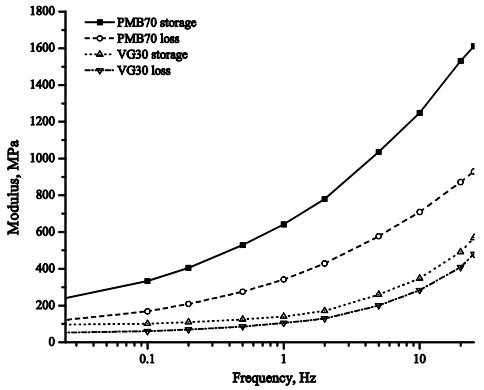



Figure 1. Storage and loss modulus variation with frequency for BC mixes

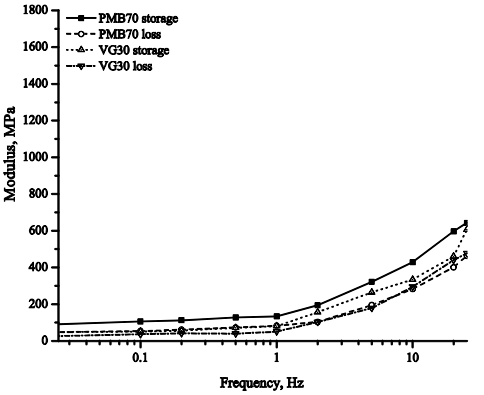



Figure 2. Storage and loss modulus variation with frequency for DBM mixes

## Flexible Pavement Composition

A five layer typical pavement structure (Figure 3) composing of asphalt and granular layer resting on a prepared subgrade of 10 % CBR for three different traffic levels as recommended by Indian Roads Congress Specification (IRC-37, 2001) is considered for the analysis. Table 2 summarizes the typical layer thickness corresponding to each traffic level.

Table 2.Flexible pavement composition for different traffic levels

| Type of layer | Thickness of each layer, mm | | | Elastic modulus, MPa | Poisson's ratio |
|---|---|---|---|---|---|
| | 50 MSA* | 100 MSA | 150 MSA | | |
| Bituminous concrete (BC) | 40 | 50 | 50 | var* | 0.35$^{+}$ |
| Dense bituminous macadam (DBM) | 110 | 130 | 150 | var | 0.35$^{+}$ |
| Granular base | 250 | 250 | 250 | 370$^{+}$ | 0.40$^{+}$ |
| Granular subbase | 200 | 200 | 200 | 170$^{+}$ | 0.40$^{+}$ |
| Prepared Subgrade | 500 | 500 | 500 | 77$^{+}$ | 0.40$^{+}$ |

*MSA= Million standard axles (traffic level), *var =Variable, $^{+}$ =as per IRC-37, 2001.

## Finite Element Analysis

### *Model*

Pavement structure shown in Figure 3 was modeled as two dimensional axisymmetrical using ABAQUS 6.8 finite element package. CAX4R (4-node, reduced-integration, axisymmetric, solid element) element types were used for

meshing. In order to reduce the CPU time, finite element mesh was refined in the region close to the axis of application of load. A total of 7497 nodes and 7300 element were formed for typical profile of 150 MSA.

Along the axis of symmetry, shear stresses and radial displacements were constrained to zero. Bottom of the subgrade and the other end was fully restrained by providing fixed end conditions. Interface conditions were assumed to be rough. A circular loading of radius 160 mm and 650 kPa contact pressure was applied. Dynamic haversine load with rest period, in order to simulate the moving load pattern was used in the present study. A loading time of 0.1 s with an increment of 0.01 s and 0.9 s rest period was repeated for 100 cycles.

Simulations were conducted in the Vega super-cluster at IIT Madras. This is a super computer with 256 nodes with each node consisting an Intel E5472, Quad-core dual processor with 16 GB RAM. Each simulation run for 100 cycles took nearly 21200 s of CPU runtime.

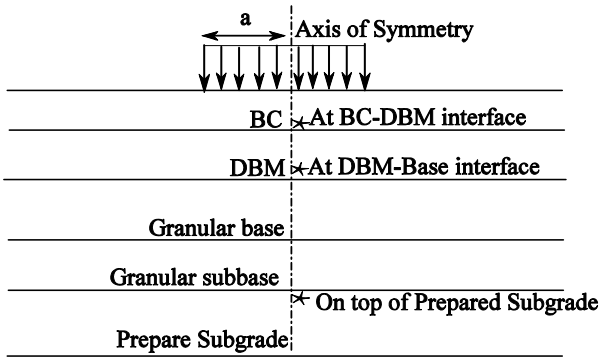



Figure 3. Pavement structure with response points and loading

***Material Models***

In this investigation, the granular materials were characterized as linear elastic materials and the material properties were taken from the Indian Roads Congress Specification (IRC-37, 2001). The asphalt layers were characterized as linear viscoelastic material. Asphalt mixes properties were modeled by assuming a linear viscoelastic Burgers' model. The constitutive relation for Burgers' model is given by:

$$\sigma + P_1 \dot{\sigma} + P_2 \ddot{\sigma} = q_1 \dot{\varepsilon} + q_2 \ddot{\varepsilon} \tag{1}$$

where, $\sigma = stress$, $\varepsilon = strain$, $P_1 = \frac{\eta_1}{R_1} + \frac{\eta_1}{R_2} + \frac{\eta_2}{R_2}$, $P_2 = \frac{\eta_1 \eta_2}{R_1 R_2}$, $q_1 = \eta_1$ and $q_2 = \frac{\eta_1 \eta_2}{R_2}$.

The viscoelastic material properties were obtained by fitting Burgers' model to the experimental data using non-linear least square regression analysis where the parameters were constrained to be positive. Time domain viscoelastic material model in ABAQUS was used in the present study. Hence, to do so stress relaxation data was generated using the calibrated Burgers' model parameters. The normalized stress relaxation data was fitted to the following relation:

$$E_r(t) = \sum_{i=1}^{N} g_i (1 - e^{-t/\tau_i}) \tag{2}$$

where, $g_i$ and $\tau_i$ are the material parameters and $E_r - \text{Relaxation modulus}$.

The expansion of this Prony series was used in ABAQUS to input time domain material parameters. In the present study, four parameter Prony series was used to model the viscoelastic response of asphalt mixes. Material parameters for different mixes at 60 $^{o}$C are tabulated in Table 3.

Table 3. Time domain viscoelastic material parameters

| Type of asphalt mix | Asphalt type | $g_1$ | $g_2$ | $\tau_1$ | $\tau_2$ |
|---|---|---|---|---|---|
| BC | VG30 | 0.783 | 0.217 | 0.051 | 2.804 |
| | PMB70 | 0.642 | 0.358 | 0.118 | 5.16 |
| DBM | VG30 | 0.990 | 0.010 | 0.082 | 13.437 |
| | PMB70 | 0.850 | 0.150 | 0.085 | 22.478 |

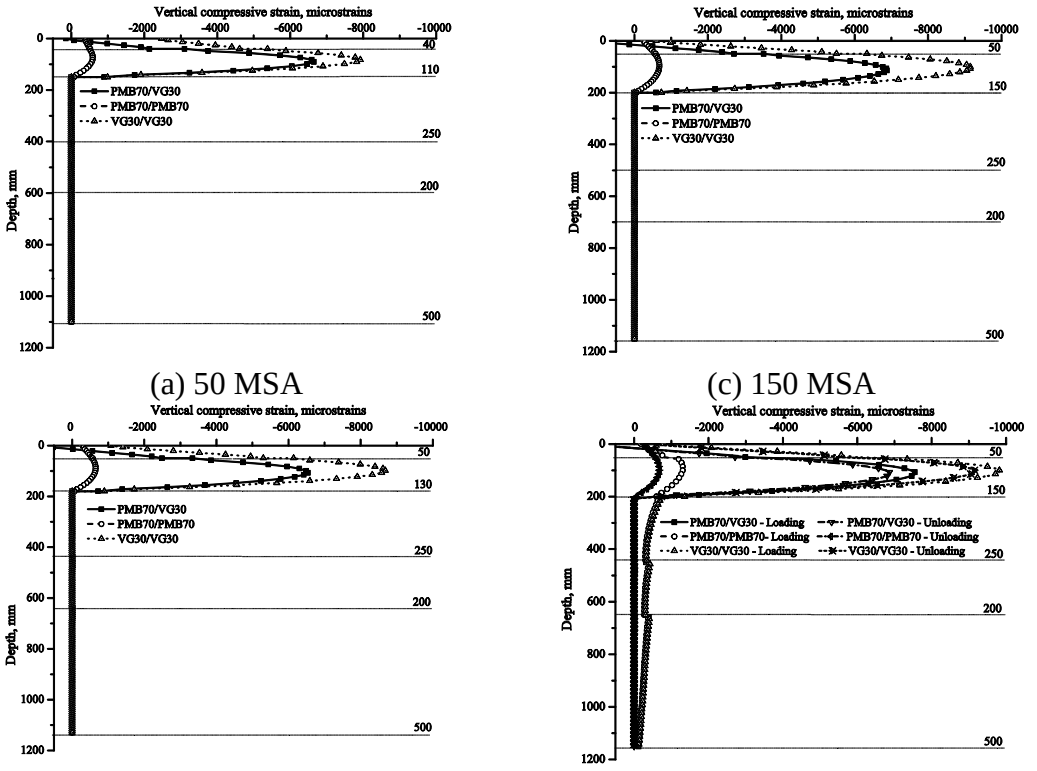
## Simulation Results and Discussions

### *Vertical compressive strain*

As discussed earlier, the simulation consisted of application of 0.1 s of haversine loading of 650 kPa peak magnitude followed by 0.9 s rest period. This constitutes one cycle and the pavement structure was subjected to 100 such cycles. Figure 4 shows the variation of vertical compressive strain along the depth of the pavement at the end of the application of 100 cycles. As it is seen, vertical compressive strain at the end of 100 cycles in the pavement structure with PMB70 in both asphalt layers is considerably less than that of pavement structure with VG30 and PMB70 in top layer alone. Also, the strain in granular and subgrade layers completely recover. Figure 4b and 4c shows the variation in strain for a higher traffic level, where the thickness of the asphalt layers alone was variable. It is very interesting to note that the vertical compressive strain along the pavement depth shows a maximum strain for pavement cross-sections having higher asphalt thickness. This variation depended on whether modified asphalts were used in the top two asphalt layers or not. For instance, there is no difference in the strain in the AC and DBM layers having polymer modified asphalts for 50, 100 and 150 MSA. There is a slight increase in strain as the total thickness of the asphalt base layers are increased (from 150 for 50 MSA to 180 for 100 MSA and then to 200 for 150 MSA). It is also observed that modified asphalt mixes experienced lower strains when compared to pavement structure with unmodified asphalt mixes. From this it could be ascertained that increase in the thickness of viscoelastic layer alone to imposed traffic loading demands might result in premature and early failures due to higher strain levels in the asphalt layers. These findings are similar to the results of Kim et al. (2009).

Response of pavement structure to loading and unloading can be noticed from Figure 4 (d) at 100$^{th}$ cycle (for 150 MSA). It is seen that during loading pavement structures with VG30 asphalt in both layers and PMB70 asphalt in top layer result in higher strain levels in asphalt layers when compared to pavement structure with PMB70 asphalt in both layer. It is interesting to notice that at depth of 200 mm (bottom of DBM layer) pavement structure with PMB70 in both layers recovers strain

by 89.70 %, whereas pavement structures with VG30 asphalt in both layers and PMB70 asphalt in top layer alone recovers by 45.96 and 50.67 % respectively. This essentially underlines the fact that asphalt mixtures with polymer modified binders exhibit a high viscoelastic solid-like behavior and hence are able to recover the strains much faster than mixtures with unmodified binders.




(a) 50 MSA (c) 150 MSA

(b) 100 MSA (d) 150 MSA

Figure 4. Variation of vertical compressive strain along the pavement cross-section

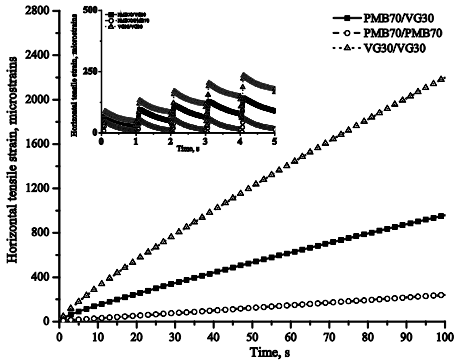
***Horizontal tensile strain***

Figures 5 (a, b, and c) shows the accumulation of tensile strain at the interface of BC and DBM layer for all traffic levels (50, 100 and 150 MSA) respectively. Marked difference in the accumulated horizontal tensile strain in the surface layer of pavement structures with modified asphalt mixes and unmodified asphalt mixes was observed. The inset figure shows the loading pattern and recovery with accumulation in each case considered in the present study. Figures 5 (d, e, and f) shows the accumulated horizontal tensile strain at the DBM-granular layer interface. It is seen from these figures there exists distinct difference in the accumulated horizontal tensile strain levels at the bottom of DBM layer (interface between asphalt and granular layers) in case of pavement structure with both layers with PMB70. The trend was same irrespective of the layer thickness considered in the present study.

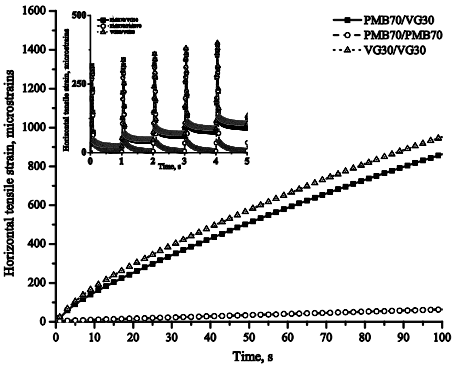
**Strain on subgrade**

Vertical compressive strain on subgrade, which is considered to be critical for rutting in the pavement structure, is shown in Figure 6 (a, b, and c) for all traffic levels. As can be seen the maximum strain experienced by the pavement structure on top of the subgrade with modified asphalt in both asphalt layers (PMB70/PMB70) was

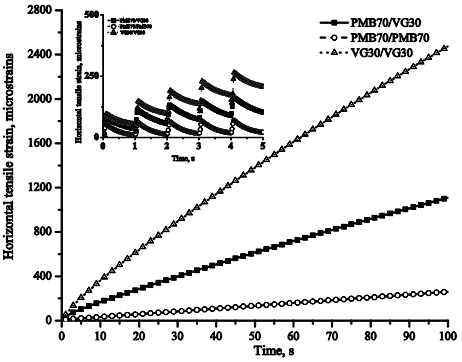
considerably lower when compared to pavement structure with unmodified asphalt and modified asphalt in top layer alone. Use of modified asphalt in both the asphalt layers in the pavement can result in substantial reduction of strain on top of subgrade. This might lead to reduction in rutting due to the component of rutting based on subgrade material property alone.



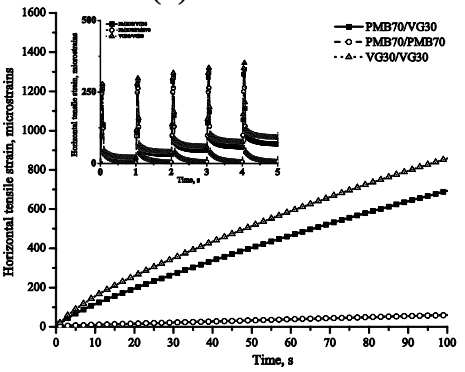

(a) 50 MSA



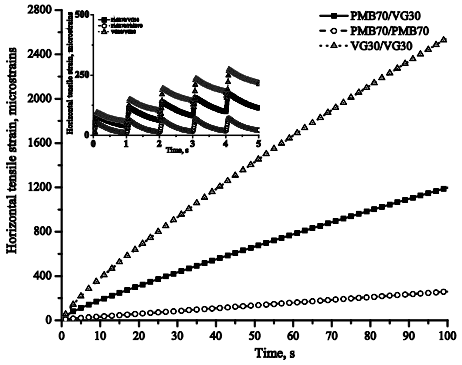

(d) 50 MSA



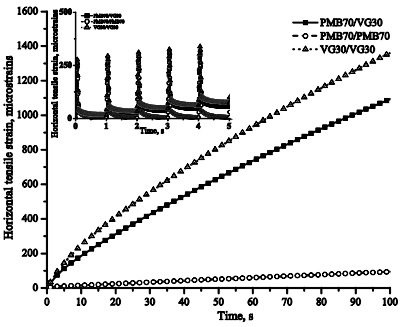

(b) 100 MSA




(e) 100 MSA




(c) 150 MSA




(f) 150 MSA

Figure 5. Accumulated tensile strain at the interface of BC-DBM and DBM-Base layers

**Quantification of Traffic Benefit Ratio**

Indian specification for guidelines and design of flexible pavements (IRC-37, 2001) relates the vertical compressive strain on the subgrade to number of cumulative standard axles to cause a rutting of 20 mm depth using the empirical relation as follows:

$$N_R = 4.1656 \times 10^{-8} [1/\varepsilon_Z]^{4.5337} \quad (3)$$

where, $\varepsilon_Z$ = vertical compressive strain on subgrade.

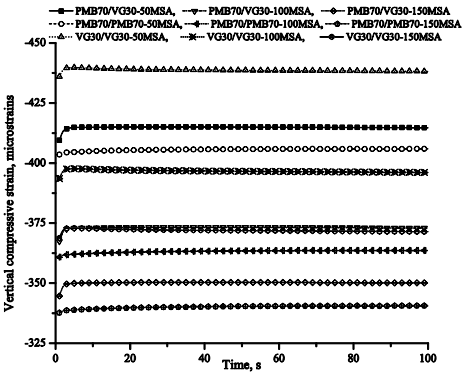



Figure 6. Variation of compressive strain variation on top of subgrade during loading

Table 4. Traffic benefit ratio of different pavement structures

| Type of layer | Traffic level, msa | TBR |
|---|---|---|
| PMB in BC alone | 50 | 1.29 |
| | 100 | 1.31 |
| | 150 | 1.32 |
| PMB in BC and DBM | 50 | 1.43 |
| | 100 | 1.48 |
| | 150 | 1.50 |

TBR (Perkins and Edens, 2002) is evaluated for pavement structures with and without modified asphalt mixes using 20.0 mm rutting as the failure criteria. Hence,

$$TBR = [\varepsilon_u / \varepsilon_m]^{4.5337} \quad (4)$$

where, $\varepsilon_u$ and $\varepsilon_m$ =vertical compressive strain on subgrade in pavement with VG30 and PMB asphalt layers respectively.

Using the relation given in equation 4, possible improvement in the life of flexible pavement (TBR) for different cases considered in the present study is listed in Table 4. Using modified asphalt in both the layers (surface and lower) benefits accruing to the pavement structure increase up to 1.5 times. From this it could be understood that using PMB70 in both layers, life of a pavement structure for a defined failure criteria (rutting in the present case) increases up to 50 %, which makes it more economical when compared to pavement structure with VG30 asphalt in both layers.

## Conclusions

In this investigation, the primary response of flexible pavements to dynamic loading was captured using two dimensional finite element model. Asphalt mixes were characterized using a linear viscoelastic model, granular and subgrade materials were modeled as elastic.

Classical elasticity based layer theory point out that increase of pavement layer thickness can substantially reduce the strains for any given material. For a viscoelastic layered structures, the layer thickness and strain gradient and its rate does not have a straight-forward relationship. It is possible that increasing the layer thickness up to some magnitude can actually lead to increase in the strain. It was seen in this investigation that the strain experienced in the top layer increased irrespective of the asphalt used for all the thickness. At the bottom of asphalt layers reduction in compressive and tensile strains can be achieved by increasing the thickness of DBM layer. Using PMB in both layers will result in considerable reduction of compressive and tensile strain at the bottom of asphalt layers. Also, substantial decrease in the compressive strain on the top of subgrade can be achieved by using PMB in both

asphalt layers. Benefits accrued (TBR) was found to be 1.3 to 1.5 times in case of pavement structures with PMB in top and both asphalt layers respectively.

**References**

AASHTO TP 62. (2007). "Standard Method of Test for Determining Dynamic Modulus of Hot- Mix Asphalt (HMA)." *American Association of State Highway and Transportation Officials*, Washington DC, USA.

Al-Hadidy, A. I., and Yi-qiu Tan. (2009) "Mechanistic analysis of ST and SBS-modified flexible pavements." *J Constr Build Mater*., (23), 2941-2950.

Al-Hadidy, A. I., and Yi-qiu Tan. (2009) "Mechanistic approach for polypropylene modified flexible pavements." *J Mat Des*., (30), 1133-1140.

Elseifi, M. A., Al-Qadi, I. L., and Yoo, P. J. (2006). "Viscoelastic modeling and field validation of flexible pavement." *J Engg Mech*., 132(2), 172-178.

Hadi, M. N. S., and Bodhinayake, B. C. (2003). "Non-linear finite element analysis of flexible pavements." *J Advances Eng Soft.*, 34, 657-662.

Helwany, S., Dyer, J., and Leidy, J. (1998). "Finite-element analysis of flexible pavements." *J Transport Engg*., 124(5), 491-499.

Hua, J., and White, T. (2002). "A study of nonlinear tire contact pressure effects on HMA rutting." *Int J of Geomech*, 2(3), 353-376.

IRC: 37 (2001). "Guidelines for the design of flexible pavements." *The Indian Roads Congress, Second revision,* New Delhi, India.

IS: 15462 (2004). "Polymer and Rubber Modified Bitumen-Specification." *Bureau of Indian Standards* (BIS), New Delhi, India.

IS: 73 (2006). "Paving Bitumen-Specification." *Bureau of Indian Standards* (BIS), New Delhi, India.

Kettil, P., Lenhof, B.,Runesson, K., Wiberg, N. E. (2007). "Simulation of inelastic deformation in road structures due to cyclic mechanical and thermal loads." *J Computers Struct.*, 85, 59-70.

Kim, J., Roque, R., and Byron, T. (2009). "Viscoelastic analysis of flexible pavements and its effects on top-down cracking." *J Materials in Civil Engg*, 21(7), 324-332.

Ministry of Road Transport and Highways, (MORT&H) 2001. "Specificationforroads and bridge work." Government of India, *Indian Roads Congress, 4th Revision*, New Delhi, India.

Mulungye, R. M., Owende, P.M.O., and Mellon, K. (2007). "Finite element modelling of flexible pavements on soft soil subgrades." *J Material Des*., (28), 739-756.

Novak, M., Birgisson B., and Roque, R. (2003). "Near-surface stress states in flexible pavements using measured radial tire contact stresses and ADINA." *J Computers Struc*., 81, 859-870.

Papagiannakis, A.T., Amoah, N., and Taha, R. (1996). "Formulation for viscoelastic response of pavements under moving dynamic loads." *J Transport Engg*., 122(2), 140-145.

Perkins, S. W., and Edens, M. Q. (2002). "Finite element and distress models for geosynthetic reinforced pavements." *Int. J. Pavement Eng.*, 3(4), 239-250.

White, T. D., Zaghloul, S. M., Anderton, G. L., and Smith, D. M. (1997). “Pavement analysis for moving aircraft load.” *J Transport Engg*., 123(6), 436-446.

# Characteristics of Mono-Strand Fiber-Reinforced Hot-Mix Asphalt Mixtures

Pyeong Jun Yoo[1], Ji Young Choi[2] and Byung-Sik Ohm[3]

[1]Highway Research Division, Korea Institute of Construction Technology, P.O. Box 411-712, Goyang, Gyeongi, Rep. of Korea; PH (82) 31-910-0175; FAX (82) 31-9100161; email: pjyoo@kict.re.kr
[2]Highway Research Division, Korea Institute of Construction Technology, P.O. Box 411-712, Goyang, Gyeongi, Rep. of Korea; PH (82) 31-910-0146; FAX (82) 31-9100161; email: legion@kict.re.kr
[3]Highway Research Division, Korea Institute of Construction Technology, P.O. Box 411-712, Goyang, Gyeongi, Rep. of Korea; PH (82) 31-910-0541; FAX (82) 31-9100161; email: bseom@kict.re.kr

## ABSTRACT

With premature deteriorations and performance limitations of hot-mix asphalt (HMA) pavements, innovative solutions are needed to enhance fatigue life or rut-resistance of HMA. Various reinforcing methods are widely utilized such as installing geogrid, geocomposite, or geomembrane layers at the bottom HMA to improve the structural integrity of HMA pavement. Such layered systems sometime improve the field performance of HMA; however, such solutions may not enhance toughness, tensile strength, or shear strength of HMA itself due to the location of interlayer systems. A new mono-strand fiber reinforcing material for HMA is proposed in this study to effectively reinforce the HMA not only at the bottom but also inner-side. Adding a small amount of fiber into HMA leads to significant increases in the phenomenological fatigue life of fiber-reinforced HMA compared to the conventional HMA, and four-point bending beam and indirect tensile tests are performed to observe the effect of fiber reinforcement.

**Key Words:** fiber, asphalt, fatigue

## INTRODUCTION

Various interlayers have been used to enhance the field performance of HMA pavement such as geogrid, geocomposite, or geomembrane; however, such solutions may not sometimes enhance the toughness or strength of HMA due to the location of interlayer, those are usually installed in between adjacent layers of a pavement system. In addition, difficulties may be arisen when installing interlayers including fixing and flattening, which may require additional labor cost. Interlayers are not a way to enhance the structural integrity of HMA itself but those are more focusing on reducing the vertical deflection or reflective damage from a slave layer to an upper layer in flexible or rigid pavement systems.

A mono-strand fiber reinforcing method for HMA is proposed in this study to overcome disadvantages in conventional reinforcing methodologies and to provide effective reinforcement of HMA itself in three-dimensional direction.

The strengthening of HMA is not an emerging technology, various chemicals are have been widely utilized to improve physical properties of HMA binder resulting in a tough HMA. Denning and Carswell (1981) showed that EVA (Ethylene Vinyl Acetate) in an asphalt binder offers advantages for HMA in that the stiffness of HMA increased at high temperatures and decreased at low temperatures. Many reports on polymer modifications of asphalt binder using LDPE (Low Density PolyEthylene), SBS (Styrene Butadiene Styrene), or SBR (Styrene Butadiene Rubber) show that a polymer modified binder improves binder ductility; hence, providing the modified binder is more durable to pavement stress and deformation due to temperature or traffic loading (Rowlett 1990, Little 1992, Brule 1996, Bonemazzi et al. 1996, Bahia et al. 2001 and Roque et al. 2004).

Although lots of pavement engineers are seeking economic ways to improve the tensile or shear strength of HMA, however, few research studies on experiment using mono-strand fibers with HMA have been found. Recently, some studies on reinforcing HMA using fibers showed that adding fibers such as nylon fibers, polypropylene fibers, or polyester micro fibers into HMA improved toughness, fracture energy, and the resistance to permanent deformation of HMA. Although some reports present that enhancing tensile or shear strength of fiber-treated HMA, increasing toughness is the representative advantage of fiber-treated HMA using micro fibers (Trottier and Banthia 1994, Lee et al. 2003, Bueno et al. 2003 and Kaloush et al. 2008).

Some difficulties in mixing and compaction of HMA mixture with micro fibers may be arisen, and also fiber's balling in HMA is an issue to solve when mixing. This means that good dispersion of fibers in HMA should be verified before paving fiber-treated HMA. To have relatively good dispersion and strengthening effect of filament-type fiber-treated HMA compared to micro fibrillate-type fiber-treated HMA, this study developed a mono-strand fiber (filament-type) as a reinforcing media for HMA.

To address improvement in resistance to fatigue damage and to see creep compliance characteristics of mono-strand fiber-reinforced HMA, four-point bending beam tests and 1,000 sec indirect creep compliance tests were performed in this study. Noticeable fiber reinforcing effects on reducing fatigue damages and compliance

behaviors of fiber-treated HMA in three temperatures (4, 25, and 40°C) were presented. Such reinforcement may result in greater stability and longer fatigue life of HMA compared to those of conventional HMA.

## MONO-STRAND FIBER-REINFORCED HMA

Heavier traffic loading and higher tire pressure have recently resulted in severe damages in HMA pavement. Although damages of HMA may be related to asphalt binder's quality, aggregate type and gradation, or mix design, some limitation in asphalt binder's performance may be one of cause of failure of HMA (Al-Qadi 2005).

Asphalt binder is a complex material that many molecules associating in microstructures through intermolecular bonds. Due to its chemical complexity, asphalt binder specifications have been developed through physical property tests representing the performance of a binder such as softening point, penetration test, ductility, and viscosity (Huang 1993, SHRP 1994, Elseifi 1999).

Viscosity specification, penetration specification, and more recently the performance grade specification by SHRP are individually defining specific limitations whether an asphalt binder meets their own criteria as a useful paving material (Huang 1993 and SHRP 1994).

Beside the physical characteristics, the mechanical behavior of asphalt binder can be treated as a function of temperature, loading time, and stress on a nonlinear horizon. To describe the mechanical characteristic of asphalt binder, various mechanical analogues are widely described as visco-elasticity, visco-plasticity, or visco-elasto-plasticity. Mechanical assumptions comprise their own constitutive relationships between stress and strain or reversely. To enhance mechanical properties like dynamic shear modulus and tensile or shear strength more than those of plain asphalt binder, an asphalt binder is usually modified using polymers like SBS or SBR (Christensen and Anderson 1992, Bonemazzi et al. 1996, Brule 1996 and Elseifi 1999).

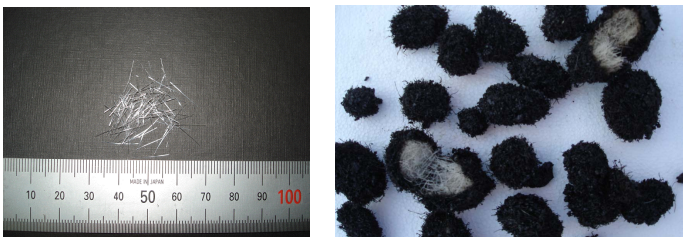
One promising method for limiting HMA failures or enhancing physical properties is using polymer modified asphalt binder in a flexible pavement exposed to severe climatic condition and heavy-weight truck traffic. Modified binder is expecting that the stiffness of HMA increases at high temperatures and decreases at low temperature (SHRP 1994 and Bruel 1996). Although several advantages of polymer modified HMA have been reported, the polymer modified binder and mixture have not yet been mechanically characterized because of the complex chemical and physical interaction between asphalt binder and polymer (Elseifi 1999).

Causing the complexity in characterizing the polymer modified binder and mixture, physical parameters like $G^*/\sin\delta$ or $G^*\sin\delta$ are widely utilized to control the performance criteria such as rutting and fatigue cracking instead of doing micro mechanical or chemical analyses (SHRP 1994).

Producing the polymer modified asphalt binder using SBS or SBR requires relatively high temperature and energy than those for plain asphalt binders. In addition, mixing and compaction temperatures for the modified HMA at a plant or in a field should maintain 30-40 °C higher than those for the plain HMA to verify specified viscosity criteria (Bonemazzi et al. 1996 and Brule 1996).

To compensate those drawbacks in the polymer modified HMA, a new mono-strand fiber-reinforced HMA proposed in this study. Adding a small amount of mono-strand fiber in plain HMA neither changes in conventional mix design processes, nor needs additional equipments to mix fibers with HMA. Shapes of the fibers developed in this study are like Figures 1.

(a) Mono-strand fibers and dispersion in a mixture



(b) Thin fibrillated mono-strand fibers and fiber's balling

**Figure 1. Characteristics of mono-strand and thin fibrillated fibers**

The dimension of mono-strand fiber, Figure 1(a), 30 mm (L)×1-2 mm (W)×1 mm (T), is decided relatively bigger than that of the fibrillated fiber, Figure 1(b), to make mono-strand fibers free from the fiber's balling at high temperature above 150 °C and to have more uniform fiber distribution in HMA than that of the thin fibrillated mono-strand fiber (Figure 1(b)).

The mono-strand fiber, Figure 1(a), has several dents on the surface of each fiber and it also has a longitudinal groove to separate itself while doing dry-mixing process with aggregates in a plant. Authors expected that enhancing frictional force on the interface between aggregates and fibers through several embossments on the surface of fiber. Enlarging the contact area of fiber with asphalt binder through additional separation of each fiber due to longitudinal grooves may expect that enhancing bond strength between fiber and asphalt binder, although further micro scale experiments should be done that would verify those effects in a future study.

Mixtures from an asphalt plant show, Figure 1(a), that good dispersion of fibers, even though any additional mixing process does not be provided except the dry-mixing with aggregate for 20-30 seconds. The surface of a cored sample also shows that randomly dispersed-fibers in yellow dots without any noticeable fiber's balling, Figure 1(a).

However, chemical reactions between the asphalt binder and the fiber's surface are hardly characterized at this point because of the chemical complexity between them. Hence, the mono-strand fiber-reinforced HMA has been tested through physical property tests such as four-point bending beam and indirect tensile tests. In

addition, 1,000 seconds creep compliance test was performed to characterize the relative creep compliance behavior between the plain and the mono-strand fiber-reinforced HMA mixture.

## MIXTURE CHARACTERIZATION

Characterization of phenomenological behaviors of the fiber-reinforced HMA is the objective of this study. Several laboratory tests were performed to investigate relative characteristics of fiber-reinforced HMA compared to the plain HMA. The asphalt binder utilized in this study was PG 64-22 (AP5). HMA mixtures both with fibers and without fibers were mixed in a laboratory according to mixture characteristics like Table 1 and the aggregate gradations in coarse-graded like Table 2.

**Table 1. Mixture Characteristics**

| Mix type | Binder Mix Design Data | | | |
|---|---|---|---|---|
| | Binder Type | Design AC (%) | Target Va (%) | $G_{mm}$ |
| Control | PG64-22 | 5.5 | 7 | 2.474 |
| Polyphalt a* | PG64-22 | 5.5 | 7 | 2.474 |

***Polyphalt a: Mono strand fiber-reinforced HMA with the fiber content of 0.4% in mixture weight.**

**Table 2. Aggregate Gradation**

| Aggregate Gradation | Percent Passing (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1" | 0.75" | 0.5" | 0.375" | No.4 | No.8 | No. 30 | No.200 |
| | 0.0 | 94.5 | 77.3 | 59.7 | 31.9 | 24.1 | 12.3 | 3.0 |

The surfaces of mono-strand fiber-reinforced HMA mixed in the laboratory are showing like Figure 2 for preparation of the Rice $G_{mm}$ and gyratory compaction.

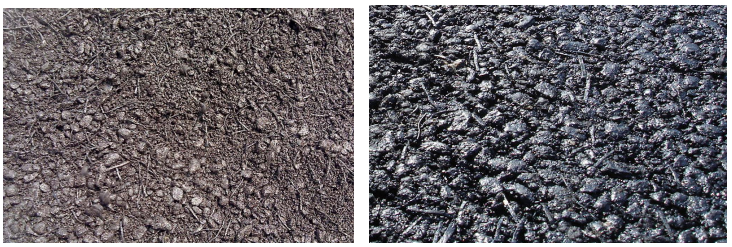


**Figure 2. Close-up of the fiber-reinforced HMA**

All specimens were compacted with a gyratory compactor of 150 mm diameter mold to approximately 122-123 mm in height. The mass of each specimen was calculated according to target $V_a$ like 7%, volume of mold, and theoretical specific gravity of $G_{mm}$. Two test specimens approximately 60 mm in height were cored and cut from the each compacted specimen for indirect tensile tests.

### Indirect Tensile Tests

As for the fiber-reinforced HMA, although an available mix-design regardless of fibers has not been developed within this research, HMA considering inclusion of

fibers may expect better contribution to enhance the fatigue life of HMA. To show phenomenological characteristics upon the fiber modification of HMA in this study, indirect tensile creep compliance and cyclic loading fatigue tests were performed and enabled to observe distinguishable properties such as creep compliance and indirect fatigue life relatively between with fibers and without fibers in HMA.

In this study, characterization of the viscoelastic properties for HMA mixtures was performed using the indirect creep compliance test. This test subjects to a constant compressive creep loading for 1,000 seconds at 4, 25, and 40 °C, Table 3 (Al-Qadi et al. 2005).

**Table 3. Creep Loading**

| Specimen Size | Temperatures, °C | Loading Rate, N/sec | Load, N |
|---|---|---|---|
| **101.6 mm × 63.5 mm** | 4 | 75 | 1800 |
| | 25 | 50 | 150 |
| | 40 | 10 | 30 |

***Indirect Creep Compliance Test***

The theoretical creep compliance in uniaxial assumes that stresses are applied suddenly and held constant $\sigma_0$, while strains are recorded. The strain in each Kelvin unit, a spring is connected to a dashpot in parallel, is governed by:

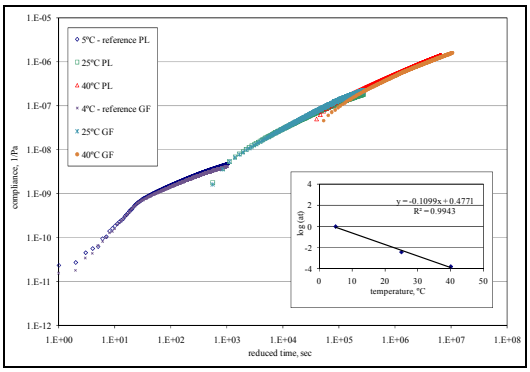
$$\sigma = E_i\varepsilon_i + \eta_i\dot{\varepsilon}_i \qquad i = 1,...,n \tag{1}$$

The initial condition, $\varepsilon_i(0) = 0$, would be applied then adding each strains for all components results in the creep compliance for the generalized Kelvin solid as the following Equation (2):

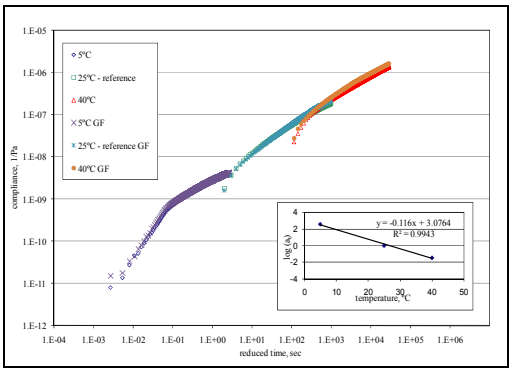
$$D(t) = \frac{\varepsilon}{\sigma_0} = D_0 + \sum_{i=1}^{n} D_i(1 - e^{-tE_i/\eta_i}) \tag{2}$$

where $D_0 = 1/E_0$, $D_i = 1/E_i$, and $\tau_i = \eta_i / E_i$., $\tau_i$ is the retardation time for element i. On the other hand, the tensile creep compliance of HMA for a biaxial stress state that exists on a specimen face engages the Hook's law first and the well-known function of D(t) is corrected to account for three-dimensional effects by Roque and Buttlar (Roque and Buttlar 1992, Buttlar et al. 1998).

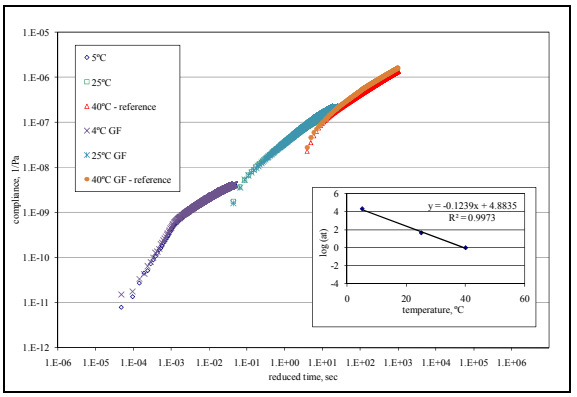
The creep compliance in this study is also utilized the procedure defined by Roque and Buttlar (1992). The master curve and shift factor are utilized to obtain viscoelastic parameters of HMA with creep compliance data. Master creep compliance curves could be fit to a power model or a generalized Kelvin model, which is defined in Equation (2). Creep compliances from laboratory tested temperatures at 4, 25, and 40°C are shown as Figures 3(a), (b), and (C). Log creep compliance vs. log time curves could be fitted second order polynomial at each temperature, and shifting the data to obtain temperature shift factors. Master creep compliances vs. reduced time curves are as follows:




(a) 4 °C Reference temperature




(b) 25 °C Reference temperature



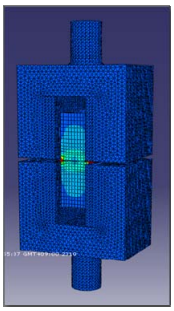

(c) 40 °C Reference temperature

**Figure 3. Comparative creep compliances**

With creep compliances between the fiber-reinforced HMA and the plain HMA compacted at the same target air void of 7%, master curves at each reference temperatures are hardly differentiated between them like Figures 3(a), (b), and (c). Although the 0.4% of fibers in mixture's weight, 40 °C GF-reference in the legend of Figure 3(c), shows slightly higher compliances at high temperature such as 40 °C, however, it may not be meaningful differences in the compliance behavior. This means that creep compliances from indirect mode tests are not proper measures to distinguish the fiber's toughening effect compared to the plain HMA. However, creep compliances would be utilized to define time and temperature dependent modulus relaxation parameters by the generalized Maxwell model after doing a conversion process, which is a viscoelastic analogy that is a simple Maxwell element is consisted of a spring and a dashpot in series such as the Prony series (Elseifi et al. 2006).

In other words, indirect tensile creep compliance test is hard to distinguish the enhancement of toughness through fiber reinforcement because it is not a failure test and creep compliance represents the localized creep behavior in the mid-part of a specimen. Therefore, this study proposed a new fixture for direct tensile creep test to make sure that compliance behavior occurs in a whole horizontal plane of the specimen in the middle like the Figure 4(a). Besides direct tensile creep test, this set-up can be used for cyclic fatigue test under completely reversed-sinusoidal

loading pulses. Direct tensile test would be performed using a rectangular specimen in tension only or tension and compression loading mode, which is in parallel to the vertical axis of the specimen. The horizontal and potential failure plane in the middle of the specimen is simulated by finite element analysis to observe the most singular stress concentration area but not yet be formulated mathematically, such that measurements may be made directly on this horizontal plane.



(a) FE Simulation

(b) Test Fixture

(C) Compaction mold

**Figure 4. Direct tensile test fixtures**

After applying the cyclic sinusoidal loading amplitude in 10 Hz for 1 second using the ABAQUS, the finite element simulation shows that the critical deformation or failure may be induced in the middle of specimen where stress concentration is the highest like Figure 4(a). The direct tensile test fixture like Figure 4(b) would be set in the Material Testing System. In addition, to make sure that the specimen for this test maintained coherent compaction energy and size, the compaction mold as Figure 4(c) was also devised.

The indirect creep compliance with cylindrical specimen is limited to verify the toughening effect of fibers in HMA, authors are expecting that the direct creep compliance test of fiber-reinforced HMA with the new fixture may result in lower creep compliances compared to the plain HMA at high temperatures above 40 °C distinguishably. Conversely, the initial relaxation modulus of the fiber-reinforced mixture may result in higher relaxation modulus than that of the plain HMA. This means that lower creep compliance and higher relaxation modulus of fiber-reinforced HMA at relatively high temperature may provide better resistant to shear flow or fatigue cracking.

**Indirect Cyclic Fatigue and Flexural Beam Fatigue Tests**

The indirect tensile strength test (IDT) of HMA is widely utilized to evaluate the tensile strength of a cylindrical specimen. Besides the indirect tensile strength, to see the distinguishable indirect toughening effect between the fiber-reinforced specimen and the plain HMA, the cyclic haversine loading pulse at 10 Hz of 150 kg was applied to the indirect cylindrical specimen (AASHTO 2007).

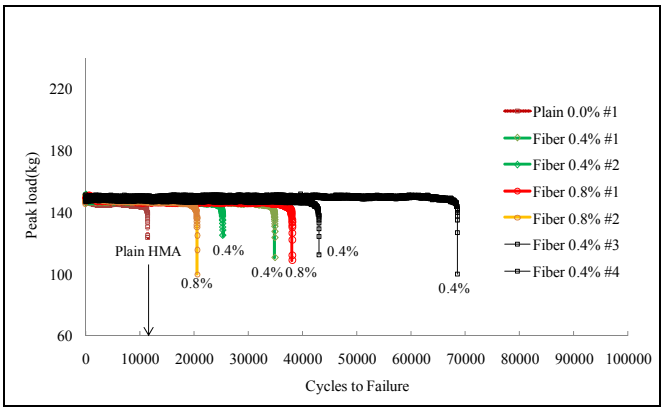



**Figure 5. Indirect cyclic tensile fatigue test results**

As shown in Figure 5, while the plain HMA without fibers failed after about 10,000 loading cycles, the 0.4% or 0.8% of fiber-reinforced HMA in specimen's weight maintained the maximum loading cycle up to about 2 to 7 times higher loading repetitions than the plain HMA did. Extending the loading cycle to the fatigue failure would be caused by the toughening effect of fibers in HMA, even though high variations in the maximum loading cycles among fiber-reinforced specimens exist due to the random distribution of fibers mixed in HMA contributing to bridging effect.

In addition, to propose a more explainable phenomenological fatigue behavior of the HMA covering the whole range of loading cycles up to the fatigue failure, alternative tests based on the four-point bending fatigue test (4PB) were performed. Results from the 4PB were utilized for fatigue characterization using stresses, strains, stiffness, and dissipated energy calculated by the common procedures defined in AASHTO T-321. The dissipated energy per unit volume per loading cycle or period was derived from the total potential energy per loading cycle like Equation (3) (AASHTO T-321, 2008).

$$\therefore\ W_i = Dissipated\quad Energy\ \ (DE) = \pi\hat{\varepsilon}^2 E''(w) = \pi\hat{\varepsilon}^2 \left|E^*\right| sin\ \delta = \pi\hat{\varepsilon}_i\sigma_i\ sin\ \delta_i \tag{3}$$

where,

$W_i$ = dissipated energy at load cycle i,

$\sigma_i$ = stress amplitude at load cycle i,

$DE$ = dissipated energy per cycle,

$\hat{\varepsilon}_i$ = strain amplitude at load cycle i, and

$\delta_i$ = phase angle between stress and strain wave signals.

Dissipated energy per loading cycle is computed using Equation (3). Summation of the dissipated energy per loading cycle results in the cumulative dissipated energy over the fatigue life. The ratio of dissipated energy change (RDEC) per 100 loading cycle-interval could be calculated by Equation (4) (Shen and Carpenter 2005):

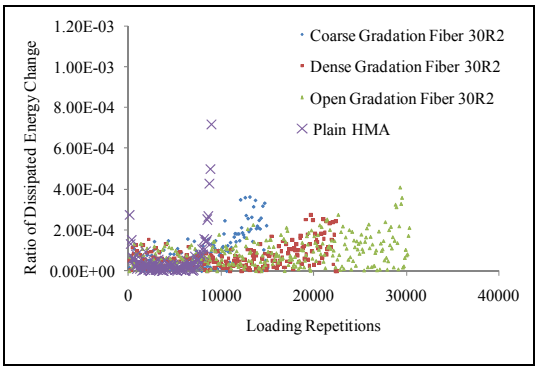
$$RDEC_a = \left|\frac{DE_a - DE_b}{DE_a \times (b-a)}\right| \tag{4}$$

where,
*RDEC*= the ratio of dissipated energy change;
*a, b* = loading cycles, 100 cyclic interval; and
*DE* = dissipated energy during a loading cycle a and b, respectively.

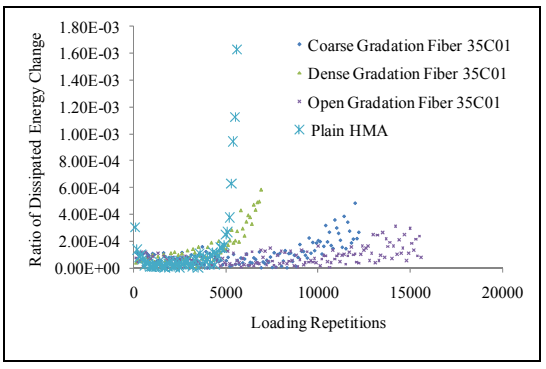
The 4PB specimen is repeatedly loaded according to the displacement-controlled mode until the fatigue failure occurs or up to the end of the tertiary stage, as used in this study. The test fixture and the two different fiber-reinforced specimens after fatigue failure are like Figures 6 (a) and (b).

(a) (b)

**Figure 6. (a) Four-point bending beam fixture and (b) Fiber-reinforced specimen after failure**




(a) Peak-to-peak strain level 1,000 με at 25°C




(b) Peak-to-peak strain level 1,500 με at 25°C

**Figure 7. Four points bending beam fatigue test results**

Although fatigue damages may be related to asphalt binder's quality, aggregate type, gradation, or mix design, all specimens in Figures 7(a) and (b) are made in same

conditions such as binder, aggregate gradation, mixing, and compaction except the with fiber's contents (0.4%) in weight of specimen.

The RDECs for the plain and fiber-reinforced HMA are plotted versus loading cycles in Figures 7(a) and (b) to distinguish the mixture's fatigue behavior in different loading amplitudes. Figures 7(a) and (b) represent that initial rapid changes of RDEC during the primary stage of loading cycles until that the initial densification is completed. Following the primary stage, the curve remains relatively constant level of RDEC indicating stable fatigue damage resistance as a secondary stage of fatigue. Upon failure, the RDEC increases again, and a much greater portion of the dissipated energy capacity of the specimen may be converted to the fatigue damage representing the tertiary stage as shown in Figures 7 (a) and (b).

While the tertiary stage after the failure starts around 8,000 loading repetitions for the plain dense-graded HMA specimen under 1,000 με strain amplitude, Figure 7 (a), all the fiber-reinforced specimens in dense, coarse, and open aggregate gradations maintain the secondary stage up to the 1.5 to 3 times higher loading repetitions than those of the plain HMA.

In case of 1,500 με loading amplitude, Figure 7 (b), the loading repetition at the tertiary stage, where the fatigue failure starts, for the fiber-reinforced HMA is about maximum 3 times higher than that of the plain HMA. The toughening effect of fiber-reinforced HMA is primary cause of extending the fatigue life up to the fatigue failure, even though more tests at different temperatures and mixture types should be done in a following research.

## CONCLUSION

To enhance the stiffness of asphalt binder, polymer modification is widely utilized in HMA pavement. Instead of the chemical modification of an asphalt binder, interlayers are often installed in between two different layers in a pavement system as a stress absorbing layer. However, such solutions may not sometimes reduce the critical strain within HMA due to the location of interlayer system and the modification of asphalt binder may result in too stiff binder so that fatigue damage potential may highly be susceptible to temperature change even compared to the conventional HMA.

To overcome those disadvantages in the conventional reinforcing methods and to reinforce HMA itself in three-dimensional direction, the mono-strand fiber reinforcing method was proposed in this study. The dimension of mono-strand fiber developed in this study was 30 mm (L)×1-2 mm (W)×1 mm (T), this relative macro dimension compared to the conventional micro filament fiber made it possible that the fiber-reinforced mixture free from the fiber's balling when mixing fibers with HMA at high temperature above 150 °C and no need additional process or equipment to mix fibers in an asphalt plant except throwing a fiber-bag into a mixer. Good dispersion of fibers in HMA without any noticeable fiber's balling was observed through the uncompacted mixtures got from an asphalt plant.

To distinguish creep compliance characteristics between the fiber-reinforced HMA and the plain HMA, 1,000 second creep compliances tests were performed; however, creep compliances were hardly distinguishable between them. Because

indirect creep compliance tests measure small deflections in the middle of a specimen and are not failure tests so that it was hardly to distinguish physical differences between them. Therefore, a new direct tensile test fixture which can apply the tensile creep loading directly to a specimen was proposed to verify the stiffening effect of fiber in HMA for a future study.

In addition, to see the distinguishable indirect toughening effect between the fiber-reinforced specimen and the plain HMA, indirect cyclic fatigue tests were performed. The 0.4% or 0.8% of fiber-reinforced HMA maintained the maximum indirect loading cycle until about 2 to 7 times higher than the plain HMA did. Extending the loading cycle to fatigue failure would be caused by the toughening effect of fiber in HMA.

To propose a more explainable phenomenological fatigue behavior of HMA considering the bridging or toughening effect of fibers, four-point bending fatigue tests were performed. Fiber-reinforced specimens maintained 1.5 to 3 times higher loading repetitions than those of the plain HMA in all cases of 1,000 με and 1,500 με amplitudes. The toughening effect of fibers in HMA may be the cause of extending fatigue life.

In conclusion, adding a small amount of fiber into HMA leads to significant increases in the phenomenological fatigue life compared to the conventional HMA, although further experiments are needed to verify those effects more extensively.

## REFERENCES

AASHTO. (2007). *Standard test method for determining the creep compliance and strength of hot-mix asphalt (HMA) using the indirect tensile test device*, AASHTO T 322-07, Washington D.C.

AASHTO. (2008). *Standard test method for determining the fatigue life of compacted hot-mix asphalt (HMA) subjected to repeated flexural bending*, AASHTO T 321-08, Washington D.C.

Al-Qadi, I. L. et al., (2005). *Fatigue life characterization of superpave mixtures at the Virginia smart road.* Final Contract Report No. FHWA/VTRC 06-CR1, Charlottesville, VA.

Bahia, H. U., Hanson, D. I., Zeng, M., and Anderson, R. M. (2001). *Characterization of modified asphalt binders in superpave mix design*, NCHRP 459, TRB, National Research Council, Washington D.C.

Bonemazzi, F., Braga, V., Corrieri, R., Giavarini, C. and Sartori, F. (1996). "Characteristics of polymers and polymer-Modified Binders." *Transportation Research Record 1535*, TRB, National Research Council, Washington, D. C.

Brule, B. (1996). "Polymer-modified asphalt cements used in the road construction industry: basic principles," *Transportation Research Record 1535*, TRB, National Research Council, Washington, D. C.

Bueno, B. S., Silva, W. R., Lima, D. C., and Minete, E. (2003). "Engineering properties of fiber reinforced cold asphalt mixes." Technical note, *Journal of Environmental Engineering*, ASCE, Vol. 129, No.10.

Buttlar, W. G., Roque, R., and Reid, B. (1998). "An automated procedure for generation of the creep compliance master curve for asphalt mixtures."

*Transportation Research Record 1630*, TRB, National Research Council, Washington, D.C., 28-36.

Christensen, D. W., and Anderson, D. A. (1992). "Interpretation of dynamic mechanical test data for paving grade asphalt cements." *Proceedings of the Association of Asphalt Paving Technologists*, volume 61, pp. 67-116.

Denning, J. H., and Carswell, J. (1981). "Improvements in rolled asphalt surfacings by the addition of organic polymers." *TRRL 989, Transport and Road Research Laboratory*, Crowthorne, England.

Elseifi , M. A. (1999). *Viscoelastic modeling of straight and modified Binders at intermediate and high temperatures*, Master Thesis, Virginia Tech, Balcksburg, VA.

Elseifi, M., Al-Qadi, I. L., and Yoo, P. J. (2006). "Viscoelastic modeling and field validation of flexible pavements." *Journal of Engineering Mechanics*, ASCE, Vol. 132, No. 2, 172–178.

Huang, Y.H. (1993). *Pavement analysis and design*, Prentice Hall, NJ, USA.

Kaloush, K. E., Biligiri, K. P., and Zeiada, W. A. (2008). *Evaluation of FORTA fiber-reinforced asphalt mixtures using advanced material characterization tests-Evergreen drive, Tampa, Arizona*, Research Report, FORTA Corporation, Grove city, PA.

Lee, S. J., Rust, J. P., Hamouda, H., Kim, Y. R., and Bordon, R. H. (2005). "Fatigue cracking resistance of fiber-reinforced asphalt concrete." *Textile Research Journal*, Vol. 75, No.2, 123-128.

Little, D. N. (1992). "Analysis of the influence of low density polyethylene modification (Novophalt) of asphalt concrete on mixture shear strength and creep deformation." *Polymer Modified Asphalt Binders*, STP 1108, ASTM, Philadelphia, PA, 186-202.

Roque, R., and Buttlar, W. G. (1992). "Development of a measurement and analysis system to accurately determine asphalt concrete properties using the indirect tensile test." *Journal of the Association of Asphalt Paving Technologists*, Vol. 61, 304-332.

Roque, R., Birgisson, B., Tia, M., Kim, B., and Cui, Z. (2004). *Guidelines for the use of modifiers in Superpave mixtures: Executive summary and volume 1 of 3 volumes: Evaluation of SBS modifier*, State Job 99052793. Florida Department of Transportation, Tallahassee, FL.

Rowlett, R.D. (1990). *Performance of Asphalt Modifiers: Classification of Modifiers and Litterature Review*. Center for Construction Materials Technology, SWL, Houston, TX.

Shen, S., and Carpenter, S. H.(2005). "Application of dissipated energy concept in fatigue endurance limit testing." *Transportation Research Record 1929*, TRB, National Research Council, Washington, D.C., 165-173.

Strategic Highway Research Program. (1994). *Binder characterization and evaluation volume 1,* SHRP-A-367, National Research Council, Washington, D.C.

Trottier, J. F., and Banthia, N. (1994). "Toughness characterization of fiber reinforced concrete." *Journal of materials in civil engineering*, vol.6, No. 2, 264-289.

# Evaluation of Guss Asphalt Applied to Steel Deck Surfacing

Jian-Shiuh Chen[1], Min-Chih Liao[2] and Chien-Chung Huang[3]

[1]Professor, Department of Civil Engineering, National Cheng Kung University, Tainan, Taiwan 701; jishchen@mail.ncku.edu.tw
[2] Post-Doctor Fellow, Department of Civil Engineering, National Cheng Kung University, Tainan, Taiwan 701; n68993242@mail.ncku.edu.tw
[3] Ph.D. candidate, Department of Civil Engineering, National Cheng Kung University, Tainan, Taiwan 701; n68951234@mail.ncku.edu.tw

**ABSTRACT:** Construction of the new Guo Shien Bridge (GSB) included a steel orthotropic bridge deck. The higher flexibility of an orthotropic deck could cause a pavement placed upon it to fatigue and crack more quickly than a pavement placed on a normal roadway. The Guss asphalt mixture was applied to the steel deck pavement for the GSB. When heated to high temperatures, the Guss mixture is designed to be pourable asphalt mastic to provide an impermeable and durable layer for the steel deck surfacing. The basis of this paper was to evaluate the mixing and placing characteristics of the Guss asphalt mixtures. Various tests including Lueer fluidity, indentation, bending flexibility and wheel tracking were conducted to analyze the engineering properties of the Guss asphalt mixture for durability. Practical guidelines were given to enhance the quality of Guss asphalt mixtures. This paper was to provide background information on orthotropic bridge deck construction practices and document the construction of the Guss asphalt mixtures on the GSB.

## INTRODUCTION

The opening of the new Guo Shien Bridge (GSB) in February of 2009 marked the use of two transportation technologies new to the highway industry in Taiwan. The new bridge is the highest steel orthotropic deck (72 m high) in Taiwan in place of a traditional concrete deck. Large bridges in other parts of the world routinely incorporate orthotropic decks to reduce weight and lower cost. The drawback is the higher flexibility of the steel deck causes pavement placed upon it to fatigue and crack more quickly than pavement placed on a normal roadway. This leads to the second new technology, Trinidad Lake Asphalt (TLA), which is added to the hot mix asphalt (HMA) as Guss asphalt on the new bridge. TLA is a naturally occurring asphalt binder used to increase the durability and stability necessary for a pavement to withstand the stresses on an orthotropic deck. This study is part of a project to evaluate the engineering properties of Guss asphalt used on the GSB.

The word orthotropic is derived from the words orthogonal anisotropic, meaning different elastic properties in perpendicular directions. An orthotropic bridge deck is one in which a steel deck plate is supported by longitudinal ribs and transverse crossbeams. The different geometries of the ribs and crossbeams give the steel deck different flexural stiffness in the transverse and longitudinal directions making it orthotropic. Fig. 1 is a schematic of a section of an orthotropic bridge deck on the GSB. The primary advantage of an orthotropic bridge deck is that it is lighter than a traditional concrete deck which reduces the total dead load carried by the rest of the structure. The reduced dead load allows the towers, cables, and other supporting members to be smaller reducing overall cost. At the same time, it is efficient on long-span structures. The reduction in dead load comes at a price. Deflections due to traffic loading in the relatively thin steel deck plate are much greater than on a concrete deck. The greater deflections translate into higher strains in the pavement which leads to reduced life due to fatigue in traditional overlays (Arnaud and Houel 2007).

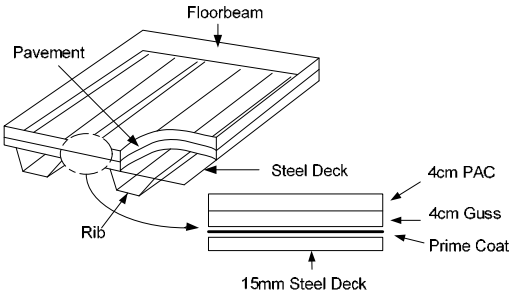



**Figure 1. Schematic of GSB orthotropic bridge deck.**

Guss asphalt has been successfully used on many orthotropic bridge decks in Europe and Japan. The characteristics of the Guss asphalt mixture are as follows: (1) It has an excellent water-proof property due to its impermeability; (2) It has an excellent resistance to the impact or shock; and (3) There is no need for compaction after the laydown. A Guss asphalt mixture is impermeable to moisture and provides a good bond with other layers. It is expected to be durable and provides a long service life on orthotropic decks (Hulsey et al. 2009; Hicks et al. 2005). On the GSB, 4-cm Guss asphalt was paved on the top of a 15-mm steel deck, followed by a 4-cm porous asphalt concrete (PAC). The prime coat is to bind the pavement system to the steel deck, and provides a good bond and protects the steel against corrosion. This pavement system needs to be designed to be flexible in order to resist fatigue cracking due to the higher strains inherent with an orthotropic deck, and yet strong enough to combat heavy traffic loading. This paper was to evaluate the engineering properties of the Guss asphalt mixture, and give highway engineers an insight into the mixing and placing characteristics of the steel deck surfacing.

## MIX DESING OF GUSS ASPHALT MIXTURE

**Binder properties.** A Guss asphalt mixture is prepared by mixing relatively hard asphalt (i.e., Pen 20/40) and Trinidad Lake Asphalt (TLA) with coarse aggregate, fine aggregate and mineral filler. Table 1 lists the specification of the Guss asphalt binder

for the GSB. It is a common practice to add 20~30% TLA to the normal paving grade binder in consideration of the rut resistance at the high temperatures. The mineral matter in the TLA is about 37%, which could provide permanent deformation resistance. The fine particles in TLA were considered as portion the design asphalt content in the mixture. Table 2 indicates the specification of the Guss asphalt prepared by mixing Pen 20/40 and TLA in the ratio of 5:1 to 3:1. The mix design discussed in this paper is based on a binder of Penn 20/40 bitumen and TLA in a 75/25 blend giving a softening point of 63°C and a penetration of 22 dmm.

**Table 1. Specification of Individual Binder in Guss Asphalt.**

| Content | Specification | | Test Method |
|---|---|---|---|
| | bitumen | TLA | |
| Pen@25ºC, dmm | 20~40 | 1~4 | AASHTO T49 |
| Softening point, ºC | 55~65 | 93~98 | AASHTO T53 |
| Ductility@25ºC, cm | ≥50 | - | AASHTO T51 |
| Loss of mass, % | ≤0.3 | - | JIS K2207 |
| Solubility (TCE), % | ≥99.0 | 52.5~55.5 | AASHTO T44 |
| Flash point (C.O.C), ºC | ≥260 | ≥240 | AASHTO T48 |
| Specific gravity | ≥1.0 | 1.38~1.42 | JIS K2249 |

**Table 2. Specification of Guss Asphalt.**

| Tests | Specification | Test Method |
|---|---|---|
| Pen@25ºC, dmm | 15~30 | AASHTO T49 |
| Softening point, ºC | 58~68 | AASHTO T53 |
| Ductility@25ºC, cm | ≥10 | AASHTO T51 |
| Loss of mass, % | ≤0.5 | JIS K2207 |
| Solubility (TCE), % | 86~91 | AASHTO T44 |
| Flash point (C.O.C), ºC | ≥240 | AASHTO T48 |
| Specific gravity | 1.07~1.13 | JIS K2249 |

**Aggregate.** Table 3 lists the range of the sieve size of the aggregate used for the Guss asphalt mixture. The quantity passing through 4.75 mm is suggested to approach the lower limit to increase the durability of Guss asphalt. More than 90% of the aggregate comprises crushed sand particles. Unlike an ordinary asphalt mixture, a Guss asphalt mixture is composed by asphalt of Pen 20/40 and TLA, and about 25% of the aggregate gradation is the mineral filler passing number 200 sieve. The gradation of fillers is also shown in Table 4. A large quantity of limestone filler mixes with the binder to form the mastic inside the Guss asphalt mixture to enhance the durability.

Since aggregate needs to be heated around 300°C in the furnace for the Guss asphalt mixture, the chemical components of the aggregate surface are likely to be dramatically oxidized and changed. As a result of the surface change, the mechanical behavior of Guss asphalt mixtures with hot-bin aggregate would be different from that with cold-bin one. Fig. 2 shows that the Guss mixture mixed with cold-bin aggregate is subjected to lower rut depth than with the hot-bin one under the wheel-tracking test. To simulate the real paving process, it is important that aggregate be obtained from

hot-bins for the purpose of Guss mixture design. Otherwise, the test results may not be representative of the Guss asphalt mixture used in the field.

**Table 3. Aggregate gradation of Guss Asphalt Mixture.**

| Sieve (mm) | 1/2'' (12.5) | #4 (4.75) | #8 (2.36) | #30 (0.6) | #50 (0.3) | #100 (0.15) | #200 (0.075) |
|---|---|---|---|---|---|---|---|
| Spec. Passing % | 95~100 | 65~85 | 45~62 | 35~50 | 28~42 | 25~34 | 20~27 |
| Design Passing % | 97.3 | 69.5 | 56.9 | 37.6 | 31.3 | 28.4 | 25.7 |

**Table 4. Filler Gradation.**

| Sieve (mm) | #50 (0.3) | #100 (0.15) | #200 (0.075) |
|---|---|---|---|
| Spec. Passing % | 100 | ≥90 | ≥70 |
| Design Passing % | 100 | 100 | 99 |

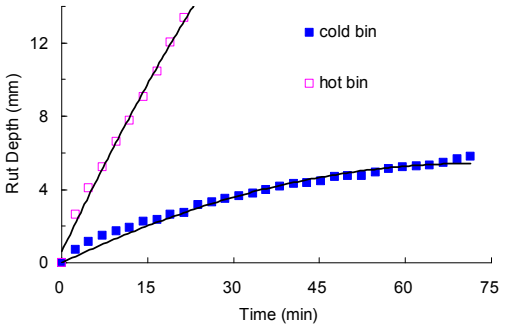



**Figure 2. Rut depth of Guss asphalt mixtures with cold- and hot-bin aggregate.**

**Criteria of Guss asphalt mixture.** The design for the Guss asphalt mixture must consider the simultaneous problems of deflection and rutting for an orthotropic steel deck plate. Various tests were carried out to meet all of the established criteria listed in Table 5. The Lueer fluidity test was to measure the workability of a Guss asphalt mixture. The indentation test and bending tests were adopted to investigate the permanent deformation and deformation compliance of the mixture. Moreover, the wheel tracking test was conducted to measure the dynamic stability of the mixture.

Fig. 3 shows the mix design procedure developed to determine the asphalt content for a Guss asphalt mixture on the GSB. In general, the quantity of the binder used in a Guss asphalt mixture is between 7 and 10% of the mixture weight. The initial design required an asphalt content of 9%, but this was gradually reduced to 8.75% as the 2$^{nd}$ binder content. This reduction was required as mix adjustments were made to produce a workable mix and to increase the dynamic stability value.

**Table 5. Specification of Guss Asphalt Mixture.**

| Test Contents | Spec | Test Method Japan Road Association |
|---|---|---|
| Lueer fluidity, 240ºC, sec. | ≤ 20 | 5.3.4 |
| Indentation, 40ºC, 52.5 kg, 30 min, mm | 1~4 | 5.3.3 |
| Wheel-tracking, 60ºC, 628 kPa, cycle/mm | ≥ 300 | 3.7.3 |
| Bending flexibility strain, -10ºC, 50 mm/min | ≥ $8.0x10^{-3}$ | 3.7.5 |

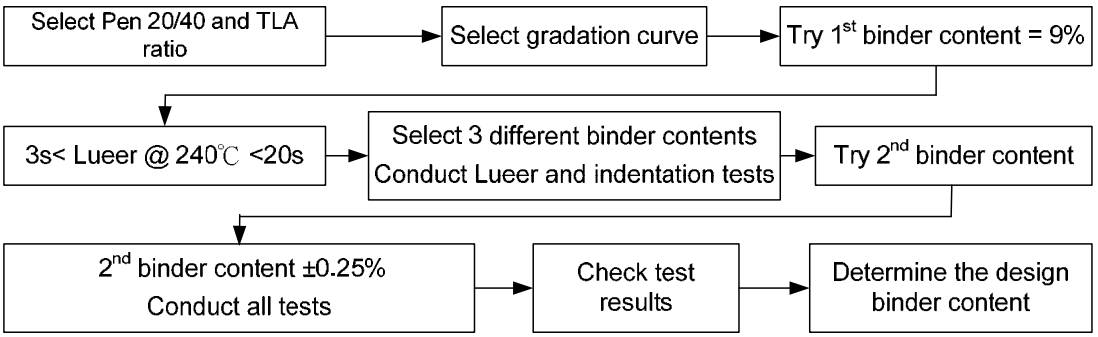



**Figure 3. Mix design procedure of Guss asphalt mixture.**

**Lueer test.** The Lueer fluidity test is consisted of a container, a support frame and a plum-shaped cylinder as shown in Fig. 4. This cylinder is made of brass with a load of 995 g marking an indicator of 50 mm on the upper part of the bar. The time required to pass through the 50-mm indicator is called the Lueer fluidity of a Guss asphalt mixture tested between 200 and 260°C. The Lueer fluidity value should be fewer than 20 sec with a Guss asphalt mixture measured at 240°C.

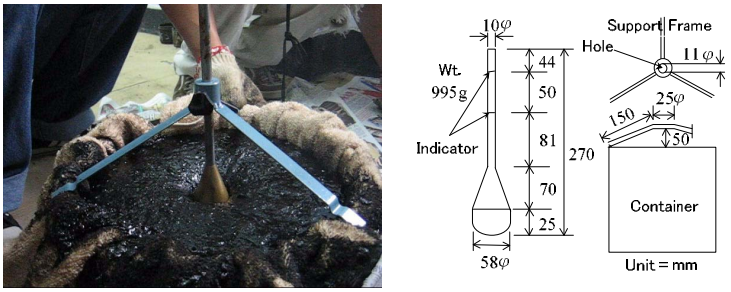
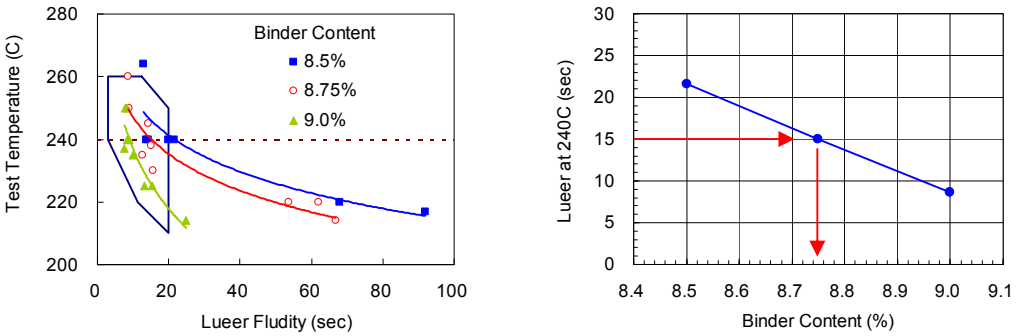



**Figure 4. Lueer fluidity test.**

Fig. 5(a) shows the test results of three different binder contents of Guss asphalt mixtures tested at temperatures between 210 and 270°C. The rectangle-like box on Fig. 5(a) is the boundary set for a Guss asphalt mixture. Any value that is outside of the box is considered as unsuitable for the paving work. Increasing the binder content would lead to the decrease in the Lueer value. A Guss asphalt mixture of 9% binder content is likely to flow like thick honey at temperature of 250°C. As the test temperature cools gradually, the Lueer fluidity value increases. The increase in the asphalt content would lead to a decrease in the Lueer Fluidity as shown Fig.5(b). In practice, a Guss asphalt mixture is poured into place and requires no compaction. To achieve this purpose, the requirement of the Lueer fluidity is set at 15 sec for a Guss

asphalt mixture tested at 240°C. According to this requirement, Fig. 5(b) demonstrates that a binder content of 8.75% was selected as the design value for the Guss asphalt mixture.




(a) Lueer vs. Temp (b) Lueer vs. Binder Content

**Figure 5. Results of Lueer fluidity test.**

**Indentation test.** The indentation test is to evaluate the deformation resistance of a Guss asphalt mixture tested at 40°C as shown in Fig. 6. Test cubes of 7-cm square of Gussasphalt are cast and allowed to cool. The cubes are then placed in a water bath maintained at 40°C for at least one hour. The specimen is initially loaded with 2.5 kg and the zero point of the dial gauge is read. The total force of 52.5 kg was then applied smoothly. There is no specific load rate for the indentation test. The loading process of the indentation test is similar to that of the penetration test for an asphalt binder. After 30 minutes the depth of penetration of the indenter is read of to the nearest 0.01 mm. At least two tests should be made and the average used. It is required that the depth of the cylinder penetration into the specimen surface range between 1 and 4 mm. If the indentation lies outside these limits, the test is repeated with another Guss asphalt mixture having a different asphalt content or a different aggregate gradation until a composition is found that complies with the indentation requirement.

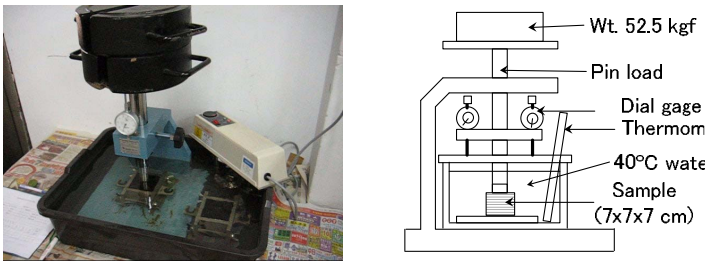



**Figure 6. Indentation test.**

Test results shown in Fig. 7 indicate that the indentation value differed little among three asphalt contents. The pin indentation depth is required to be between 1.0 and 4.0 mm. For heavy duty pavements, the indentation was reduced to 1.0 to 1.5 mm. This test was made to confirm that an asphalt content of 8.75% was satisfactory for the indentation depth of 1.5 mm.

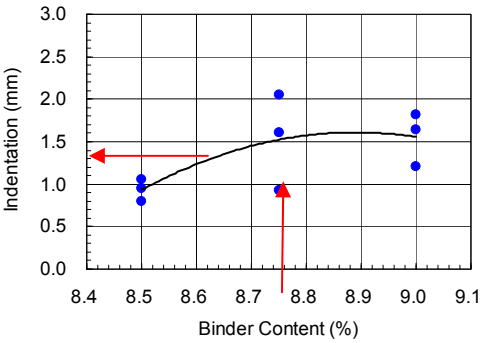



**Figure 7. Results of indentation test.**

**Bending test.** The three-point bending test is to provide information on low-temperature flexural strain. Specimens were kept at a -10°C chamber for 6 hours before testing. Fig. 8 shows that a beam (30x10x5 cm) is tested by applying a load at the midpoint of the specimen at a constant rate of 50 mm/min within a chamber controlled at -10°C. The bending flexibility strain is calculated as follows:

$$\varepsilon_{max} = \frac{6hd_{max}}{L^2}$$

where h is the height, $d_{max}$ is a mid-span beam displacement corresponding to the failure point, and L is the span of the two rest points.

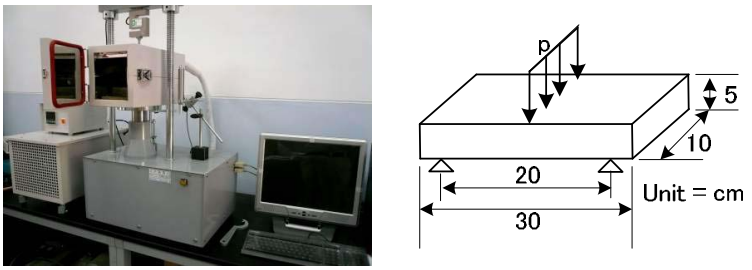



**Figure 8. Bending flexibility test.**

The deformation compliance of a Guss asphalt mixture tested at a low temperature is shown in Fig. 7. All three mixtures had higher strain than the minimum strain requirement of $8.0x10^{-3}$. The high flexibility strain of the Guss asphalt mixture indicated that it has the resistance to fatigue cracking.

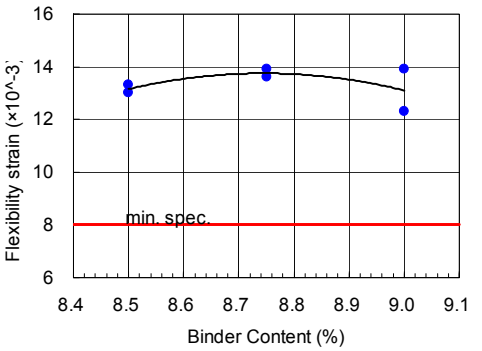



**Figure 9. Results of bending test.**

**Wheel tracking test.** A wheel-tracking tester was performed to evaluate a mixture's susceptibility to permanent deformation. This equipment is similar to the Hamburg Wheel-Tracking Device. The test was conducted on a 300x300x70 mm sample at 60°C under dry conditions. A smooth solid-steel wheel traveling at a speed of 42 cycles/min was used to correlate with rutting. The dynamic stability (DS) is calculated as follows:

$$DS = \frac{15}{d_{60} - d_{45}} \times 42$$

where $d_{60}$ and $d_{45}$ are the rut depth measured between the testing time of 45$^{th}$ and 60$^{th}$ min., respectively. The minimum requirement of the dynamic stability is 300 cycle/mm for a Guss mix. Realistically, a Guss asphalt mixture tends to be without interlocking of the aggregate, and may show a weakness of rutting. This is why the dynamic stability value of the Guss asphalt mixture is not specified to be as high as that of a conventional dense mixture, which is to be around 1500 cycle/mm. The Guss asphalt mixture is generally used in the base course of the steel deck paving as an impermeable layer to obtain the compatibility against bending and to stop the infiltration of moisture.

Fig. 10 indicates that an asphalt content of 9% would lead to a relatively low dynamic stability. With a relatively high asphalt content, aggregate is likely to suspend within the Guss mix with few interlocking points. Guss mixes with abnormally high values of dynamic stability are also less desirable because surfaces with such mixes may crack prematurely under heavy volumes of traffic.

The final selected mix design is usually the most economical one that satisfactorily meets all of the established criteria listed in Table 5. The design asphalt content is a compromise selected to balance all of the Guss mix properties. The Guss mix design criteria produced a narrow range of acceptable asphalt content that passed all of the guideline as shown in Fig. 11. The design asphalt content of the Guss asphalt mixture was suggested to be the one at the median of all the established criteria, which was 8.75%. The asphalt content selection was then adjusted within this narrow range to achieve mix properties that satisfied the requirement of the GSB project. In the end, the final design asphalt content was chosen to be at 8.55% due to heavy traffic and cost considerations.

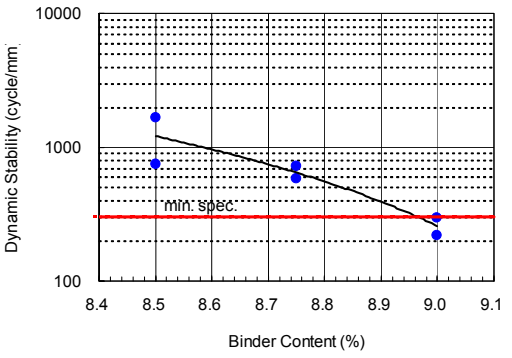



**Figure 10. Dynamic stability.**

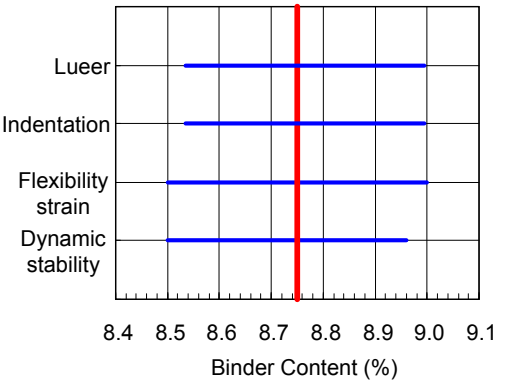



**Figure 11. Range of acceptable asphalt content.**

## CONSTRUCTION

**Preparation.** Construction of a Guss asphalt mixtures on the GSB began by shot-blasing the deck surface and placing the protection membrane system. To prepare the surface for membrane application, the steel deck plate was shot blasted to a near white surface as shown in Fig. 12(a). Within 3 hours after shotblasting, the prime coat proceeded with first application of rubberized bitumen emulsion transversely, and followed by spreading another coat of the waterproofing membrane longitudinally as shown in Fig. 12(b). Completion of the prime coat system consisted of spreading the sealant along the edges of concrete and steel structures as shown in Fig. 12(c). The sealant is used to prevent water and moisture from seepage into the steel deck.

(a) shot blasting

(b) prime coat

(c) sealant

**Figure 12. Surface preparation of a steel deck.**

**Production.** Guss asphalt mixture production was carried out in a batch asphalt plant. Modifications to the plant are necessary to heat and dry the aggregate, as well as to heat the filler which represents a substantial amount (about 25%) in the Guss mix. Mixing was accomplished by adding the aggregate and the filler into the pugmill. All materials were dry mixed for about 20 sec. The binders (i.e., Pen20/40 and TLA) were mixed with the heated aggregate for approximately 90 sec at around 200 to 220°C. At this stage, the Guss mix batches were dropped into a mobile 20-tone cooker as shown in Fig. 13. It is also a common practice on small-scale contracts to mix the bitumen and aggregate first, and then to add TLA to the cooker to blend for a certain period of time before laying. The Guss asphalt mix was agitated and heated at a high temperature in the cooker in order to obtain the workability and stability for the laydown. The Guss asphalt mix was held in the cooker for about 1 to 3 hours with the mixing temperature (say 220ºC) raised to 240~260ºC for application.

**Figure 13. Cooker in a batch plant.**

**Figure 14. Application of Guss mix.**

**Application.** The Guss asphalt mixture was delivered in a cooker to site, ready for feeding to the paver. While transporting the Guss asphalt mixture from the plant to the construction site, it is necessary to stir the mix continuously to prevent aggregate segregation and provide a uniform temperature. At a high temperature, the Guss asphalt mixture was sufficiently fluid to be poured into place without compaction, forming a voidless pavement. The laying equipment was set up on crawlers which were set to the desired height of the surface, as shown in Fig. 14. The spread of the paver can be varied up to the roadway width, say 3.5 m. The cooker fed into a heated non-vibrating screed bar, which spread the Guss asphalt mixture to level, moving forward at a rate of 1 to 2 m per minute. The steel deck surface was kept dry all the time to prevent the development of blisters caused by trapped moisture and expanding steam.

## CONCLUSIONS

The engineering properties of a Guss asphalt mixture could be evaluated by experimental tests including Lueer fluidity, indentation, bending flexibility and wheel tracking. Aggregate should be obtained from hot-bins for the mix design of a Guss asphalt mixture. Since a Guss asphalt mixture could be poured into place with no compaction, the requirement of the Lueer fluidity was found to be appropriate to set at 15 sec for a Guss asphalt mixture tested at 240°C. To combat heavy traffic, the indentation depth could be set at 1.5 mm. The Guss asphalt mixture tested in this study met the specifications of the bending flexibility and the dynamic stability. The steel deck surface must be shot-blasted and prime-coated to improve the binding properties between the Guss mix and the steel deck. The Guss asphalt mixture produced at the temperature of 200~220ºC must be placed in a cooker to be agitated and heated up to 240~260ºC before application.

## ACKNOWLEDGEMENTS

The authors are very grateful for the National Science Council and the Taiwan Area National Expressway Engineering Bureau providing financial and field supports.

## REFERENCES

Arnaud, L., and Houel, A. (2007) "Fatigue Damage of Asphalt Pavement on an Orthotropic Steel Deck," *Road Materials and Pavement Design*, Vol.8, 2007, pp.505-522.

Hicks, R.G., Dussek, I.J. and Seim, C. (2000) "Asphalt Surfaces on Steel Bridge Decks," Transportation Research Record: *Journal of the Transportation Research Board,* No. 1740, National Research Council, Washington, D.C., 2000, pp.135-142.

Hulsey, J.L., Yang, L., and Raad, L. Hulsey, J.L., Yang, L., and Raad, L. (1999) "Wearing Surfaces for Orthotropic Steel Bridge Decks," Transportation Research Record: *Journal of the Transportation Research Board*, No. 1654, National Research Council, Washington, D.C., pp.89-96.

# Laboratory Evaluation of Effects of Soaked Duration on Moisture Susceptibility of Warm Mix Asphalt Mix

Feipeng Xiao[1], Wenbin Zhao[2] and Serji N. Amirkhanian[3]
[1]Research Assistant Professor, Department of Civil Engineering, Clemson University, Clemson, South Carolina, USA, feipenx@clemson.edu
[2]Graduate Research Assistant and [3]Formerly Professor, Department of Civil Engineering, Clemson University, Clemson, South Carolina, USA,

ABSTRACT: Moisture damage in flexible pavements might cause stripping in the asphalt pavement and ultimately lead to premature failure. In addition, rising energy prices, global warming, and more stringent environmental regulations have resulted in an interest in warm mix asphalt (WMA) technologies as a mean to decrease the energy consumption and emissions associated with conventional hot mix asphalt production. In this study, the objective was to conduct a laboratory investigation of long-term influence of anti-stripping additives on moisture damage in WMA mixtures. The test results indicated that the aggregate source did not show a remarkable effect on moisture susceptibility of mixture in this study. The addition of anti-stripping additives increased the ITS, TSR, and flow values; however, the differences were not statistically significant compared to the control mixtures. Mixtures containing Asphamin® did not perform as well as others. In addition, an increase of condition duration decreases the anti-stripping resistance regardless of mixture type.

## 1. INTRODUCTION

The phenomenon of breaking the bond between the aggregate and the binder is known as stripping. A typical situation is the gradual loss of strength over the years, which causes many surface manifestations like rutting, corrugations, shoving, raveling, cracking, etc (Xiao et al. 2008; Caro et al. 2008). To prevent moisture susceptibility, proper mix design is essential. Of the many ways to prevent stripping in a pavement, the use of anti-stripping agents (ASAs) is the most common (Lu and Harvey 2006). One of the most commonly used ASAs in the United States is hydrated lime. Others include liquid ASAs such as amines, di-amines, liquid polymers, and solids like Portland cement, fly-ash, flue dust, etc. Pavement contractors usually prefer liquid ASAs as they are relatively easy to use (Lu and Harvey 2006). However, ASAs from an approved list of sources should not be blindly added as some ASAs are aggregate and asphalt specific, and therefore, may not be effective to be used in all mixes; they could even be detrimental at times. Thus, a proper study of each mix should be done by systematically testing the mix for moisture susceptibility using tests like indirect tension strength (ITS), Lottman's and boiling water tests; among many other tests, in the laboratory.

Recently, the "warm mix asphalt" (WMA) is widely being used in the hot HMA industry as a mean of reducing energy requirements and lowering emissions. WMA can significantly reduce the mixing and compacting temperatures of asphalt mixtures, by either lowering the viscosity of asphalt binders, or causing foaming in the binders. Reduced mixing and paving temperatures decreases the energy required to produce HMA, reduces emissions and odors from plants, and makes for better working conditions at both the plant and the paving site (Prowell et al. 2007).

Moisture damage is usually not limited to one mechanism rather than the result of a combination of many processes. From a chemical standpoint, the literature is clear that though neither asphalt nor aggregate has a net charge, but components of both have nonuniform charge distributions, and both behave as if they have charges that attract the opposite charge of the other material (Abo-Qudais et al. 2005). The foaming process caused by WMA additive makes the charge re-distribution more complex and thus may affect the moisture susceptibility of mixture. Especially, at the mixture temperature of 100 to 140 °C (212 to 280 °F), the aggregate may not be completely dried during mixing process though some of states in US and other countries have specifications that require a completely dry aggregate in WMA mixtures (Xiao et al. 2008). There are not many research projects conducted in the area of determining the effects of the moist aggregates with WMA additives which may result in moisture damage and further lead to the failure of the pavement.

Most laboratory studies conducted to date have assessed the short-term moisture susceptibility of mixes. The ITS test, for instance, tests the moisture susceptibility of mixes after conditioning the samples in water for only 24 hours. This may not always be representative of the actual field conditions, and thus might be, in some cases, a misrepresentation of the actual moisture susceptibility of the pavement itself. In a previous study conducted by Lu and Harvey (2006), the long term effects of moisture on the effectiveness of the ASAs were studied. It was observed that most of the detrimental effects of moisture occurred in the first four months. This paper reports the research findings of the evaluation of long term moisture susceptibility of several mixes. Indirect tensile strength (ITS) tests were performed on dry and wet conditions after 1 day, 60 days, and 90 days (Gandhi et al. 2009)

## 2. EXPERIMENTAL MATERIALS AND PROCEDURES

The Superpave method of mix design for a nominal maximum aggregate size of 12.5 mm was followed for this study (Figure 1). A total of 18 mix designs (1 binder sources × 2 aggregate sources × 3 different ASA treatments × 3 different WMA additives) were conducted. The bulk specific gravity, maximum specific gravity, voids in mineral aggregate (VMA) and voids filled with asphalt (VFA) were obtained or calculated to determine the optimum asphalt content of all the 18 mixes.

The aggregates used in this study were obtained from two sources, denoted as B and C. The engineering properties of aggregate sources are shown in Table 1. Aggregate B (schist) is a metamorphic rock while aggregate source C (granite) is composed predominantly of quartz and potassium feldspar. Aggregate C has a higher percentage of $Al_2O_3$ and $SiO_2$ compared to aggregate B. These properties may affect the performance of mixtures containing anti-striping and WMA additives.

Table 1 Engineering properties of Aggregates B and C

| Aggregate Source | LA Abrasion Loss (%) | Absorption (%) | Specific Gravity | | | Soundness % Loss at 5 Cycles | | | Sand Equivalent | Hardness |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Dry (BLK) | SSD (BLK) | Apparent | 11/2 to3/4 | 3/4 to3/8 | 3/8 to #4 | | |
| B | 34 | 0.60 | 2.780 | 2.800 | 2.830 | 0.4 | 0.6 | 0.9 | 35 | 5 |
| C | 24 | 0.50 | 2.620 | 2.640 | 2.660 | 0.3 | 0.7 | 0.9 | 50 | 6 |

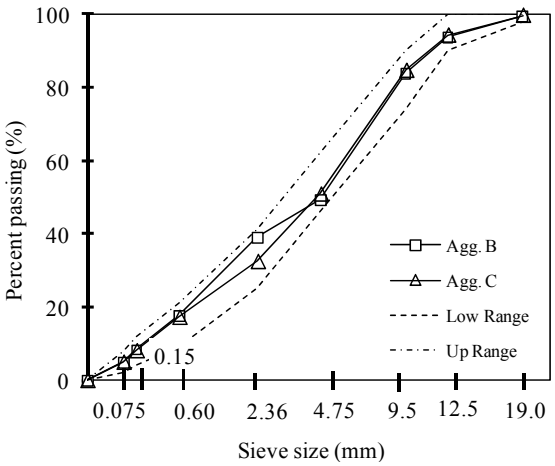



Figure 1 Aggregate gradations

Two different anti-stripping agents were used in this project. They were commercially available hydrated lime and one liquid ASA denoted as 1 and 2, respectively. The third treatment was the control, no ASAs, denoted by '0' in this paper. In the mixes containing hydrated lime as the ASA, 1% hydrated lime by weight of the aggregate, was added in a slurry form. In the mixes containing liquid ASAs, 0.5% liquid ASA, by weight of the binder, was added to the binder. This ASA content rate was based on the 0.25% - 0.75% recommended by the suppliers.

Asphamin® and Sasobit® were used in this study as two WMA additives. Asphamin® (A) is Sodium–Aluminum–Silicate which is hydro thermally crystallized as a very fine powder which contains approximately 21% crystalline water by weight. By adding it to an asphalt mix, the fine water spray is created as all the crystalline water is released, which results in volume expansion in the binder. This, in general,

increases the workability and compactability of the mix at lower temperatures. Sasobit® (S) is a long chain of aliphatic hydrocarbons obtained from coal gasification using the Fischer-Tropsch process. After crystallization, it forms a lattice structure in the binder which is the basis of the structural stability of the binder containing Sasobit®. The mixtures without any WMA were referred to as virgin mixture (V) in this paper.

After determining the optimum binder content for each of the mix designs, ITS samples were made for each of the mix type. Six samples for each conditioning duration (e.g. 1, 60, and 90 days) were prepared to test the *ITS*. The air void of each sample is 7 ± 1%. Three of which were stored as dry samples at 25 ± 1˚C, and three were stored as wet samples. If the samples were to be tested after 1 day, the wet samples were submerged in a water bath (60 ± 1˚C) for 24 hours followed by submersion in another water bath at 25 ± 1˚C for 2 hours before testing. For conditioning durations of more than 1 day, the wet samples were submerged in a water bath at 25 ± 1˚C for one day short of that specific duration (e.g. 59 and 89 days) and then they were submerged in a water bath (60 ± 1˚C) for 24 hours followed by submersion in water at 25 ± 1˚C for 2 hours before testing. All the wet samples were vacuum saturated to a saturation level of 70% to 80% before immersing in water.

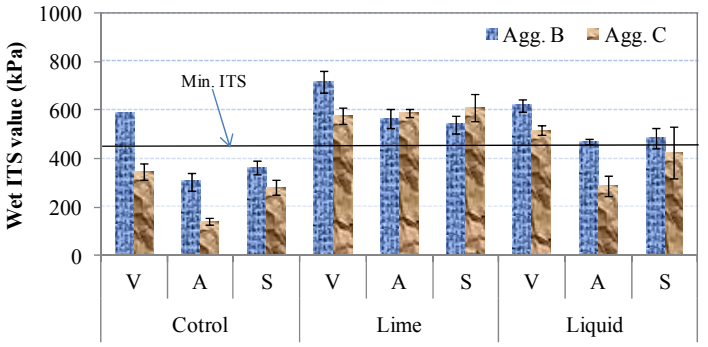
The wet ITS, tensile strength ratio (TSR), and flow were used as the measure of stripping for each of the mixes. To study the effects of ASAs and aggregates on the mixes, Analysis of variance (ANOVA) was then performed to test the null hypothesis (ITS, TSR, and flow of each treatment and block variables are not significantly different from each other) at the 5% level of significance.

## 3. EXPERIMENTAL RESULTS AND DISCUSSIONS

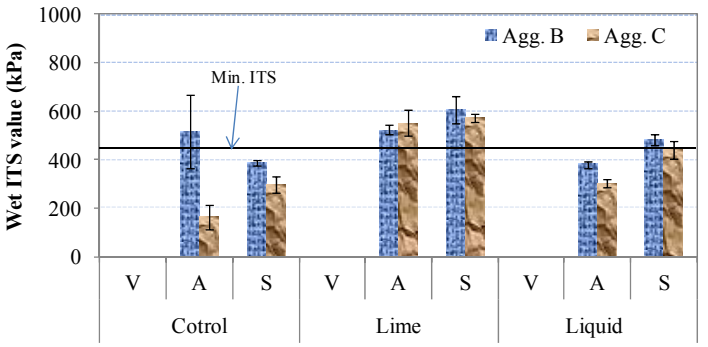
### *3.1Wet ITS analysis*

The wet ITS values of mixtures that were conditioned under water for 90 days are shown in Figure 2(a). It can be noted that, generally, the mixture made with aggregate B has a slightly higher wet ITS value for all the mixture types. Also, additional lime and anti-stripping liquid, in most cases, slightly increases the wet ITS values. In general, hydrated lime seemed to be the most effective ASA. Moreover, with respect to the effect of WMA additive on the wet ITS value, the mixture without WMA additive shows a higher value. In addition, Figure 2(a) illustrates that the mixture containing Asphamin® has the lowest ITS values regardless of aggregate types and anti-stripping additive. On the other hand, as shown in Figure 2(a), it can be noted that some wet ITS values of mixtures containing WMA additives, especially for Asphamin®, are less than 448 kPa (65 psi). The results indicate for the materials tested for this research work, Asphamin® did not performed as well as Sasobit® for the mixtures containing no antistripping additives.

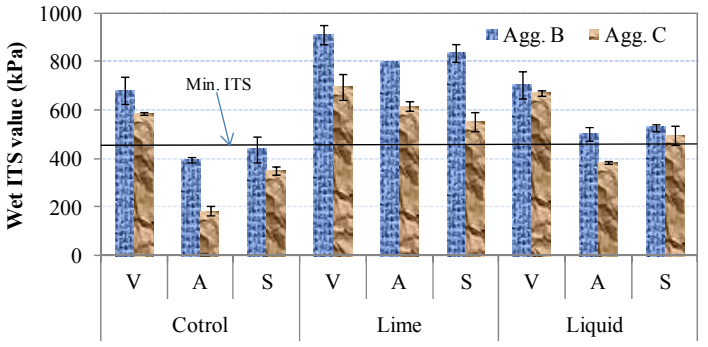
As shown in Figures 2(b) and 2(c), generally, after conditioning the samples for 1 or 60 days, the wet ITS values of these samples show a similar trend with those values of samples that were conditioned for 90 days. However, these wet ITS values of mixtures conditioned 1 day are the highest among three conditioning durations, as expected. Thus, it can be considered that the conditioning duration reduces the anti-stripping resistance of the materials.




(a)




(b)




(c)

Figure 2 Wet ITS values after conditioning, (a) 90 days; (b) 60 days; (c) 1 day

The statistical analysis (t-statistic) results in Table 2 indicate that, with respect to the effect of aggregate sources (B and C) on the wet ITS value, in most cases, there is no significant difference between two aggregate sources regardless of mix types. Table 3 illustrates that the influence of anti-stripping additives on the wet ITS value is generally significant for WMA mixtures but is not significant for those ITS values for mixtures without WMA (virgin). As mentioned earlier, it seems that, as expected, the anti-stripping additive plays an important role in improving the wet ITS values. On the other hand, statistical analysis results also show that there are significant differences of wet ITS values amongst three mixtures (control, hydrated lime and anti-stripping liquid). As shown in Table 4, statistical analysis results based on the effect of WMA additives illustrate that there are no significant differences in the wet ITS values as the mixtures used the hydrated lime as an anti-stripping additive. However, significantly different wet ITS values can be found in mixtures containing any WMA additives as they do not included any anti-stripping additive or included liquid additive.

Table 2 Statistical analysis of moisture properties in terms of aggregates B and C

| | wet ITS | | | TSR | | | flow | | |
|---|---|---|---|---|---|---|---|---|---|
| | 90 | 60 | 1 | 90 | 60 | 1 | 90 | 60 | 1 |
| Control | N | N | N | N | N | N | N | N | Y |
| Lime | N | N | Y | Y | N | Y | N | N | - |
| Liquid | N | N | N | N | N | N | Y | N | N |

*Note: Y: significant difference; N: No significant difference*

Table 3 Statistical analysis of moisture properties in terms of anti-stripping additive types (control, lime and liquid)

| | wet ITS | | | TSR | | | flow | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0~1 | 0~2 | 1~2 | 0~1 | 0~2 | 1~2 | 0~1 | 0~2 | 1~2 |
| Virgin | N | N | N | N | N | N | - | - | - |
| Asphmin | Y | N | Y | Y | Y | Y | Y | Y | N |
| Sasobit | Y | Y | Y | Y | Y | N | Y | Y | N |

*Note: 0- control; 1-Lime; 2-Liquid; Y: significant difference; N: No significant difference*

Table 4 Statistical analysis of moisture properties in terms of WMA additive types (virgin, Asphamin® and Sasobit®)

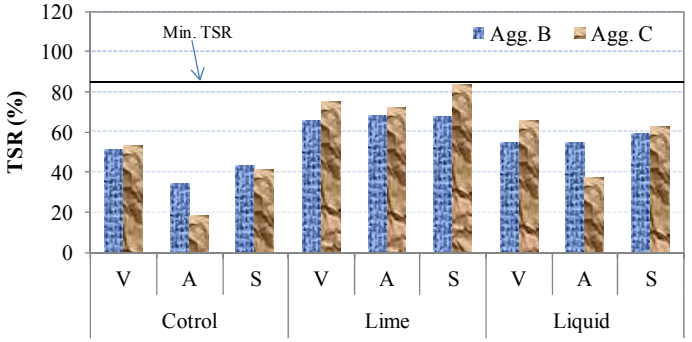
| | wet ITS | | | TSR | | | flow | | |
|---|---|---|---|---|---|---|---|---|---|
| | V~A | V~S | A~S | V~A | V~S | A~S | V~A | V~S | A~S |
| Control | Y | Y | N | Y | N | N | - | - | Y |
| Lime | N | N | N | N | N | N | - | - | N |
| Liquid | Y | Y | Y | Y | N | Y | - | - | N |

*Note: V- virgin; A-Asphmin; S-Sasobit; Y: significant difference; N: No significant difference*

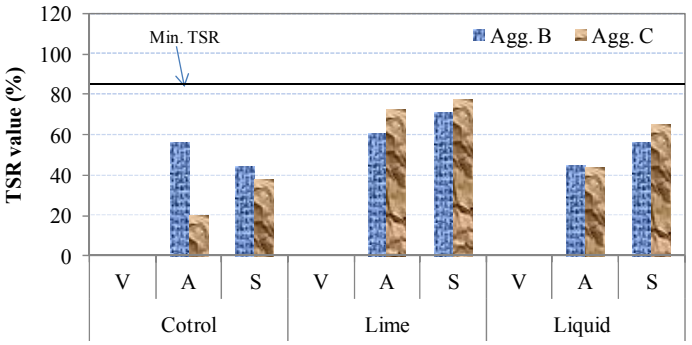
*3.2TSR analysis*

The TSR values of mixtures after 90 days of conditioning are shown in Figure 3(a). It can be noted that there is not noticeable trend between two aggregate sources. However, similar to the wet ITS values, addition of lime and anti-stripping liquid slightly increases the TSR values and hydrated lime seemed to be more effective than liquid. Moreover, with respect to the effect of WMA additive on the TSR values, Figure 3(a) illustrates that the mixture containing Asphamin® has the lowest TSR values. In addition, Figure 3(a) shows that all TSR values are less than 85%, a minimum value of SCDOT's specification. As a result, similar to the wet ITS analysis, it seems that although the addition of anti-stripping additives can increase the TSR value the mixture exhibit moisture susceptibility.

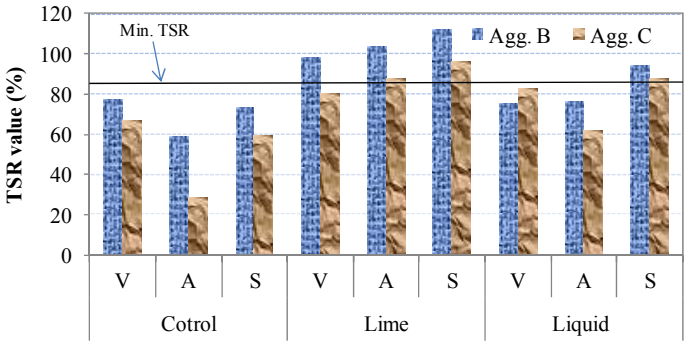
Similar analysis was performed for the samples that were conditioned for 1 and 60 days. As shown in Figures 3(b), generally, TSR values of these samples show a similar trend with those values of samples that were conditioned for 90 days. All TSR values are less than 85%. However, for those samples conditioned for 1 day (Figure 3(c)), their TSR values are the highest and some of them are higher than 85% (hydrated lime samples). From this standpoint, as mentioned earlier, the conditioning duration was also considered to reduce the anti-stripping resistance of mixture.




(a)




(b)




(c)

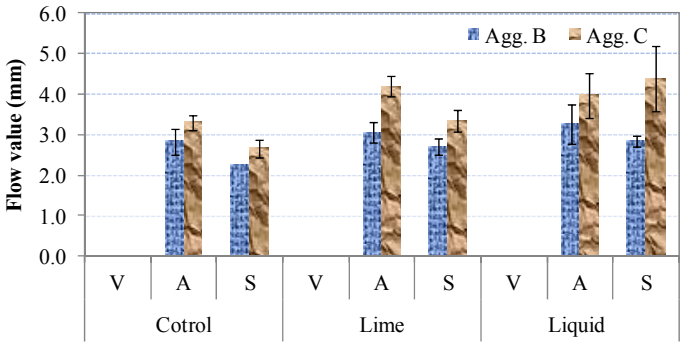
Figure 3 TSR values of mixes after conditioning, (a) 90 days; (a) 60 days; (c) 1 day

The statistical analysis for TSR was shown in Table 2, it can be seen that, with respect to the effect of aggregate sources (B and C) on the TSR value, in most cases, there is no significant difference between two aggregate sources. Table 3 illustrates that the influence of anti-stripping additive on the TSR value is also significant for WMA mixtures but is no significant for those TSR values from mixtures without WMA mixture (virgin). On the other hand, statistical analysis results of TSR values also show that there are significantly differences amongst three mixtures (control, hydrated lime and anti-stripping liquid). These TSR statistical analysis is similar to the Wet ITS analysis. However, Table 4 illustrate that, in general, there are no significant differences in the TSR values of mixtures containing various WMA additives.

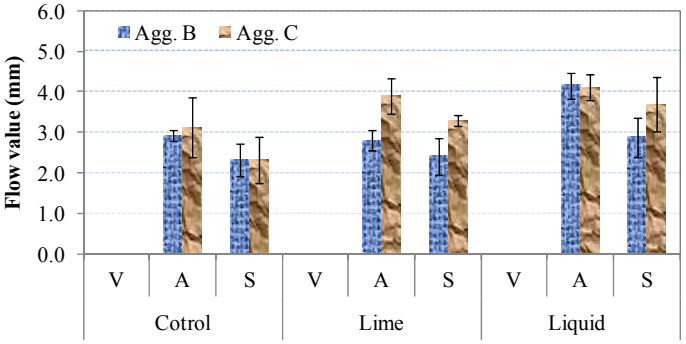
*3.3Wet flow analysis*

Figure 4 illustrates that samples made with aggregate C having a higher flow value for all WMA mixtures. The mixture containing Asphamin® additive exhibits a slightly higher flow value. As expected, the samples conditioned for longer duration (90 days) had higher flow values compared to samples tested after 1 day.

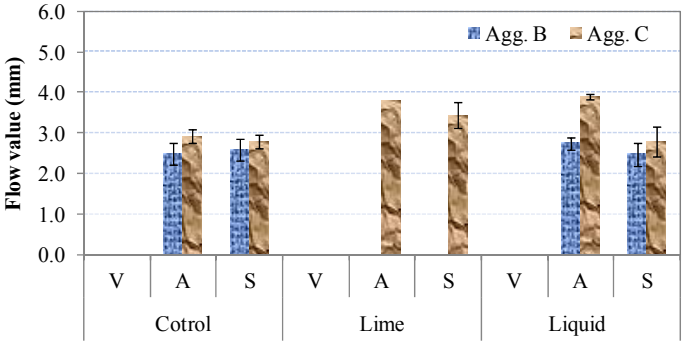
The statistical analysis result shows that the influence of aggregate sources on flow is not significant (Table 2). However, Table 3 illustrates that the influence of anti-stripping additive on the flow value is generally significant, but there is no significant difference in flow values between mixtures containing two anti-stripping additives. As shown in Table 4, statistical analysis results illustrate that there are no significant differences in the flow values for the mixtures containing WMA additive.




(a)




(b)




(c)

Figure 4 Flow values of mixes after conditioning, (A) 90 days; (B) 60 days; (C) 1 day

## 4. FINDINGS AND CONCLUSIONS

The following conclusions were drawn based upon the experimental results obtained from a laboratory investigation of effects of utilizing anti-stripping additives in WMA mixtures:

- In general, the mixture from aggregate B had a slightly higher ITS value and a slightly lower flow values compared to the aggregate source. However, in

most cases, there is no significant differences in ITS, TSR, and flow values (at $\alpha = 0.05$).

- The addition of anti-stripping additives increased the ITS, TSR, and Flow values. In addition, the samples containing hydrated lime indicated better moisture susceptibility compared to the liquid ASA used in this study. The statistical analysis shows that there are not significant differences between ITS, TSR, and flow values for mixtures containing WMA additive.
- Asphamin® shows a negative effect on anti-stripping resistance. Statistical result indicates that as the hydrated lime was employed in the mixture, there were no significant differences in ITS, TSR, and flow values. However, the effects of WMA additive are significant as the mixture with liquid or without anti-stripping additive.
- The increase of condition duration reduces the ITS and TSR values, but the flow value increase. As a result, an increase of condition duration decreases the anti-stripping resistance regardless of mixture type.

## 5. REFERENCES

Abo-Qudais, S., M., Mohammad W., 2005, New chemical antistripping additives for bituminous mixtures, *Journal of ASTM International*, v. 2(8), p 87-97

Caro, S., Masad, E., Bhasin, A., Little, D.N. 2008, Moisture susceptibility of asphalt mixtures, Part 1: Mechanisms, *International Journal of Pavement Engineering*, v 9 (2), Moisture Induced Damage in Asphalt, p 81-98

Gandhi T., Xiao F., and Amirkhanian S. 2009, Estimating indirect tensile strength of mixtures containing anti-stripping agents using an artificial neural network approach, *International Journal of Pavement Research and Technology*, V.2(1), p 1-12

Lu Q, and Harvey, J.T., 2006, Laboratory Evaluation of Long-term Effectiveness of Antistripping Additives, *Journal of the Transportation Research Board 2006,* Annual Meeting CD ROM

Prowell, B., Hurley, G., and Crews, E. 2007, Field performance of warm-mix asphalt at national center for asphalt center for asphalt technology test track, *Journal of the Transportation Research Board* No. 1998, p 96-102

Xiao F. and Amirkhanian S.N. "Laboratory Investigation of Moisture Damage in Rubberized Asphalt Mixtures Containing Reclaimed Asphalt Pavement" *The International Journal of Pavement Engineering,* Vol.10, No.5, pp.319-328, 2009

Xiao, F., Jordan J., Amirkhanian, S., 2009. Laboratory investigation of moisture damage in warm mix asphalt containing moist aggregate, *Journal of Transportation Research Board* (TRB), Vol. 2126, pp. 115-124, 2009

# Effect of Test Methods on Viscosity Temperature Susceptibility Characterization of Asphalt Binders for the Mechanistic-Empirical Pavement Design Guide

Syed Waqar Haider[1], M. Waseem Mirza[2], Ashvini K. Thottempudi[3], Javed Bari[4], and Gilbert Y. Baladi[5]

[1]Assistant Professor, Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824; PH (517) 353-9782; FAX (517) 432-1827; email: syedwaqa@egr.msu.edu
[2]Professor, Department of Transportation Engineering and Management Engineering, University of Engineering and Technology Lahore, Pakistan; mw.mirza@hotmail.com
[3]Graduate Research Assistant, Department of Civil and Environmental Engineering, Michigan State University; thottemp@msu.edu
[4]Sr. Pavement Design and Development Engineer, Arizona Department of Transportation (ADOT), Phoenix, AZ-85009; jbari@azdot.gov
[5]Professor, Department of Civil and Environmental Engineering, Michigan State University;baladi@egr.msu.edu

## ABSTRACT

The Superpave asphalt binder specifications are intended to improve pavement performance by limiting the potentials for plastic deformation and cracking. Hot-mix asphalt (HMA) behavior depends on the binder's—temperature susceptibility, and the rate of loading. The binder viscosity temperature susceptibility (VTS) characterization plays a major role in determining the HMA dynamic modulus ($E^*$) master curve, especially when input Level 3 is used in the M-E PDG. For input Levels 2 and 3, $E^*$ could be predicted by the Witczak equation, whereas, three test methods are specified to determine the asphalt binder VTS. These tests are (i) dynamic shear rheometer (DSR) for determining shear modulus ($G_b^*$) and phase angle ($\delta$); (ii) conventional tests such as penetration, kinematic and absolute viscosities, and ring and ball softening point; and (iii) rotational viscosity. All test methods can be utilized to determine the ASTM "A-VTS" parameters. In this paper, the A-VTS parameters for twenty two (22) neat and polymer modified binders were obtained from the three test methods and compared. The results show that A-VTS parameters from the three tests could be significantly different. Further, the impact of the VTS parameters' variations on the $E^*$ master curve and on the predicted flexible pavement performance are discussed.

## INTRODUCTION

Asphalt binder is a viscoelastic material i.e., its behavior depends on both temperature and loading frequency. As asphalt is a byproduct of petroleum crude, the source of crude has a significant impact on the complex chemical nature of the binders. Therefore, considering the intricate chemical composition and variability involved in characterizing an asphalt binder, its properties are characterized into two broad groups: physical and rheological properties (Bari and Witczak 2007; Bari and

Witczak 2006; Dongre et al. 2005). Traditionally, the physical and rheological characterization of binders is accomplished by conducting conventional tests (CT): penetration, softening point, viscosity (absolute and kinematic) and ductility etc. However, rheological characterization of binders has gained a widespread acceptance in practice due to new developments e.g., Superpave testing protocols (Asphalt Institute 2001; Asphalt Institute 2003). The rheological properties of binders include temperature and shear susceptibility, rate of loading dependence and stiffness. These properties are determined by performing Superpave tests: dynamic shear rheometer (DSR) and bending beam rheometer (BBR) tests after preparing samples by using rolling thin film oven (RTFO) and pressure aging vessel (PAV) to simulate short- and long-term binder aging, respectively. In addition, rotational viscometer (RV) is used to characterize its temperature-viscosity behavior over a wide range of temperature (Asphalt Institute 2003). Flexible pavement distresses, which are related to these binder properties, significantly impact the short- and long-term pavement performances. Consequently, asphalt binder characterization plays a vital role in predicting flexible pavement performance. While conventional tests (absolute and kinematic viscosity) can be used to measure viscosity of asphalt binders at limited temperatures, RV can determine temperature-viscosity relations at a wide range of different temperatures and loading rates. The measure of binder's resistance to flow under external forces is termed as the binder viscosity ($\eta$). Also, viscosity is determined as the ratio of shear stress ($\tau$) to shear strain ($\gamma$) rate at a given temperature. From rheological perspective, binders behave as Newtonian liquids at very high temperatures—the ratio is nearly constant, while at low or intermediate temperatures those behave like non-Newtonian fluids—the ratio varies.

Considering the asphalt binder behavior, the Superpave asphalt binder specifications are intended to improve pavement performance by limiting the potentials for plastic deformation and cracking. As asphalt binder rheological properties control the viscoelastic behavior of hot-mix asphalt (HMA); therefore, HMA behavior depends on the binder viscosity temperature susceptibility and the rate of loading. The mechanistic-empirical pavement design guide (M-E PDG) uses a dynamic modulus ($E^*$) master curve to characterize the HMA behavior for different temperatures and loading rate (NCHRP Project 1-37A 2004). The binder viscosity temperature susceptibility (VTS) characteristics play a major role in determining the HMA master curve. Three test methods are specified to determine the asphalt binder VTS. These test methods include: (i) dynamic shear rheometer (DSR) for determining $G_b^*$ and $\delta$; (ii) conventional tests such as penetration, kinematic and absolute viscosities, and ring and ball softening point; and (iii) rotational viscosities.

In this study, the above mentioned three test methods were employed to characterize 22 original, modified, or blended binders. The overall goal of the study is to facilitate implementation of Superpave specification in Pakistan. Consequently, appropriate binder performance grades (PG) will be recommended for pavement construction in different regions according to the indigenous climatic (high and low temperature) needs. The main objectives of this paper are to: (a) compare binder VTS parameters obtained from three test methods, and (b) evaluate the impact of these differences (if any) on $E^*$ and flexible pavement performance predictions using the M-E PDG.

## VISCOSITY-TEMPERATURE SUSCEPTIBILITY OF BINDERS

The binder viscosity temperature susceptibility (VTS) characteristics play a major role in determining the HMA master curve, especially when design levels 2 and 3 of the M-E PDG are used. For these two input levels, $E^*$ could be predicted by the Witczak equation (Bari and Witczak 2006), whereas, three test methods are specified to determine the asphalt binder VTS. The ASTM model given by Equation (1) is used to define the viscosity-temperature relationship.

$$\log\log\eta = A + VTS\log T_R \quad (1)$$

where:

$\eta$ = viscosity, cP
$A$ = regression intercept
$VTS$ = regression slope (viscosity-temperature susceptibility parameter)
$T_R$ = temperature, degree Rankine

All three test methods, mentioned above, can be utilized to determine A and VTS parameters empirically from the laboratory data as discussed below.

**Conventional testing.** Some practical situations or available data may demand simple binder characterization without conducting a full range of binder tests. Therefore, existing relationships can be used in those cases to determine viscosities at different temperatures. For unmodified asphalt binders, the ring and ball softening point corresponds to a viscosity of 13,000 Poise and penetration test results can be converted to viscosity using Equation (2) (Mirza and Witczak 1996).

$$\log\eta = 10.5 - 2.26\log(Pen) + 0.00389\left[\log(Pen)^2\right] \quad (2)$$

where:

$Pen$ = Penetration value using 100 g loading for 5 seconds

In addition, conventional tests are used to measure absolute and kinematic viscosities of binders at 60 and 135 $^{o}$C, respectively. Thus by using conventional tests, the binder viscosity-temperature susceptibility can be determined between temperature range of 25 to 135 $^{o}$C. Once, the viscosities are determined at different temperatures, Equation (1) can be used to determine A-VTS parameters empirically.

**Rotational viscometer.** A rotation viscosity test is used to determine the flow characteristics of the asphalt binder at high temperatures (80 to 185 $^{o}$C). The resulting temperature-viscosity relationship is used for determining mixing and compaction temperatures of the hot-mix asphalt (HMA). The measured viscosities from the test can also be used to determine A-VTS parameters empirically by using Equation (1).

**Dynamic shear rheometer (DSR).** In the current Superpave binder characterization methodologies, the binder complex shear modulus ($G_b^*$) and the phase angle ($\delta$) are obtained from the DSR test. As compared to A-VTS parameters obtained by conventional tests, the DSR data can be used to quantify the effects of both temperature and loading rate on binder shear stiffness. However, measured binder $G_b^*$

and $\delta$ data can be used to estimate viscosities by using Equation (3) (Bari and Witczak 2007; NCHRP Project 1-37A 2004)

$$\eta = \left(\frac{|G_b^*|}{\omega}\right)\left(\frac{1}{\sin\delta_b}\right)^{a_0+a_1\omega+a_2\omega^2} \quad (3)$$

where:

$\eta$ = viscosity, cP
$|G_b^*|$ = binder shear modulus, Pa
$\delta_b$ = binder phase angle, degree
$\omega$ = angular frequency, rad/s
$a_0$, $a_1$, $a_2$ = fitting parameters, 3.639, 0.1314, and -0.0009, respectively

The Equation (3) is reduced to Equation (4) for $\omega = 10$ rad/s , which is the specified test frequency for temperature sweep in the Superpave performance grading system

$$\eta = \left(\frac{|G_b^*|}{10}\right)\left(\frac{1}{\sin\delta_b}\right)^{4.8628} \quad (4)$$

where; all variables are same as defined above

It should be noted binder viscosity obtained through the above mentioned methods is independent of loading frequency, meaning that a constant binder viscosity (Newtonian behavior) is assumed for different frequencies (Bari and Witczak 2007). The DSR test can be used for characterizing binder behavior at intermediate temperatures. Due to measurement limitations, a DSR temperature range of between 7 to 82 $^{o}$C is typical. The A-VTS parameters can be determined empirically by using Equation (1) for defining viscosity-temperature relationship for an asphalt binder. The brief detail of binders and laboratory testing results are presented in the next section.

## LABORATORY TESTING RESULTS

To evaluate local available materials in Pakistan, samples of the virgin asphalt binders produced by the two major oil refineries were obtained. It should be noted that the properties of the asphalt binders are function of several variables. For example, these properties depend on the source of the crude oil, refining process, storage facilities, and mixing procedures used in refineries. Both refineries produce limited grades of asphalt binders. Typical binder grades produced and obtained for the study from these refineries include: Attock Refinery [Pen grade 80/100, 60/70 and 60/70 polymer modified binder (PMB)], and Karachi Refinery [Pen grades 80/100, 60/70, and 40/50]. The Superpave specifications are intended to improve performance by limiting the potential of the asphalt binder to contribute to permanent deformation, low temperature, and fatigue cracking in asphalt pavements. The evaluation of local binders is intended to explain how each of the new test parameters relates to the observed pavement performance and how to select the asphalt binder grade for a specific project. Consequently, the evaluation will determine whether or not the local binders meet the Superpave performance grading specifications in the local climate.

Further, considering the prevailing environment and traffic conditions in Pakistan, blending of different types of asphalt binders may be needed to produce grades meeting the Superpave requirements. In addition, under certain climatic and loading conditions, polymer modification may be required. In order to optimize the use of polymers in a cost-effective manner, polymer with varying percentage mixed with the neat asphalt were examined in this study. Only two polymer types were used in the testing. Reactive polymer was used with the binders from Attock refinery while non-reactive polymer was employed with the binders from Karachi refinery. The test matrices for the blending and polymer modification for the two refineries (Attock and Karachi) are given in Tables 1 and 2, respectively.

**Table 1. Blending of original binders and polymer modification-Attock Refinery**

| Binder type | 80/100 | 60/70 | 60/70 P | Polymer* |
|---|---|---|---|---|
| 80/100 | - | - | | 1.35, 1.7, 2.0 |
| 60/70 | 0, 20, 50, 100[1] | - | | 1.35, 1.7, 2.0 |
| 60/70 (P) | - | - | 100 | - |

* Elvaloy® RET, a reactive elastomeric terpolymer with the chemical designation EGA (ethylene/glycidyl/ acrylate).

**Table 2. Blending of original binders and polymer modification-Karachi Refinery**

| Binder type | 80/100 | 60/70 | 40/50 | Polymer** |
|---|---|---|---|---|
| 80/100 | - | - | 0, 20, 50 | 2.5, 3.5, 4.5 |
| 60/70 | 0, 20, 50, 100[1] | - | - | - |
| 40/50 | - | 50 | - | - |

** Elvaloy® AC, a non-reactive polymer, [1] Percentage of column binder with row binder types for blending. For example, 0 means 0% 60/70 with 100% 80/100 binder.

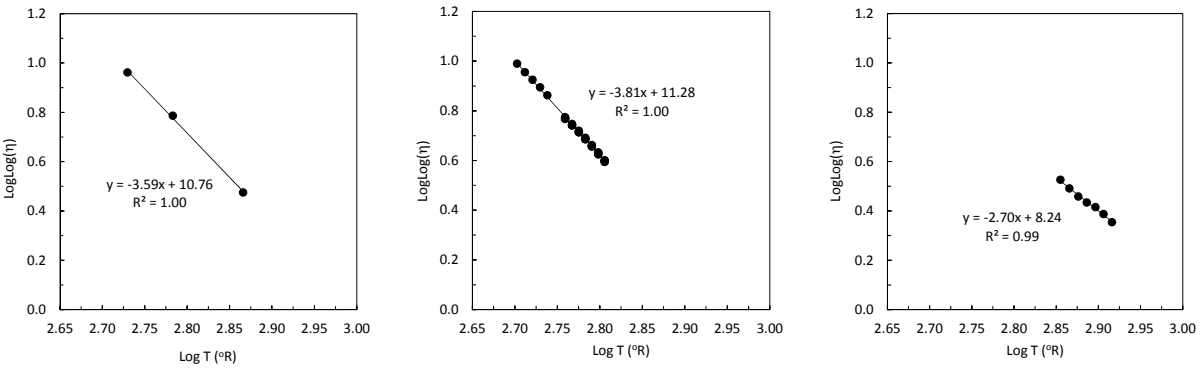
The above two test matrices resulted in a total of 22 binder types. The codes for all binders tested and analyzed are presented in Table 3. All these 22 binders (neat and modified) were individually characterized using conventional (penetration, softening point and viscosity) and Superpave (DSR, BBR and RV) tests. For each binder type, three replicates were tested and averages of the measurements were considered in the analyses.

**Comparison of A-VTS from different methods.** According to the Equation(1), the respective dependent [loglog(η)] and independent [logT in Rankine] variables were plotted for each binder type and test method. The intercept and slope of the linear line represent the A and VTS parameters in Equation(1), respectively. Figure 1 shows an example of the linear relationship for the binder #1 for the three test methods. It should be noted that test temperature range is dependent in the test method. The temperature ranged from 25 to 135 $^{o}$C, 100 to 185 $^{o}$C, and 7 to 82 $^{o}$C for conventional, RV, and DSR tests, respectively.

**Table 3. The binder coding for 22 binders used in the study**

| No. | Binder Code | Binder Description |
|---|---|---|
| 1 | APMB | Attock Polymer Modified |
| 2 | A6/7 | Attock Pen 60/70 |
| 3 | A8/10 | Attock Pen 80/100 |
| 4 | K4/5 | Karachi Pen 40/50 |
| 5 | K6/7 | Karachi Pen 60/70 |
| 6 | K8/10 | Karachi Pen 80/100 |
| 7 | BA6/7(20)8/10(80) | Blended Attock: Pen 60/70 (20%) and Pen 80/100 (80%) |
| 8 | BA6/7(50)8/10(50) | Blended Attock: Pen 60/70 (50%) and Pen 80/100 (50%) |
| 9 | PA8/10(1.35) | Polymer Modified Attock 80/100 with 1.35% Elvaloy |
| 10 | PA8/10(1.70) | Polymer Modified Attock 80/100 with 1.70% Elvaloy |
| 11 | PA8/10(2.00) | Polymer Modified Attock 80/100 with 2.00% Elvaloy |
| 12 | PA6/7(1.35) | Polymer Modified Attock 60/70 with 1.35% Elvaloy |
| 13 | PA6/7(1.70) | Polymer Modified Attock 60/70 with 1.70% Elvaloy |
| 14 | PA6/7(2.00) | Polymer Modified Attock 60/70 with 2.00% Elvaloy |
| 15 | BK6/7(20)8/10(80) | Blended Karachi Pen 60/70 (20%) and Pen 80/100 (80%) |
| 16 | BK6/7(50)8/10(50) | Blended Karachi Pen 60/70 (50%) and Pen 80/100 (50%) |
| 17 | BK8/10(20)4/5(80) | Blended Karachi Pen 80/100 (50%) and Pen 40/50 (80%) |
| 18 | BK8/10(50)4/5(50) | Blended Karachi Pen 80/100 (50%) and Pen 40/50 (50%) |
| 19 | BK6/7(50)4/5(50) | Blended Karachi Pen 60/70 (50%) and Pen 40/50 (50%) |
| 20 | PK8/10(2.5) | Polymer Modified Karachi 60/70 with 2.5% AC |
| 21 | PK8/10(3.5) | Polymer Modified Karachi 60/70 with 3.5% AC |
| 22 | PK8/10(4.5) | Polymer Modified Karachi 60/70 with 4.5% AC |

Note: Shaded binders were used for the M-E PDG analysis.




(a) Conventional (b) DSR (c) RV

**Figure 1. The A-VTS parameters plots for binder #1**

Figure 2 shows the VTS parameter for all binders tested by all methods. In general, the VTS parameters from RV test is the lowest than those measured from the other two test types (CT and DSR). The VTS parameters calculated based on CT are slightly lower than those measured by the DSR test for all binder types used in the study. The descriptive statistics shown in Table 4 confirm these general observations. The 95% confidence mean intervals for VTS parameter by test type are also given in

the table. The statistical comparison of multiple means using one way analysis of variance (ANOVA) showed that mean of parameter A obtained by using CT and DSR tests is significantly different from that of RV while the overall means between CT and DSR test methods are not different at 5% significance level. Similar results were obtained when VTS parameters were compared among different methods.

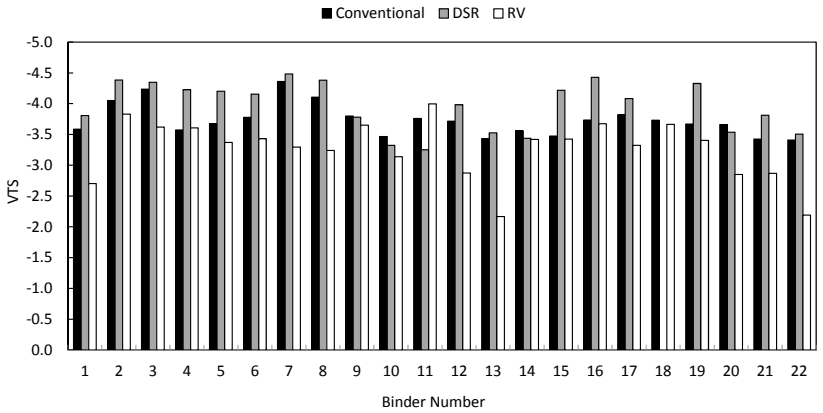
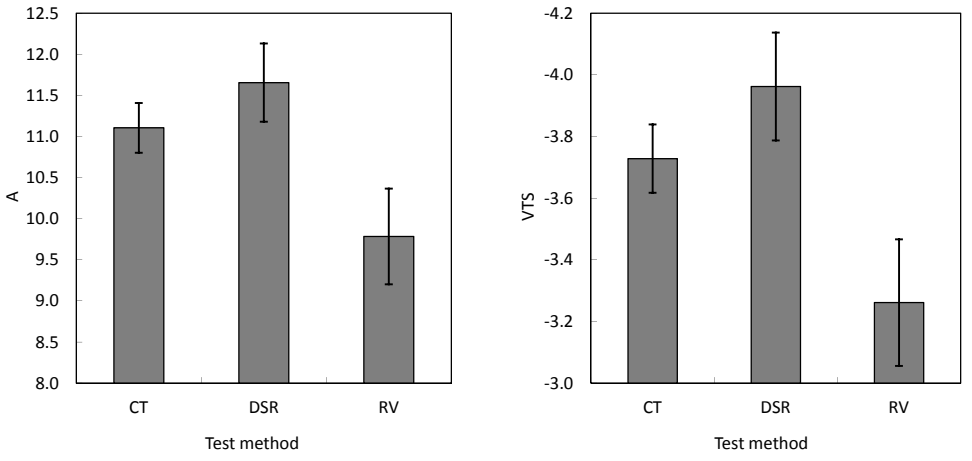



**Figure 2. The VTS parameters for all binders by test method**

Figure 3 shows the mean comparisons of A-VTS parameters among different test methods. The 95% confidence intervals are also shown in the figure. The means between test methods are not significantly different from each other if the 95% confidence intervals overlap. On the other hand, non-overlapping intervals show a significant statistical difference between means.

**Table 4. Descriptive statistics of VTS by test method**

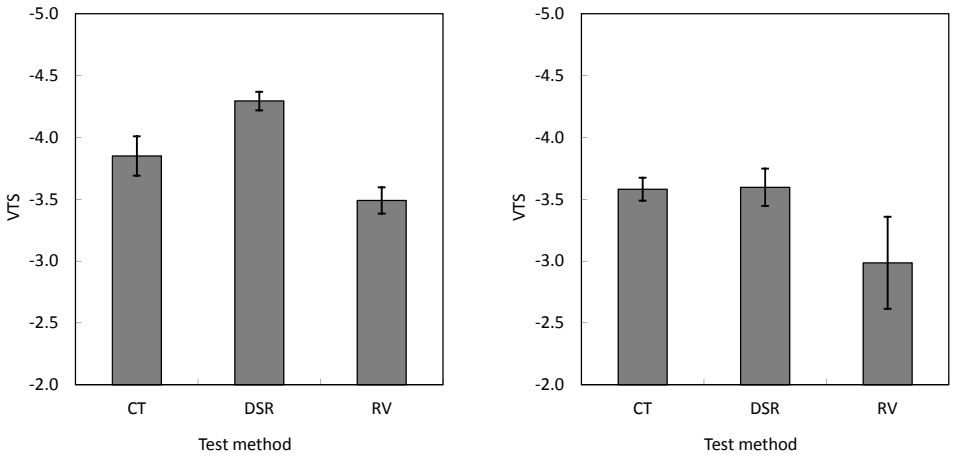
| Test Type | n | Mean | Std. Dev. | Std. Error | 95% CI | |
|---|---|---|---|---|---|---|
| | | | | | Lower | Upper |
| CT | 22 | -3.73 | 0.26 | 0.055 | -3.84 | -3.62 |
| DSR | 21 | -3.96 | 0.40 | 0.087 | -4.14 | -3.79 |
| RV | 22 | -3.26 | 0.48 | 0.103 | -3.47 | -3.06 |




(a) A (b) VTS

**Figure 3. Average A-VTS parameters measured by different test methods**

In order to further investigate the reasons for the A-VTS mean differences between test methods, the binders were stratified according to type i.e., neat vs. modified. The A-VTS mean comparisons were made again within each binder type. Figure 4 shows the results of those comparisons. Figure 4a shows that for neat binders, the A-VTS parameters are significantly different among all test methods while the difference only exists between RV and other two test types in case of modified binders.




(a) Neat binders (b) Modified binders

**Figure 4. Average A-VTS parameters measured by binder type**

The above results were based on the comparison A-VTS parameters means by test methods from the statistical standpoint. However, in order to identify the practical differences among test methods, the impact of A-VTS parameters on $E^*$ master curves and predicted pavement performance was investigated using the M-E PDG version 1.0 software.

**Impact of A-VTS characterization on $E^*$ and pavement performance.** In this analysis, two binder types were identified based on the DSR test results. The neat binder #7 had the largest VTS parameter (-4.5) while the modified binder #11 possessed the lowest (-3.2). The results for same binder types were considered for other test methods (i.e., CT and RV). The $E^*$ mater curves were generated by matching the A-VTS parameters using the M-E PDG by changing binder PGs for Level 3 inputs. A typical flexible pavement cross-section having 8-inch thick HMA and base/subbase (MR=40,000 psi) layers each over an A-6 subgrade soil (MR=14,000 psi) were assumed. A wet-freeze climate for Michigan and 10 million ESALs for a 20 years design life were used in the analysis. Figure 5 presents the $E^*$ master curves for the two binders types within each test method. Basically, the figure shows the sensitivity of the master curves to A-VTS parameters. Higher the absolute value of the VTS parameter, the binder will be more temperature susceptible and consequently the master curve will shift to the left on the reduced time scale (lower time or higher frequency) or vice versa. This also means the $E^*$ of the binder will be significantly reduced at low frequencies or at higher temperatures. As expected, the impact of A-VTS on $E^*$ seems to be lower at low temperatures or high frequencies. The same six $E^*$ masters curves were used to analyze a flexible pavement with the

M-E PDG software having 8-in thick HMA and base layers. The objective of the analysis was to evaluate the impact of variations in the master curves on the predicted flexible pavement performance. Figure 6 illustrates the predicted pavement performance after 20 years of service life. It should be noted that all variables in the analysis were kept at the same level except the *E** master curves.

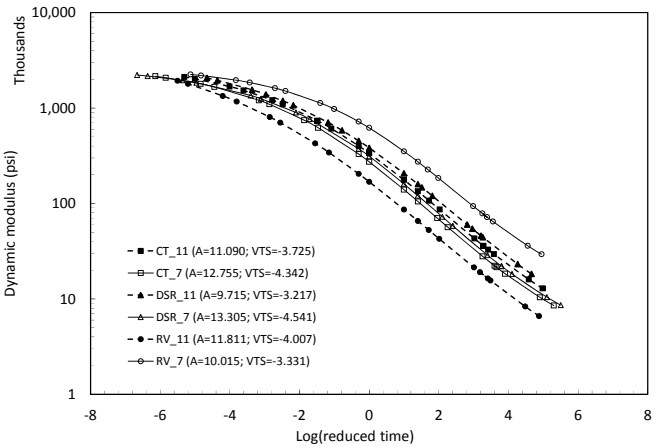
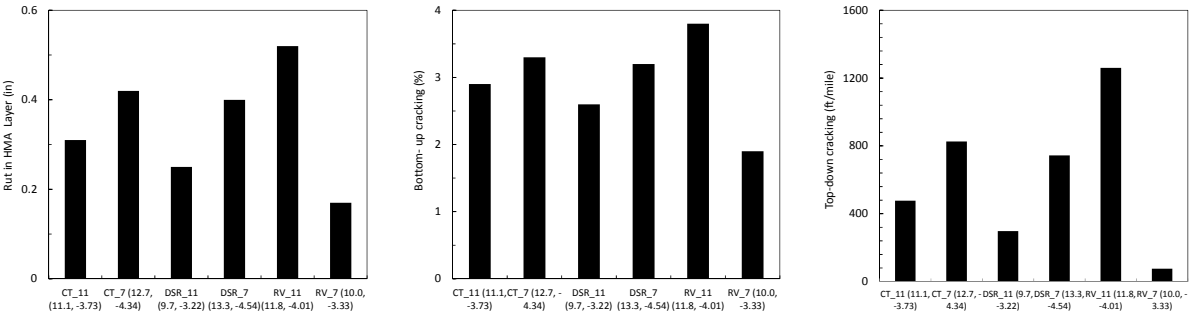



**Figure 5. Effect of VTS parameters on HMA master curve**

For convenience, the A-VTS parameters are also shown in the figure. The results of the performance prediction show that A-VTS parameters (also implies to the variations in *E** master curves) have an important effect on the pavement performance, especially within each test method. The HMA layers with higher absolute VTS value are expected to have more HMA rutting, fatigue and longitudinal cracking in the long-run. When comparing performance trends among test method, the results show that relative to DSR results (as reference) if the A-VTS are determined from CT, the pavement performance will be slightly over predicted. However, if A-VTS are used based on the RV test, the performance prediction will be significantly different from the results based on other two test methods.




(a) HMA rutting (b) fatigue cracking (c) Longitudinal cracking

**Figure 6. Effect of VTS parameters on pavement performance**

## DISCUSSION OF RESULTS

The binder viscosity temperature susceptibility (VTS) characteristics play a major role in determining the HMA master curve, especially when design input Level 3 is

used in the M-E PDG. The Level 3 inputs for A-VTS parameters are based on default values depending on the binder PG, viscosity grading, and penetration grading systems. Because the *E** master curve determination is sensitive to A-VTS parameters, the default A-VTS parameters according to different grading systems (i.e., penetration, viscosity and PG grading) may not be able to capture the unique binder behavior and thus the expected pavement performance. It will be useful if the A-VTS parameters could be directly used as an input in the M-E PDG to circumvent this concern. As mentioned before, the range of temperature used to characterize a binder is an important factor considering its Newtonian and non-Newtonian behavior. The high temperature range used during RV binder testing clearly shows that A-VTS parameters are significantly different from other test types considered in this paper. Therefore, the most logical choice of temperature range to be used in binder testing should correspond to actual surface temperatures that pavement experience at a particular location. The DSR test temperature range corresponds closely to the field temperatures of the all test methods considered in the paper.

## CONCLUSIONS

In this paper, the A-VTS parameters for twenty two (22) neat and polymer modified binders were obtained from the three test methods and compared. The results show that A-VTS parameters from the three tests could be significantly different. Furthermore, the impact of the VTS parameter on the *E** master curve shows that it is significantly affected by the temperature susceptibility of the binder. As a consequence, the pavement performance can be affected. The results also show that among various test methods, CT, RV, and DSR can be used to determine A-VTS parameters for the pavement design purposes; however, the designer should be aware of the consequence of using an input A-VTS which may not truly represent the field conditions (i.e., the temperature range used in testing).

## ACKNOWLEDGEMENT

The authors acknowledge the Pakistan-United States Science and Technology Cooperative Program for funding this research.

## REFERENCE

Asphalt Institute. (2001). "Superpave Mix Design, Superpave Series No. 2 (SP-2)." Asphalt Institute, Lexington, KY.

Asphalt Institute. (2003). "Performance Graded Asphalt Binder Specification and Testing, Superpave Series No. 1 (SP-1)." Asphalt Institute, Lexington, KY.

Bari, J., and Witczak, M. (2007). "New Predictive Models for Viscosity and Complex Shear Modulus of Asphalt Binders: For Use with Mechanistic-Empirical Pavement Design Guide." *Transportation Research Record: Journal of the Transportation Research Board*, 2001(-1), 9-19.

Bari, J., and Witczak, M. W. "Development of a new revised version of the Witczak E Predictive Model for hot mix asphalt mixtures." Savannah, GA, United States, 381-424.

Dongre, R., Myers, L., D'Angelo, J., Paugh, C., and Gudimettla, J. (2005). "Field evaluation of Witczak and Hirsch Models for predicting dynamic modulus of hot-mix asphalt." *Jounal of Association of Asphalt Paving Technologist*, 74, 381-442.

Mirza, M. W., and Witczak, M. W. "Development of a global aging system for short and long term aging of asphalt cements." Portland, OR, USA, 393.

NCHRP Project 1-37A. (2004). "Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement structures." *Final Report*, National Cooperative Research Program (NCHRP), Washington D.C.

# Characterizing Temperature Susceptibility of Asphalt Binders Using Activation Energy for Flow

Syed Waqar Haider[1], M. Waseem Mirza[2], Ashvini K. Thottempudi[3], Javed Bari[4], and Gilbert Y. Baladi[5]

[1]Assistant Professor, Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824; PH (517) 353-9782; FAX (517) 432-1827; email: syedwaqa@egr.msu.edu
[2]Professor, Department of Transportation Engineering and Management Engineering, University of Engineering and Technology Lahore, Pakistan; mwmirza1485@hotmail.com
[3]Graduate Research Assistant, Department of Civil and Environmental Engineering, Michigan State University; thottemp@msu.edu
[4]Sr. Pavement Design and Development Engineer, Arizona Department of Transportation (ADOT), Phoenix, AZ-85009; javed_bari@yahoo.com
[5]Professor, Department of Civil and Environmental Engineering, Michigan State University;baladi@egr.msu.edu

## ABSTRACT

The temperature-viscosity relationships for twenty two neat and modified asphalt binders were evaluated using conventional tests (CT) such as penetration, kinematic and absolute viscosities, and ring and ball softening point; rotational viscometer (RV); and dynamic shear rheometer (DSR). The temperature ranged from 25 to 135 $^{o}$C, 100 to 185 $^{o}$C, and 7 to 82 $^{o}$C for conventional, RV, and DSR tests, respectively. Results from several binders studies have revealed that binders having similar penetration, viscosity and performance grade (PG) could show dissimilar rheological behaviors. Hence, there is a need to fully characterize asphalt binders to capture their rheology over a wide range of temperatures and loading frequencies. The activation energy (AE) concept was utilized in this study to characterize the temperature susceptibility of the asphalt binders. Correlations between AE and binder rheological properties (e.g., $G^*/\sin\delta$ and $G^*\sin\delta$) were evaluated. In addition, a relationship between AE and useful temperature range (UTR) determined from the performance grades (PG) was developed. The advantages of characterizing the relative temperature susceptibility of the binders using AE are discussed. Finally, discussion on mixing and compaction temperatures, especially for polymer modified binders, by using AE is also included in this paper.

## INTRODUCTION

Asphalt binders are generally viscoelastic liquids; their short-time loading behavior is like that of elastic solid, long-time behavior resembles that of viscous liquids and at intermediate-time loading they behave as viscoelastic liquids. Also, for a fixed time of loading, low temperatures lead to a short-time loading behavior while at higher temperatures their response simulates long-time loading behavior (Jongepier and Kuilman 1969). Rheological properties of viscoelastic materials reflect the

relationship between load and deformation. These properties are fundamental in understanding the behavior of these materials. The rheological properties of these materials play important roles in the evaluation and selection of paving materials, and in the analysis and design of asphalt pavements (Zeng et al. 2001). Due to ease of testing and complex nature of the material, asphalt binder properties are characterized into two broad groups: physical and rheological properties (Bari and Witczak 2007; Bari and Witczak 2006; Dongre et al. 2005). The need for both types of properties is further emphasized in the literature because of significant difference in the rheological behavior of binders even though they may possess the similar physical properties or performance grade (PG) (Bahia et al. 2001; Haider et al. 2009; Nam et al. 2004). Traditionally, the physical characterization of binders is accomplished by conducting convention tests (CT): penetration, softening point, viscosity (absolute and kinematic) and ductility, etc. The rheological test protocols are used by the most of highway agencies to better characterize the binder's viscoelastic behavior as expected in the field conditions (Asphalt Institute 2003). These properties are determined by performing Superpave tests: dynamic shear rheometer (DSR) and bending beam rheometer (BBR) tests after preparing samples by using rolling thin film oven (RTFO) and pressure aging vessel (PAV) to simulate short- and long-term binder aging, respectively. In addition, rotational viscometer (RV) is used to characterize binder's temperature-viscosity behavior that determines the mixing and compaction temperatures of the binder (Asphalt Institute 2003).

Viscosity is the measure of the shear resistance to the flow of the liquids and is dependent on the temperature of the liquid. As the temperature increases, the thermal energy of the molecules increases and the resistance to the flow decreases; hence the viscosity of the liquid decreases. When liquid flows, layers of liquid molecules slide over each other while intermolecular forces resist the motion and cause resistance to flow. This phenomenon results in an activation energy (AE) barrier for viscous flow which must be exceeded for flow to occur (Salomon and Zhai 2002a; Salomon and Zhai 2002b). Equation (1) is known as "Arrhenius" equation and is used to model the viscosity-temperature dependency of asphalt binders.

$$\eta = Ae^{\frac{E_f}{RT}} \tag{1}$$

where;

$\eta$ = viscosity, Pa.s
$A$ = regression constant
$E_f$ = activation energy for flow, kJ/mol
$T$ = temperature, degree Kelvin
$R$ = universal gas constant=8.314 J/mol/k

Equation (1) can be linearized by taking natural logarithm as follows:

$$\ln \eta = E_f / RT + \ln A \tag{2}$$

Equation (2) can be used to determine the AE for flow. The typical activation energy for a modified binder was found to be around 67 kJ/mol (Maze 1996). In another study, the average activation energy values for unmodified and modified mixtures were around 205 kJ/mol and 202 kJ/mol, respectively (Pellinen et al. 2004).

Salomon and Zahi (2002b) used rotational viscometer to determine the viscosity-temperature curves for various binder grades (3 Pen. grades and 6 PGs). The measured activation energy for various binders ranged from 55 to 90 kJ/mol. The binders having same PGs exhibited significantly different AEs. The binders with low activation energy were found to be less susceptible to the temperature changes. The modified asphalt binders showed non-Newtonian behavior within the temperature range of 110 to 160 $^{o}$C. It was also found that higher the asphaltene content in the binder, higher is the activation energy. Aging of the binder due to oxidation increases the intermolecular forces, which results in a higher resistance to flow and the activation energy. Increase in the polymer content in asphalt binder decreases the activation energy of the flow. Dongre et al. (2005) reported AE between 192 to 243 kJ/mol for six binders (PG 62-30 to PG 78-30). The AE was found to be highly correlated to the Useful Temperature Range (UTR) of the asphalt binders. Therefore, it was also concluded that AE can be used to estimate $E^*$ without the need of testing binders at various frequencies and temperatures. The UTR was defined as the algebraic sum of high and low PG of asphalt binders.

Pavements in Pakistan are subjected to a large variety of conditions, the surface temperature may vary from -10 to more than 80 $^{o}$C (cold and warm regions) while at the same time stresses may be applied very slowly or quickly by overloaded trucks moving at creep speeds. Thus it is essential to study the properties of the asphalt binders over a wide range of temperatures and loading times to safeguard premature rutting on flexible pavements. Binder properties at long loading times are conveniently determined by viscometers while for short loading times mechanical dynamic testing is more suitable (Robson 1969). Three test methods were used in this study to characterize the 22 original, modified, or blended binders' viscosity temperature susceptibility. These test methods include (i) dynamic shear rheometer (DSR) for determining complex shear modulus ($G_b^*$) and phase angle ($\delta$); (ii) conventional tests (CT) such as penetration, kinematic and absolute viscosities, and ring and ball softening point; and (iii) rotational viscosity (RV). The overall goal of the study is to facilitate the implementation of Superpave specification in Pakistan. As a result, the binder performance grades (PG) will be recommended for different regions for pavement construction according to the local climatic (high and low temperature) needs. The main objectives of this paper are to (a) compare binders AE obtained from three test methods, and (b) evaluate the temperature susceptibility of binders using the AE concept.

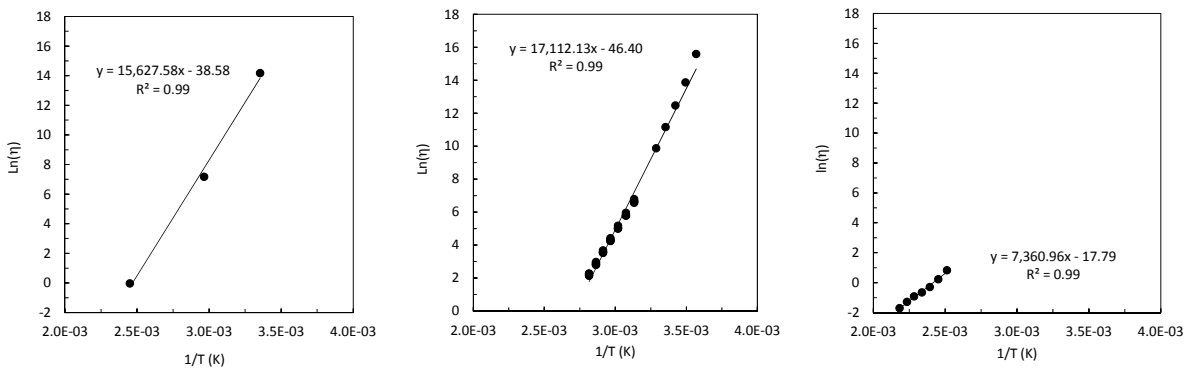
## LABORATORY TESTING RESULTS

To evaluate locally available paving materials in Pakistan, samples of the asphalt binders produced by the two oil refineries were obtained. Both refineries produce limited grades of asphalt binders. Typical binder grades produced and obtained for the study from these refineries include: Attock Refinery [Pen. Grades 80/100, 60/70 and 60/70 (PMB)], and Karachi Refinery [Pen. Grades 80/100, 60/70, and 40/50]. It should be noted that the properties of the asphalt binders are function of several variables. For example, these properties depend on the source of the crude oil, the refining process, the storage facilities, and the mixing procedures used in refineries.

The Superpave asphalt binder specifications are intended to improve pavement performance by limiting the potential for the asphalt binder to contribute to permanent deformation, low temperature and fatigue cracking in asphalt pavements. The evaluation of local binders is needed to explain how each of the Superpave test parameters relates to the observed pavement performance and how to select the asphalt binder grade for a specific project. Consequently, the evaluation will determine whether or not the local binders meet the Superpave performance grading (PG) specifications in the local climate.

Further, considering the prevailing environment and traffic conditions in Pakistan, blending of different types of asphalt binders may be needed to produce grades meeting the Superpave requirements. In addition, under certain climatic and loading conditions, polymer modification may be required. In order to optimize the use of polymers in a cost-effective manner, polymer with varying percentage with neat asphalt were examined in this study. Only one polymer type was used in the testing. The codes and descriptions for all binders tested and analyzed are presented in Table 1. The 22 binders (neat and modified) were individually characterized using conventional (penetration, softening point and viscosity) and Superpave (DSR, BBR and RV) tests. For each binder type three replicates were tested and averages of the measurements were considered in the analyses.

**Activation energy of flow by different test methods.** According to Equation (2), the dependent variable [ln (η)] and independent variable [1/T in Kelvin] were plotted for each binder type and test method. The intercept and slope obtained from the linear regression represent the $E_f/R$ and ln$A$ parameters in Equation (2). The slope parameter was multiplied by the Universal Gas Constant (UGC) to determine the AE. Figure 1 shows an example of the linear relationship of binder #1 for the three test methods. It should be noted that the test temperature range is dependent in the test method. The temperature ranged from 25 to 135 $^{o}$C, 100 to 185 $^{o}$C, and 7 to 82 $^{o}$C for conventional, RV and DSR tests, respectively.




**(a) Conventional** (b) DSR (c) RV

**Figure 1. Calculation of AE for binders #1**

The slopes shown in Figure 1 should be multiplied by UGC for AE estimation. These variations in slopes for binder #1 show that AE is affected by the test method. In

addition to several other differences among test methods, the testing temperature range could be one of the most important factors in this regard.

**Table 1. The binder coding for 22 binders used in the study**

| No. | Binder Code | Binder Description | PG Grade |
|---|---|---|---|
| 1 | APMB | Attock Polymer Modified | 76-6 |
| 2 | A6/7 | Attock Pen 60/70 | 58-12 |
| 3 | A8/10 | Attock Pen 80/100 | 58-12 |
| 4 | K4/5 | Karachi Pen 40/50 | 64-6 |
| 5 | K6/7 | Karachi Pen 60/70 | 64-12 |
| 6 | K8/10 | Karachi Pen 80/100 | 58-12 |
| 7 | BA6/7(20)8/10(80) | Blended Attock: Pen 60/70 (20%) and Pen 80/100 (80%) | - |
| 8 | BA6/7(50)8/10(50) | Blended Attock: Pen 60/70 (50%) and Pen 80/100 (50%) | - |
| 9 | PA8/10(1.35) | Polymer Modified Attock 80/100 with 1.35% Elvaloy | 64-12 |
| 10 | PA8/10(1.70) | Polymer Modified Attock 80/100 with 1.70% Elvaloy | 70-12 |
| 11 | PA8/10(2.00) | Polymer Modified Attock 80/100 with 2.00% Elvaloy | 70-12 |
| 12 | PA6/7(1.35) | Polymer Modified Attock 60/70 with 1.35% Elvaloy | 64-12 |
| 13 | PA6/7(1.70) | Polymer Modified Attock 60/70 with 1.70% Elvaloy | 70-12 |
| 14 | PA6/7(2.00) | Polymer Modified Attock 60/70 with 2.00% Elvaloy | 70-12 |
| 15 | BK6/7(20)8/10(80) | Blended Karachi Pen 60/70 (20%) and Pen 80/100 (80%) | - |
| 16 | BK6/7(50)8/10(50) | Blended Karachi Pen 60/70 (50%) and Pen 80/100 (50%) | - |
| 17 | BK8/10(20)4/5(80) | Blended Karachi Pen 80/100 (50%) and Pen 40/50 (80%) | - |
| 18 | BK8/10(50)4/5(50) | Blended Karachi Pen 80/100 (50%) and Pen 40/50 (50%) | - |
| 19 | BK6/7(50)4/5(50) | Blended Karachi Pen 60/70 (50%) and Pen 40/50 (50%) | - |
| 20 | PK8/10(2.5) | Polymer Modified Karachi 60/70 with 2.5% AC | 64-12 |
| 21 | PK8/10(3.5) | Polymer Modified Karachi 60/70 with 3.5% AC | 64-12 |
| 22 | PK8/10(4.5) | Polymer Modified Karachi 60/70 with 4.5% AC | 64-12 |

Figure 2 shows the estimated AE for all binders by three test methods. The AE values calculated based on RV testing is the lowest among all test methods. The AE values from DSR testing are always higher than other test methods for all the binders tested. Note that the binder type i.e., neat (N) or modified (M) is also included in this figure. The relative ranking of the binders depends on the test method according to these results. For example, binder #11, #19 and #7 show the maximum AE values based on RV, DSR and conventional tests, respectively. According to the literature, mentioned before, higher the AE of flow, more sensitive the binder will be to temperature changes. Conversely, lower AE indicates that the asphalt binder is less sensitive to temperature changes. Table 2 shows the descriptive statistics for AE for all 22 binders for various test methods. From the results, it can be observed that RV test showed the highest coefficient of variation (CoV) (about 15%) while AE determined from the conventional tests has the lowest CoV (about 5%). This also implies that higher range of AE was shown by RV test results, which can capture more binder behavior variations among the 22 binders considered. Figure 3 presents the mean comparisons for AE obtained by different tests. The 95% confidence intervals are also shown in the figure. The AE Means between test methods are not significantly different from

each other if the 95% confidence intervals overlap. On the other hand, non-overlapping intervals show a statistical significant difference between Means.

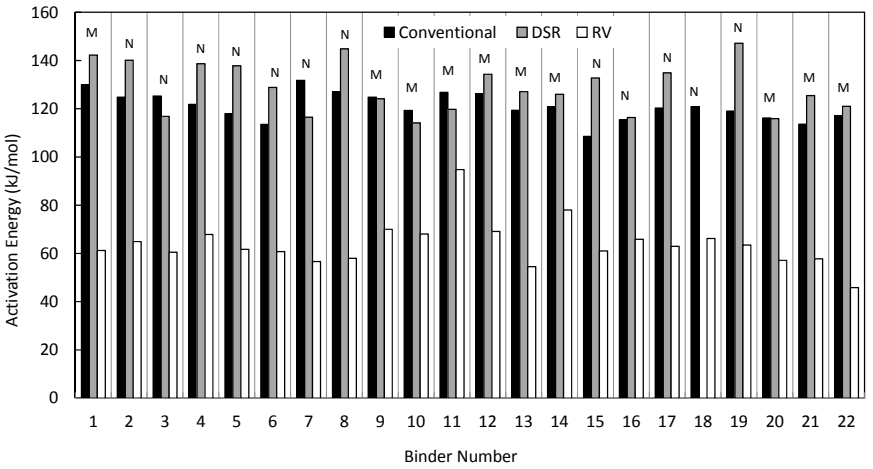



**Figure 2. AE for all binders for different test methods**

**Table 2. Descriptive statistics for AE**

| Descriptive Statistics | Test type | | |
|---|---|---|---|
| | CT | DSR | RV |
| Avg. | 120.9 | 128.8 | 63.9 |
| Std. | 5.76 | 10.45 | 9.48 |
| CoV | 4.76% | 8.12% | 14.82% |
| N | 22 | 21 | 22 |
| Min. | 108.50 | 114.07 | 45.84 |
| Max. | 131.79 | 147.19 | 94.73 |

The one way analysis of variance (ANOVA) analysis also showed that the AE Means among different test types are significantly different from each other at 5% significance level.

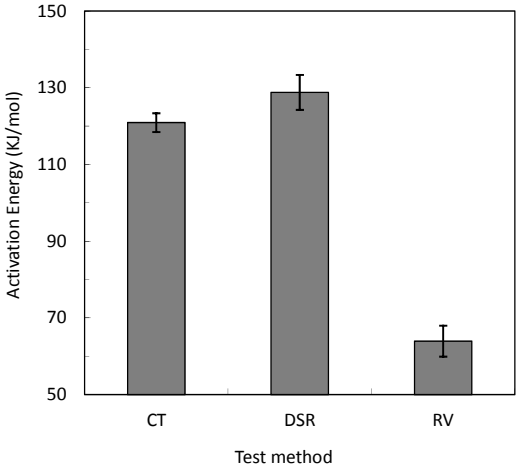
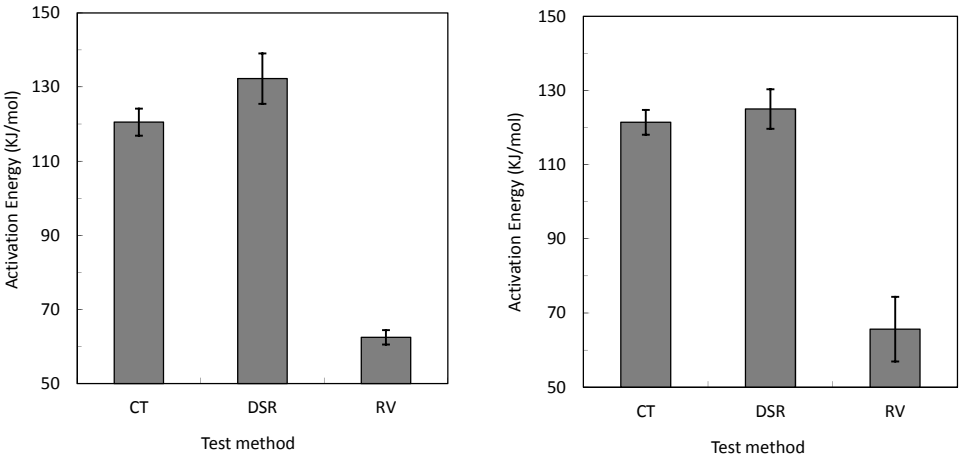



**Figure 3. Comparison of AE by different test methods**

In order to further investigate the reasons for the differences in AE Mean between different test methods, the binders were stratified according to type i.e., neat vs.

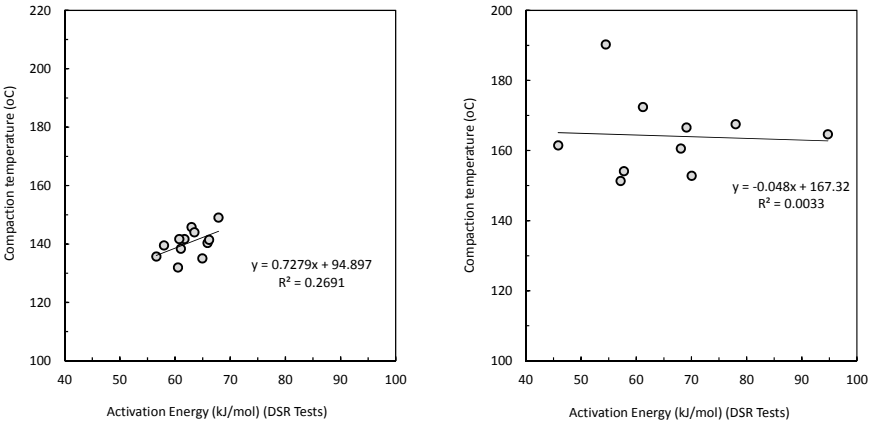
modified. The AE Mean comparisons were made again within each binder type. Figure 4 shows the results of those comparisons. Figure 4a shows that for neat binders, the AE Means are significantly different among all test methods while the difference only exists between RV and other two test types in case of modified binders.




(a) Neat binders (b) Modified binders

**Figure 4. Comparison of AE by binder and test type**

Because the AE values obtained from RV test methods showed the maximum variations for all binders and the same test is used to determine mixing and compaction temperatures of the binders, the correlation were investigated between AE and compaction temperatures for the binders. Figure 5 illustrates the relationships between AE from the RV test and compaction temperature obtained through the conventional viscosity criteria. A poor correlation can be observed between AE and compaction temperatures for the neat binders (see Figure 5a) while no relationship could be found for the modified binders (see Figure 5b). These results imply that AE may not be used to estimate mixing or compaction temperatures for the binders, especially for the modified ones.




(a) Neat binders (b) Modified binders

**Figure 5. Relationship between AE and compaction temperatures**

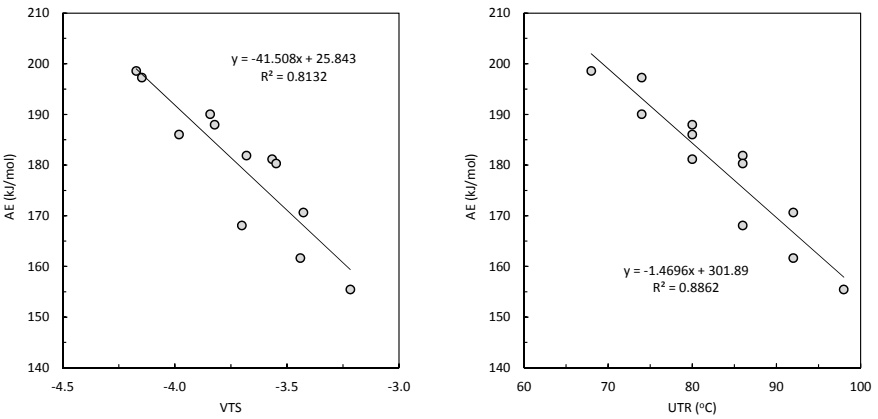
**Relation between activation energy of flow and performance grade.**

In order to investigate the relationship between AE and PG grades, the M-E PDG software (version 1.0) was used. A typical flexible pavement structure was analyzed by only changing the PG grade at Level 3. From the A-VTS information generated by the M-E PDG, the AE was calculated by using Equation (2). The pavement performance at the end of the 20 years analysis period in terms of distresses was predicted. Table 3 shows the details of PG grades, A-VTS parameters, predicted pavement performance, and UTR.

**Table 3. The M-E PDG results for different binder performance grades**

| No | PG | A | VTS | AE (kJ/mol) | IRI (in/mile) | HMA rut (in) | Total rut (in) | Long. crack. (ft/mile) | Fatigue crack. (%) | UTR |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 58-10 | 12.32 | -4.17 | 199 | 123.20 | 0.30 | 0.65 | 449.00 | 2.80 | 68 |
| 2 | 58-16 | 12.25 | -4.15 | 197 | 123.20 | 0.30 | 0.64 | 443.00 | 2.80 | 74 |
| 3 | 58-22 | 11.78 | -3.98 | 186 | 124.10 | 0.32 | 0.67 | 519.00 | 2.90 | 80 |
| 4 | 58-28 | 11.02 | -3.70 | 168 | 126.00 | 0.36 | 0.71 | 669.00 | 3.20 | 86 |
| 5 | 64-10 | 11.44 | -3.84 | 190 | 120.40 | 0.25 | 0.58 | 252.00 | 2.40 | 74 |
| 6 | 64-16 | 11.38 | -3.82 | 188 | 120.70 | 0.25 | 0.59 | 272.00 | 2.50 | 80 |
| 7 | 64-22 | 10.98 | -3.68 | 182 | 121.50 | 0.27 | 0.60 | 326.00 | 2.60 | 86 |
| 8 | 64-28 | 10.32 | -3.44 | 162 | 122.80 | 0.29 | 0.63 | 435.00 | 2.80 | 92 |
| 9 | 70-10 | 10.70 | -3.57 | 181 | 118.40 | 0.21 | 0.53 | 147.00 | 2.20 | 80 |
| 10 | 70-16 | 10.65 | -3.55 | 180 | 118.40 | 0.20 | 0.53 | 145.00 | 2.20 | 86 |
| 11 | 70-22 | 10.30 | -3.43 | 171 | 119.40 | 0.22 | 0.55 | 197.00 | 2.30 | 92 |
| 12 | 70-28 | 9.72 | -3.22 | 155 | 120.90 | 0.25 | 0.59 | 297.00 | 2.60 | 98 |

Figure 6 shows the relationships of VTS vs. AE and UTR vs. AE based on the M-E PDG results. The higher the AE, higher is the VTS, meaning that binder having a higher AE will be more temperature susceptible and vice versa (see Figure 6a). A negative correlation is observed with the AE and UTR (see Figure 6b), showing that lower the AE, higher will be the UTR or harder the binder or vice versa. This result contradicts with the previous findings in the published literature. It should be noted that the temperature range between 4.5 and 54.5 $^{o}$C is considered in the M-E PDG analysis for generating temperature-viscosity relationship.




(a) Temperature susceptibility (b) Temperature sensitivity

**Figure 6. Relationship between AE and VTS using the M-E PDG**

In order to investigate the impact of aging, the AE values were determined from the DSR test results in which RTFO and PAV aged binders were used. Figure 7 shows the relative comparison among all binders considered. In general, higher aging has increased the AE for the binders, indicating higher energy for flow is needed for the aged binders.

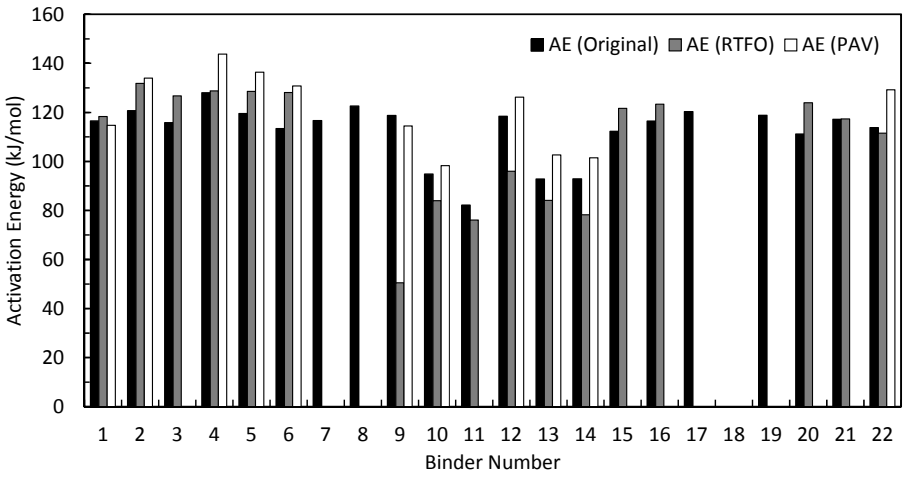
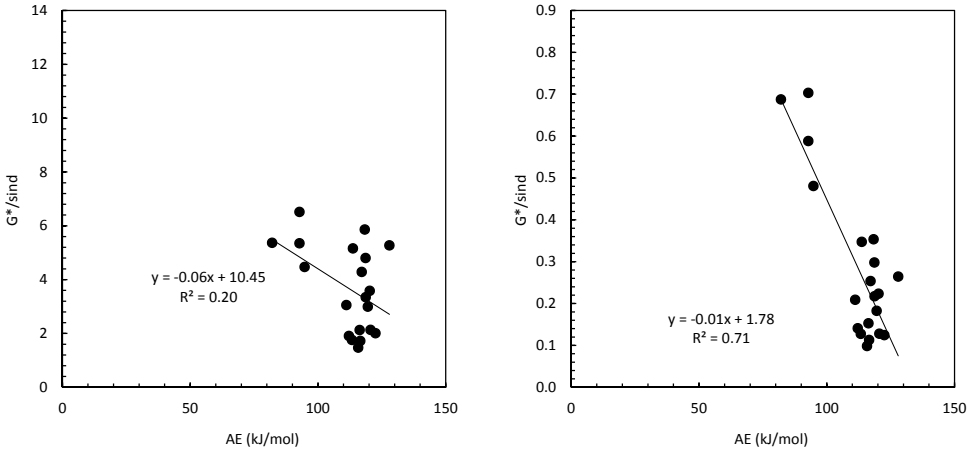



**Figure 7. AE comparison for all binders for different aging conditions (DSR)**

The temperature sweep data from DSR tests for all binders were utilized to investigate correlations between the AE and rheological parameters like $G^*/\sin\delta$ & $G^*\sin\delta$ at given temperatures. Figure 8 shows examples of such two relationships between AE and $G^*/\sin\delta$. For all binders, it was observed that at intermediate temperatures, the AE showed a poor association with $G^*/\sin\delta$ while at higher temperatures the AE correlates reasonably well with $G^*/\sin\delta$. It was observed that higher the AE, lower the $G^*/\sin\delta$, especially at higher temperatures. These results imply that the binders having higher AE are more susceptible to rutting. Similar trends were also observed for the $G^*\sin\delta$ parameter. The results apply to all original, RTFO and PAV aged binders considered in the study.




(a) 58 °C (b) 82 °C

**Figure 8. AE relationship with DSR parameters**

## DISCUSSION OF RESULTS

Asphalt binders behavior is significantly affected by temperature and rate of loading. The binder reacts to a load (shear or axial) in an elastics manner at low temperatures or high loading rate. It changes its response like a viscoelastic material at intermediate temperatures while at very high temperatures binders behave as a viscous material. While the binder grade (stiff vs. soft) impacts the range of temperature or loading rate, the general behavior of the materials remains the same. Three temperature ranges were used in this study depending on the test method: 25 to 135 $^{o}$C, 100 to 185 $^{o}$C, and 7 to 82 $^{o}$C for conventional, RV and DSR tests, respectively. It is expected that binders will behave as viscoelastic material within the temperature ranges used for conventional and DSR testing while those will exhibit viscous behavior for the RV testing. The activation energy (AE) concept was utilized in this study to characterize the temperature susceptibility of the binders; however, contradicting results were found, especially if the DSR temperature sweep and the M-E PDG are used. Setting aside the difference in various test methods, temperature under which the tests are performed was found to be one of the most significant factors defining the binder behavior. The AE determined from the M-E PDG and DSR data show similar trends for binders (i.e., lower AE for harder binders). It could be as a consequence of viscoelastic binder behavior within the test temperature range. In order to capture the viscous behavior of the binders, RV test data could be used to quantify the temperature susceptibility of binders.

## CONCLUSIONS

The temperature-viscosity relationships for twenty two neat and modified asphalt binders were evaluated using conventional and other binder tests such as penetration, kinematic and absolute viscosities, and ring and ball softening point; rotational viscometer (RV); and dynamic shear rheometer (DSR). The binders having similar penetration, viscosity and performance grade (PG) could show dissimilar rheological behaviors. Hence, there is a need to fully characterize the binder behavior to capture their rheology over a wide range of temperatures and load frequencies, especially if those are modified. The results of binder temperature susceptibility data show that the AE of the binders depends on the test methods, which employ different temperatures used to characterize the binders. Correlations between AE and G*/sin$\delta$ and G*sin$\delta$ parameters show that the binders with higher AE will be more viscous at higher temperatures. A negative correlation between AE and useful temperature range (UTR) of binders were shown by the M-EP DG. It is recommended that RV data could be used to characterize binders using activation energy concept. The AE concept did not show a good correlation with mixing and compaction temperatures of the binders, especially for the modified binders.

## ACKNOWLEDGEMENT

The authors acknowledge the Pakistan-United States Science and Technology Cooperative Program for funding this research.

## REFERENCE


Asphalt Institute. (2003). "Performance Graded Asphalt Binder Specification and Testing, Superpave Series No. 1 (SP-1)." Asphalt Institute, Lexington, KY.

Bahia, H., Hanson, D., Zeng, M., Zhai, H., Khatri, M., and Anderson, R. (2001). "NCHRP Report 459: Characterization of Modified Asphalt Binders in Superpave Mix Design." *Transportation Research Board*.

Bari, J., and Witczak, M. (2007). "New Predictive Models for Viscosity and Complex Shear Modulus of Asphalt Binders: For Use with Mechanistic-Empirical Pavement Design Guide." *Transportation Research Record: Journal of the Transportation Research Board*, 2001(-1), 9-19.

Bari, J., and Witczak, M. W. (2006). "Development of a new revised version of the Witczak E Predictive Model for hot mix asphalt mixtures." *0270-2932*, Association of Asphalt Paving Technologist, White Bear Lake, MN 55110, United States, Savannah, GA, United States.

Dongre, R., Myers, L., D'Angelo, J., Paugh, C., and Gudimettla, J. (2005). "Field evaluation of Witczak and Hirsch Models for predicting dynamic modulus of hot-mix asphalt." *Jounal of Association of Asphalt Paving Technologist*, 74, 381-442.

Haider, S. W., Baladi, G. Y., Aurangzeb, Q., and Akram, T. (2009). "Rheology of asphalt binders in Pakistan." *978 88 8202 029 3*, Politecnico Di Torino, Torino, Italy.

Jongepier, R., and Kuilman, B. (1969). "Characteristics of the Rheology of Bitumens." *Journal of Association of Asphalt Paving Technologist*, 38, 98-122.

Maze, M. (1996). "Viscosity of EVA Polymer-Modified Bitumens: Modeling."

Nam, K., Delgadillo, R., and Bahia, H. (2004). "Development of Guidelines for PG Binder Selection for Wisconsin." *WisDOT Highway Research Study 0092-01-01*.

Pellinen, T. K., Witczak, M. W., and Bonaquist, R. F. (2004). "Asphalt Mix Master Curve Construction using Sigmoidal Fitting Function with Non-linear Least Squares Optimization." *0895-0563*, American Society of Civil Engineers, Reston, VA 20191-4400, United States, New York, NY., United States.

Robson, G. R. (1969). "The Dynamic Mechanical Properties of Bitumen." *Journal of Association of Asphalt Paving Technologist*, 38, 123-139.

Salomon, D., and Zhai, H. "Asphalt Binder Flow Activation Energy and Its Significance for Compaction Effort." 12–14.

Salomon, D., and Zhai, H. (2002b). "Ranking Asphalt Binders by Activation Energy for Flow." *Journal of Applied Asphalt Binder Technology*, 2(2).

Zeng, M., Bahia, H. U., Zhai, H., Anderson, M. R., and Anderson, P. (2001). "Rheological Modeling of Modified Asphalt Binders and Mixtures." *Journal of Association of Asphalt Paving Technologists*, 70, 403-444.

# STUDY ON FREQUENCY DEPENDENCY OF SOL-GEL TRANSITION OF ASPHALT

A.Padmarekha & J.Murali Krishnan *

**Abstract**

One important factor related to the asphalt pavement performance is the transitory nature of asphalt. Heating asphalt from - 40 to 150 °C, one can observe the transition from glassy solid to viscoelastic solid, viscoelastic solid to viscoelastic fluid and finally to Newtonian fluid. This study focuses on capturing the viscoelastic solid-fluid transition. Specifically, this transition was captured here by the storage modulus (G') and loss modulus (G") cross-over when the sample is subjected to a temperature sweep with constant strain amplitude and frequency. Since this measure depends on frequency of testing, this study also quantifies the influence of frequency.

Two types of asphalt of the same grade but manufactured by different process (blending and air blowing) were tested at unaged, short-term aged and long-term aged conditions. The G'-G" cross-over point changes at low frequency and the frequency dependency vanishes above 40 Hz. This frequency independent G'-G" cross-over temperature can then be truly called as the viscoelastic transition temperature or sol-gel transition temperature. It was interesting to note that the aging conditions did not alter the frequency dependency. Also there was no significant difference in cross-over temperature of blended and air blown asphalt.

KEYWORDS: Air blown asphalt, Blended asphalt, Aging condition, Oscillatory shearing, Temperature sweep, Sol-gel transition, Frequency dependency

* Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, India.

E-mail: apadmarekha@gmail.com & jmk@iitm.ac.in

## 1 Introduction

Asphalt is a widely used material in pavement construction. Better understanding

of asphalt rheology will help in constructing crack and rut-free pavement. One important factor related to the asphalt pavement performance is the transitory nature of asphalt. Heating asphalt from -40 to 150 °C, one can observe the transition from glassy solid to viscoelastic solid, viscoelastic solid to viscoelastic fluid and finally to Newtonian fluid. Viscoelastic solid to viscoelastic fluid transition of asphalt is expected to occur in 15 to 60 °C regime. In a typical day, for most part of the time an asphalt pavement in India is subjected to temperature regimes in this range. The mechanical response characteristic of the pavement during traffic loading is also dependent on whether asphalt mixture exhibit viscoelastic solid-like behavior or viscoelastic fluid-like behavior. To better understand the stress-strain response of the asphalt pavement, it is hence necessary that one understands the transitory nature of asphalt.

Mechanical response of asphalt depends on the temperature during load application. For instance, Schweyer and Davis (1973) classified asphalt response at high temperature (above 60 °C) as Newtonian fluid, near transition temperature (between 0 and 60 °C) as viscoelastic and far-transition range (between glass transition temperature and 0 °C) as elastic. Storm et al. (1996) in their transition study on asphalt suggested that the asphaltene colloidal particles are well dispersed in the fluid (resin) in the 65-150 °C temperature range showing Newtonian behavior. As the temperature is reduced, the asphaltene begin to overlap and induces a Newtonian to viscoelastic fluid transition in 45-65 °C temperature range. A detailed review on asphalt transition was done by Krishnan and Rajagopal (2005).

Literature shows that as the temperature increases, the internal structure of asphalt changes. Pfeiffer and Saal (1940) viewed asphalt structure as a colloidal systems consisting of dispersions of micelles in viscous oil. The micelles and viscous oil have different thermo-mechanical behavior at any temperature and hence as the temperature is varied, one can expect more than one transition, one corresponding to micelles and the other corresponding to viscous oil. Differential scanning calorimetry (DSC) study of Planche et al. (1998) and Masson and Polomark (2001) showed that asphalt has more than one glass transition temperature. It is interesting to note that very little investigation has been carried out on capturing the transitory response of the asphalt. Before we continue, it is necessary to understand how one quantifies the transition temperature experimentally in the laboratory.

Nijenhuis and Winter (1989) and Li and Aoki (1997) defined viscoelastic fluid-like to viscoelastic solid-like transition as gelation or sol-gel transition. There are several ways of defining the sol-gel transition point. One of the frequently used methods for characterizing sol-gel transition is through oscillatory shear experiments. In these experiments, thematerial is simultaneously subjected to temperature sweep with a constant strain amplitude and frequency. The evolution of storage modulus (G') and loss modulus (G") as well as the phase angle are monitored as the material is sheared. Initially when the temperature is high, the loss modulus values are larger than the storage modulus. However as

the temperature reduces, the storage modulus increases and at a specific temperature where the storage and loss modulus values are equal is defined as the sol-gel transition. While this idea of capturing the transition point is very appealing, one major disadvantage is the influence of frequency on the transition point. As the material is tested at increasing frequencies, it is seen that the cross-over point also increases. Interesting discussion related to the use of G'-G" cross-over as the sol-gel transition point and the influence of frequency on this can be found in Winter and Mours (1997).

The main objective of this work is to investigate the influence of frequency on the sol-gel transition point and see how this influence changes as the material processing method and/or aging conditions are changed.

# 2 Experimental Investigation

## 2.1 Material

Air blown and blended asphalt obtained from Arab Mix crude were tested. They met the viscosity grade 30 (VG30) standards of Indian Standards IS73 (2006). In air blowing process, vacuum residue is subjected to blowing operation to obtain the required grade of asphalt. In blending process, petroleum pitch obtained from solvent de-asphalting process is blended with heavy extract to obtain the required grade of asphalt. Blending proportion depends on the grade of asphalt required. To obtain VG30 grade of asphalt, 91-92% of pitch is blended with 8-9% of heavy extract. In air blowing process high penetration value material is converted to low penetration value material, whereas in blending process low penetration value pitch is converted to high penetration value asphalt. Since these two material goes through completely different processing method, they are expected to show different thermo-rheological behavior. This is true even if both the materials were processed from the same crude source.

In this investigation, all the materials were tested in unaged, short-term aged and long-term aged conditions. Short term aging was carried out in rolling thin film oven as per ASTM standard D2872-04. Long term aging was carried out in pressure aging vessel as per ASTM standard D6521-08.

## 2.2 Experimental Protocol

Investigations were carried out using Anton-Paar dynamic shear rheometer, model MCR301. The materials were subjected to oscillatory shear superimposed over temperature sweep. Cone and plate geometry of 25 mm diameter and gap setting of 0.0475 mm were used. All the samples were tested in accordance with the ASTM standard D7175-2008. The samples were placed in between cone and plate assembly at 60 °C and equilibrated for 10 minutes. Preliminary test trials were conducted to find the linear viscoelastic limit. An amplitude sweep test was conducted at 20 °C and strain amplitude of 0.01%

was finally chosen. The material was sheared in the oscillatory mode while subjecting the material simultaneously to a temperature sweep, lowering temperature from 60 to 20 °C at a rate of 10 °C per minute. Experiments were conducted for 5 to 50 Hz frequency at the increment of 5 Hz and data was recorded as the function of temperature. When the material is subjected to simultaneous oscillatory shear and temperature decrease, the torque increased. However, in some cases, especially at high frequency and low temperature, the resultant torque decreased due to de-bonding between the plate and the material. These tests were treated as failure trials and the experiments were repeated with a fresh sample.

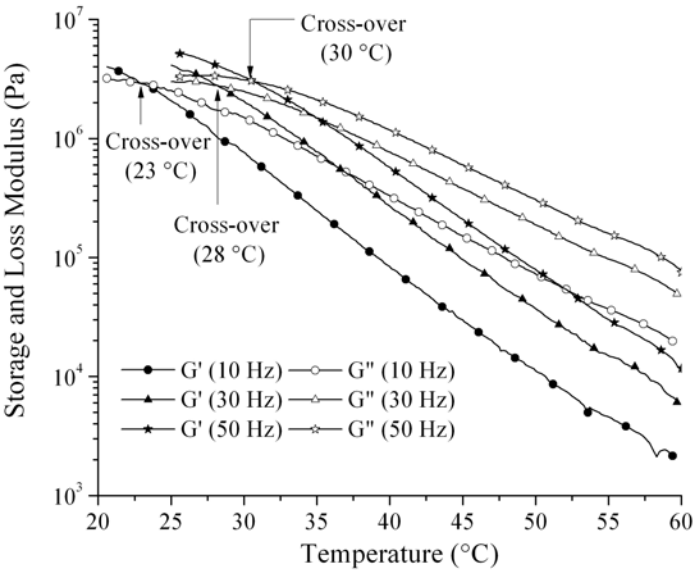
## 3Resultsand Discussion

Figure 1 shows the storage and loss modulus plot of blended and air blown unaged asphalt at different frequencies. At 60 °C, loss modulus is greater than the storage modulus for all frequencies. As the temperature decreases, storage and loss modulus increases. G' increases at a higher rate when compared to G" and the point were the G' crosses G" is taken as the sol-gel transition temperature. It is clear from figure 1 that the cross-over temperature changes with the frequency of testing. The cross-over of blended asphalt occurred at 22 °C for 10 Hz frequency, at 28 °C for 30 Hz frequency and at 30 °C for 50 Hz frequency. There is no appreciable difference in cross-over temperature between air blown and blended asphalt.

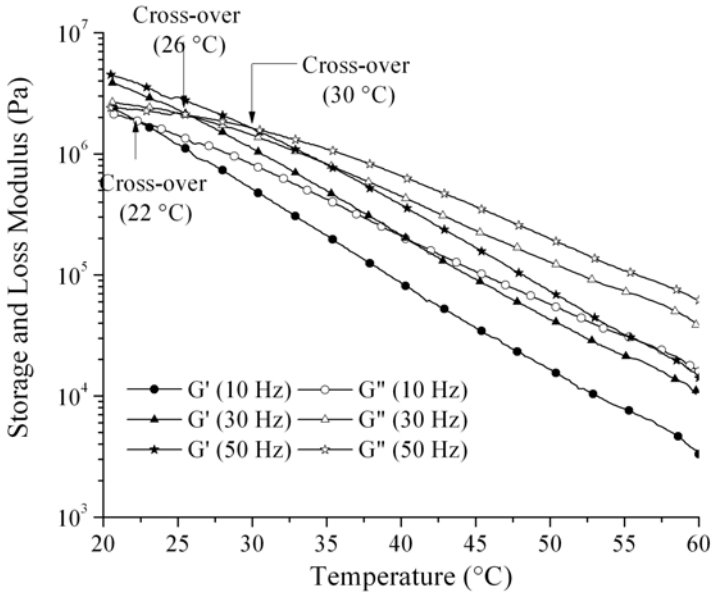
Figure 2 compares G'-G" values for all aging conditions at 50 Hz frequency for blended and air blown asphalt. As the material ages cross-over occurs at higher frequency. G' value for aged sample is always higher than that of unaged sample. This shows that the elastic component of the response dominates as the material ages and hence results in sol-gel transition at higher temperature.

Table 1 shows the cross-over temperature of blended and air blown asphalt for all aging conditions. The influence of frequency on cross-over temperature is more at lower frequencies. G'-G" cross-over reaches a constant value above 40 Hz.

The loss modulus plot in figure 3 clearly shows that the rate of change of G" increase with decrease in temperature for all frequencies reached asymptotically a constant value. However G' shows steady increase with decrease in temperature. Also, G' curves are parallel at higher temperature. As the temperature decreases, G' at higher frequency also attains asymptotically a constant value. Hence at higher frequency and lower temperature, G' and G" reaches a constant value resulting in frequency independent cross-over temperature.

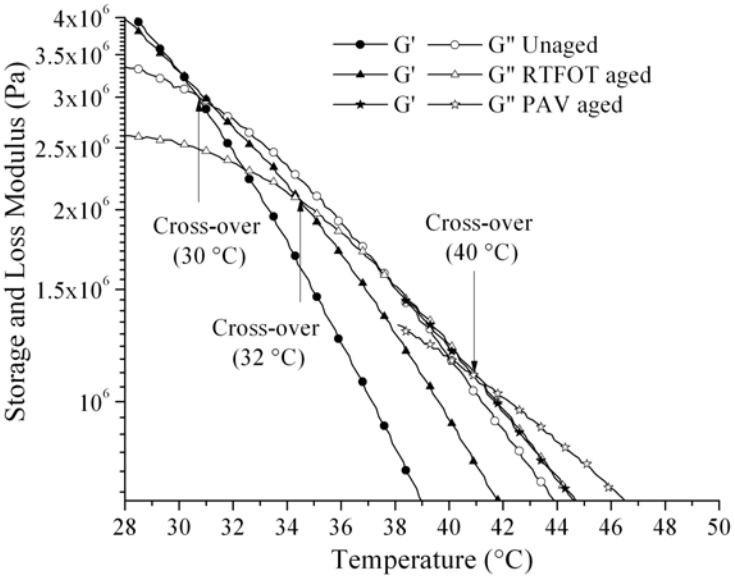



(a) Blended Asphalt

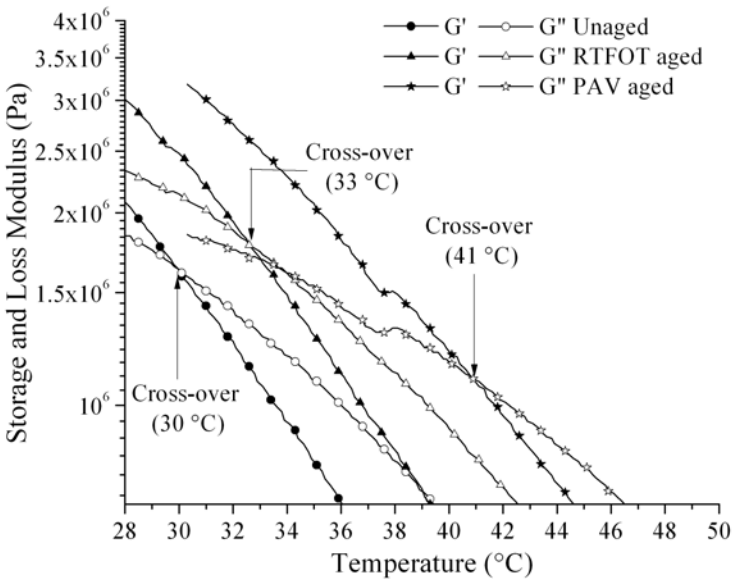



(b) Air Blown Asphalt

Figure 1: Storage and Loss Modulus of Unaged Asphalt at 10, 30 and 50 Hz Frequency

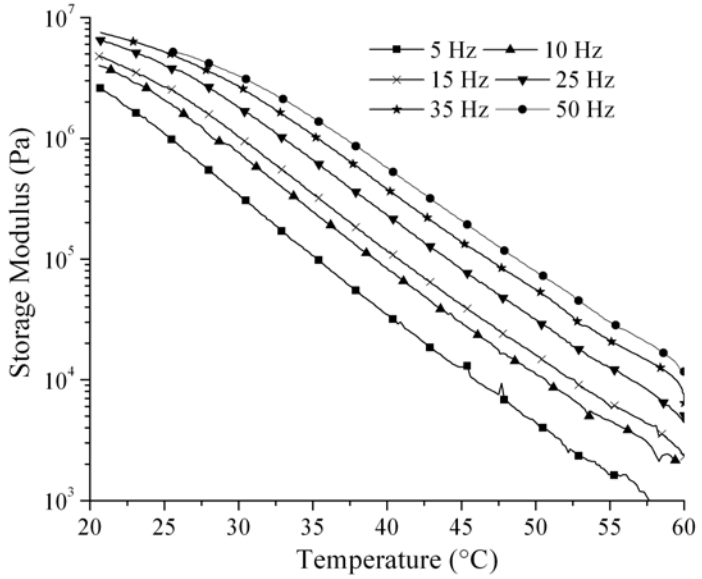



(a) Blended Asphalt

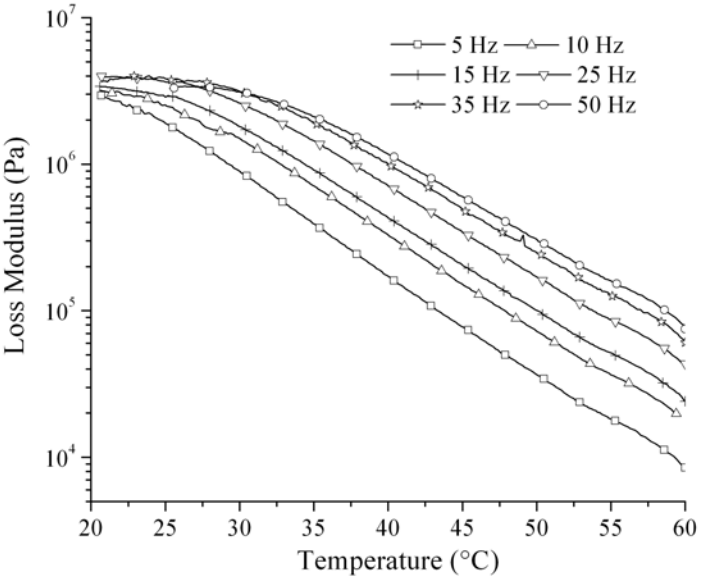



(b) Air Blown Asphalt

Figure 2: Storage and Loss Modulus of Unaged, RTFOT Aged and PAV Aged Samples at 50 Hz Frequency




(a) Storage Modulus of Blended Asphalt




(b) Loss Modulus of Blended Asphalt

Figure 3: Trend in Storage and Loss M odulus of Blended Unaged Asphalt at Different Frequencies

Table 1: Cross-Over Temperature of Samples Tested in °C

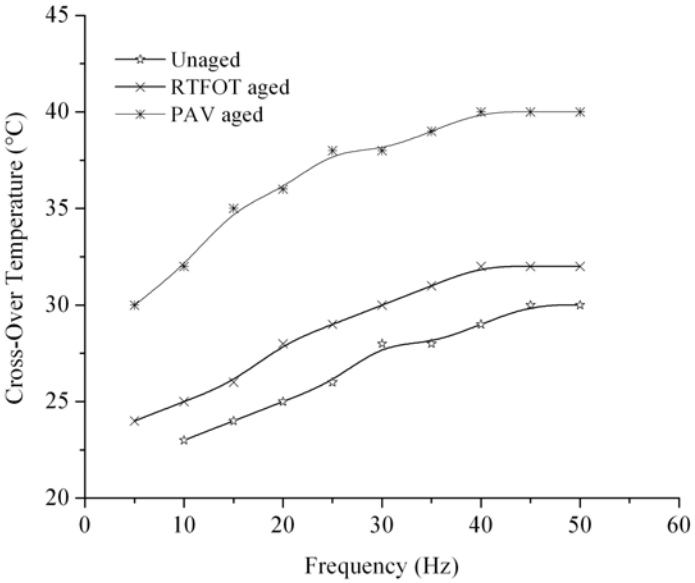
| Frequency (Hz) | Blended Asphalt | | | Air Blown Asphalt | | |
|---|---|---|---|---|---|---|
| | Unaged | RTFOT | PAV | Unaged | RTFOT | PAV |
| 5 | 24 | 30 | 21 | 24 | 33 | |
| 10 | 23 | 25 | 32 | 22 | 25 | 35 |
| 15 | 24 | 26 | 35 | 23 | 26 | 37 |
| 20 | 25 | 28 | 36 | 24 | 28 | 38 |
| 25 | 26 | 29 | 38 | 25 | 28 | 39 |
| 30 | 28 | 30 | 38 | 26 | 30 | 40 |
| 35 | 28 | 31 | 39 | 28 | 32 | 40 |
| 40 | 29 | 33 | 40 | 30 | 33 | 41 |
| 45 | 30 | 33 | 40 | 30 | 33 | 41 |
| 50 | 30 | 33 | 40 | 30 | 33 | 41 |
| Standard Deviation | 2.6 | 3.4 | 3.5 | 3.5 | 3.5 | 2.8 |
| Coefficient of Variation (%) | 9.3 | 11.5 | 9.3 | 13.5 | 12.0 | 7.0 |

Figure 4 shows the trend in change of cross-over temperature as a function of frequency of blended and air blown asphalt. Trend lines for unaged, RTFOT aged and PAV aged samples in both the figure are parallel to each other. For the samples tested here, and based on the idea that G'-G" signifies sol-gel transition, this shows the frequency dependence does not change as the material is aged.

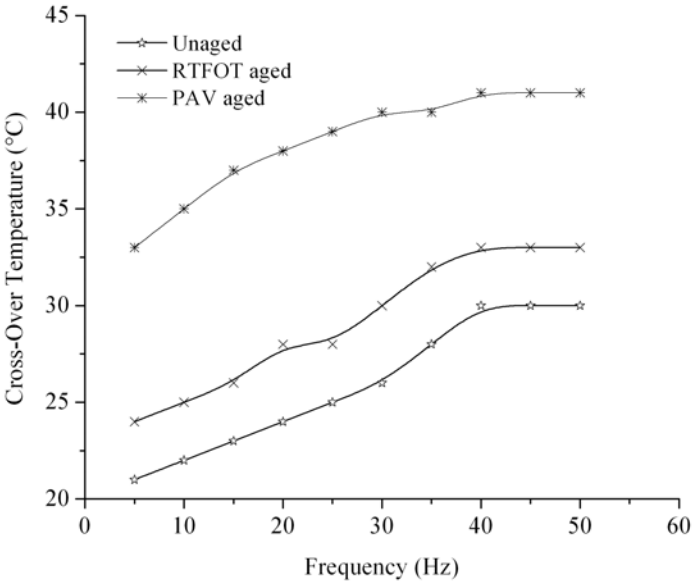
## 4 Conclusion

In this investigation, the viscoelastic solid-fluid transition was captured using G' and G" cross-over when the sample was subjected to a temperature sweep with constant strain amplitude and frequency. At higher frequencies above 40 Hz, the frequency dependence of sol-gel transition diminished. It was also interesting to note that the aging conditions did not alter this frequency dependence.

This study is one of the first attempts in capturing the asphalt transition using oscillatory test data. A detailed investigation on sol-gel transition considering phase lag, torque normal force during oscillatory shear are currently being conducted. Also this study captures the sol-gel transition of binder alone. A detailed study is required to check whether binder transition is the representation of the asphalt mix transition.




(a) Blended Asphalt




(b) Air Blown Asphalt

Figure 4: Cross-over Temperatures of Blended and Air Blown Asphalt

# References

IS73 (2006). “Paving bitumen - specification (third revision).” The Bureau of Indian Standards, New Delhi, India.

Krishnan, J. M. and Rajagopal, K. (2005). “On the mechanical behavior of asphalt.” Mechanics of Materials, 37, 1085–1100.

Li, L. and Aoki, Y. (1997). “Rheological images of poly(vinyl chloride) gels. 1. the dependence of sol-gel transition on concentation.” Macromolecules, 30 (25), 7835–7841.

Masson, J. F. and Polomark, G. M. (2001). “Bitumen microstructure by modulated differential scanning calorimetry.” Thermochimica Acta, 374, 105–114.

Nijenhuis, K. T. and Winter, H. H. (1989). “Mechanical properties at the gel point of a crystallizing poly(vinyl chloride) solution.” Macromolecules, 22 (1), 411–414.

Pfeiffer, J. P. and Saal, R. N. J. (1940). “Asphaltic bitumen as colloid system.” The Journal of Physical Chemistry, 44, 139–149.

Planche, J. P., Claudy, P. M., Letoffe, J. M., and Martina, D. (1998). “Using thermal analysis methods to better understand asphalt rheology.” Thermochimica Acta, 324 (1-2), 223–227.

Schweyer, H. E. and Davis, R. L. (1973). “Asphalt rheology in the near-transition temperature range.” Highway Research Record, 468, 1–15.

Storm, D. A., Barresi, R. J., and Sheu, E. Y. (1996). “Development of solid properties and thermochemistry of asphalt binders in the 25-65°C temperature range.” Energy & Fuels, 10 (3), 855–864.

Winter, H. H. and Mours, M. (1997). “Rheology of polymers near liquid-solid transitions.” Advances in Polymer Science, 134, 165–234.

# DETECTION OF POLYMERS IN ASPHALT BINDERS

Martin McCann, Ph.D., P.E., 1; Joseph F. Rovani, 2; Kenneth P. Thomas, Ph.D. 3

1 U.S. Forest Service, Region 8 – Engineering, 1720 Peachtree Road, NW, Atlanta, Georgia, 30309, U.S.A., 404.347.2459, 404.347.2399 fax, mmccann@fs.fed.us; 2 Western Research Institute, 365 North Ninth Street, Laramie, Wyoming 82072-3380, U.S.A., 307.721.2275, 307.721.2300 fax, rovani@uwyo.edu; 3 Western Research Institute, 365 North Ninth Street, Laramie, Wyoming 82072-3380, U.S.A., 307.721.2326, 307.721.2300 fax, kpthomas@uwyo.edu

**ABSTRACT**

Superpave™ performance-graded asphalt binder technology continues to advance with some producers focusing on the use of additives to enhance the performance of asphalt binders. A common practice is to modify an asphalt binder with a polymer to extend the high- and low-temperature range. However, agencies require a method to verify the presence of the polymer. The method that is presented serves to demonstrate that gel-permeation chromatographic (GPC) can be used to detect the presence of polymers in asphalt binders.

Typically, the polymers used in asphalt binders have a number average molecular weight ranging from 300 to 600 times larger than the number average molecular weight of asphalt binders. The GPC technique separates the polymer from the asphalt binder as a function of molecular size.

The analysis of eight unknown samples, some of which were reported to contain a polymer, was conducted. The GPC results demonstrated that three of the asphalt binders were polymer modified while the remaining five asphalt binders were not modified. After the results were reported, the actual composition of the asphalt binders was revealed and found to be in agreement. Thus, GPC can be used as a simple screening test to identify when asphalt binder has been modified with a polymer.

**KEYWORDS**

Gel-Permeation Chromatography, Size-Exclusion Chromatography, Polymer, Asphalt Binder, Molecular Weight, Molecular Size

## INTRODUCTION

The use of polymer-modified asphalt binders has dramatically increased since the advent of the Strategic Highway Research Program (SHRP) and the adoption of the Superpave™ performance grade (PG) asphalt binder specifications. In combination with certain chemical modifiers, extending agents, or cross-linking agents, i.e. polymers, the high- and low-temperature properties of unmodified asphalt binders can be extended [Norton 2003]. Consequently, users of polymer-modified asphalt binders for the construction of asphalt pavements have noted improved performance. Because of the demonstrated improved performance of polymer-modified asphalt pavements, state agencies and suppliers have embraced the use of polymer-modified asphalt binders.

However, agencies require a method to verify the presence of the different types of polymers in the mix. An infrared method, AASHTO T 302, is currently available that can identify the presence and the concentration of styrene-butadiene polymers (SBR, SB, and SBS) in asphalt binders. However, additional methods are necessary that can be used to determine the chemical nature and concentration of the other polymers, such as crumb rubber and Elvaloy®, which are also used to modify asphalt binders.

This paper discusses the application of gel-permeation chromatography (GPC), which is synonymous with size-exclusion chromatography (SEC), for the detection of polymers in asphalt binders. In addition, the method is tentatively validated using eight asphalt binders obtained from the Federal Highway Administration. The method can be easily conducted in agency laboratories that have suitably trained personnel and the appropriate equipment.

## BACKGROUND

A polymer is characterized by the presence of numerous repeating units and a very high molecular weight. For example, in the simplest case the repeating unit in polyethylene is ethylene; while in a more complex case, such as styrene-butadiene, the repeating unit is styrene-butadiene. Because of the polymerization process, the molecular weights of the individual molecules that comprise the polymer and their relative distribution vary. Consequently, the molecular weight of polymers is typically reported as either the number average or the weight average molecular weight. As an example, the number average molecular weight of SBS used in asphalt pavements is about 400,000. For comparison, the range of the number average molecular weights determined for the SHRP core asphalt binders are from 700 to 1300 [Branthaver et al. 1993]. Thus, the number average molecular weight of polymers is significantly larger than that determined for typical asphalt binders. Because of this disparity in the number average molecular weight between the polymer and the asphalt binder, it was apparent that GPC could be used to detect the presence of a polymer in an asphalt binder.

Chromatography is an analytical method that is used to separate a mixture of compounds on the basis of charge, affinity, solubility, and molecular size. This is accomplished by dissolving the mixture in a solvent and passing the dissolved mixture called the solute through a column containing spherical polyacrylamide beads. For the results from a GPC experiment to accurately reflect the true

molecular distribution of the mixture, the molecules must be completely separated and soluble in the solvent. The selection of the optimal pore diameter of the polyacrylamide beads is based on an estimate of the molecular weights of the compounds in mixture to be separated. As the solute moves through the column, the smaller molecules pass through the pores in the beads while the larger molecules travel around the beads. The smaller molecules which constitute the asphalt binder have a more convoluted path while the larger molecules have a shorter, more direct route through the column. Therefore, larger molecules that spend less time in the column are eluted first and have a shorter retention time in comparison to the smaller molecules which have a longer retention time. The molecular size distribution of the solute can be converted into a molecular weight distribution by comparing the GPC results to the retention time of known standards. ASTM standard methods that utilize GPC are available for determining the molecular weight distribution and molecular weight averages of polystyrene (D 5296), polyolefins (D 6474), and hydrocarbon and terpene resins (D 6579). However, there is no standard method that utilizes GPC for the determination of polymers in asphalt binders.

The separation of asphalt binder into fractions that are based on molecular size using GPC dates back many years. For example, Altgelt [1965] separated three asphalt binders and their asphaltenes and maltenes and two air-blown asphalt binders into various fractions. Based on the molecular weight of the fractions, he demonstrated that separation based on molecular size had been achieved, that the molecular weight of asphalt binders vary, and solvent precipitation is less effective than GPC for separating asphalt binders by molecular size. Subsequently, numerous researchers have used the technique to characterize asphalt binders to gain insight into asphalt binder behavior [Brule 1979; Donaldson et al. 1988; Jennings et al. 1992; Herrington and Ball 1996; Huang and Bertholf 1997; Siddiqui and Ali 1999], relate the GPC results to other properties [Huynh et al. 1978; Garrick and Wood 1986; Price and Burati 1989; Garrick and Biskur 1990; Kim et al. 1993], relate the GPC results to field performance [Jennings and Pribanic 1989], and study the impact of polymer modification of asphalt binders using GPC [Wei et al. 1996; Lee et al. 1997; Asi et al. 1997; Negulescu et al. 2006]. The preceding citations are not intended to be all inclusive but rather to give the reader some insight into the capabilities of the technique as it pertains to the study of asphalt binders.

## OBJECTIVES

This paper presents the analysis of eight asphalt binders using GPC to determine which of the asphalt binders were modified. The experimental objectives were:

a) Can GPC be used to separate a polymer from the asphalt binder?

b) Do the data obtained conclusively distinguish between unmodified and polymer-modified asphalt binders?

## EXPERIMENTAL DESIGN

The eight asphalt binders used in this study had previously been used for a study by the West Coast User/Producer Group. Upon completion of that study, the binders were delivered to the Federal Highway Administration (FHWA). Subsequently, the FHWA requested Western Research Institute (WRI) to evaluate the use of the GPC to determine the presence or absence of a polymer in the eight asphalt binders. At the beginning of the study, the asphalts binders were referenced by a tracking number one through eight, and the grade of asphalt binder was unknown to the WRI staff performing the GPC analysis. After the analysis was complete, a determination was made of whether the asphalt binder was modified. The grade of asphalt binders and whether they were modified was then revealed to WRI researchers by FHWA.

The GPC column for this experiment was a Phenogel 5 $\mu$m Linear(2) (300 x 7.80 mm) obtained from Phenomenex. The column is capable of separating molecules in the size range from 200 to greater than 600,000 Daltons. The isocratic solvent delivery system consisted of a Waters Corporation model 515 high performance liquid chromatography pump. The operating conditions for the pump were set at a flow rate of 1.0 milliliter per minute and a temperature of 40°C. Since several samples were run sequentially, a Waters Corporation model 717 HPLC auto-sampler was used. For the column used in this experiment, Table 1 illustrates the retention times obtained for the polystyrene standards kit (Waters Corporation). The kit was composed of eight different polymers ranging in molecular weight from 436 to 629,000. In addition, naphthalene, having molecular weight equal to 128, was added to the calibration standard. For the experimental configuration used, polymers, typically having molecular weights greater than 50,000, will have a retention time less than 9.2 minutes.

**TABLE 1 Retention time for calibration standards of different molecular weights**

| Calibration Standard Sample | Retention Time minutes | Molecular Weight |
|---|---|---|
| 1 | 7.5 | 629,000 |
| 2 | 8.1 | 210,000 |
| 3 | 8.7 | 70,600 |
| 4 | 9.2 | 28,600 |
| 5 | 9.7 | 10,900 |
| 6 | 10.2 | 3,000 |
| 7 | 10.6 | 1,300 |
| 8 | 11.0 | 436 |
| 9 | 12.0 | 128 |

Assuming the molecular size of the SHRP asphalt binders are representative of the experimental asphalt binders, the retention time will be 10.6 minutes and greater. A refractive index detector, Waters Corporation model 2414, was used to detect the presence of the eluent. The retention time data was electronically collected in increments of 0.003 of a second.

The experimental sequence was conducted as follows. The asphalt binder was dissolved in toluene, filtered, and the concentration of the asphalt binder adjusted to a concentration of 0.2 percent. Because of the low concentration of polymers in some of the asphalt binders, a conclusive determination of polymer modification was not possible at the 0.2 percent concentration. Therefore, a second set of tests was performed with a concentration of 7.0 percent asphalt binder in toluene. The GPC software was then used to calculate the concentration and the molecular weight distribution of the solute using the naphthalene and polystyrene standards as calibrants. Typically, eight samples at two concentrations and the standards can be analyzed in an 8-hour day, with the analysis time per sample being 15 minutes.

**RESULTS**

The GPC data for asphalt binder 1 is shown in Figure 1. Initial testing was done at a concentration of 0.2 percent asphalt binder in toluene. The initial molecules to elute have a molecular weight of approximately 10,900. As the retention time increases, the associated smaller molecules continue to pass through the refractive index detector. The spike in millivolts can be attributed to the increase in concentration of asphalt binder molecules and the associated decrease in molecular weight of the asphalt binder molecules. The peak in the data at approximately 10.8 minutes can be associated with asphalt binder molecules having an average molecular weight of 800. This average molecular weight is typical of SHRP asphalt binders. The decrease in millivolts from 10.8 minutes to 12.5 minutes demonstrates the decrease in molecular weight of the asphalt binder and a decrease in the concentration of molecules passing through the detector. In addition, the area under the plotted curve represents the volume of molecules contained in the sample. The evidence demonstrates no molecules larger than 10,900 were present in the sample, and therefore, asphalt binder 1 was determined to be an unmodified binder. Asphalt binder 1 was an unmodified binder and the grade was PG 64-22.

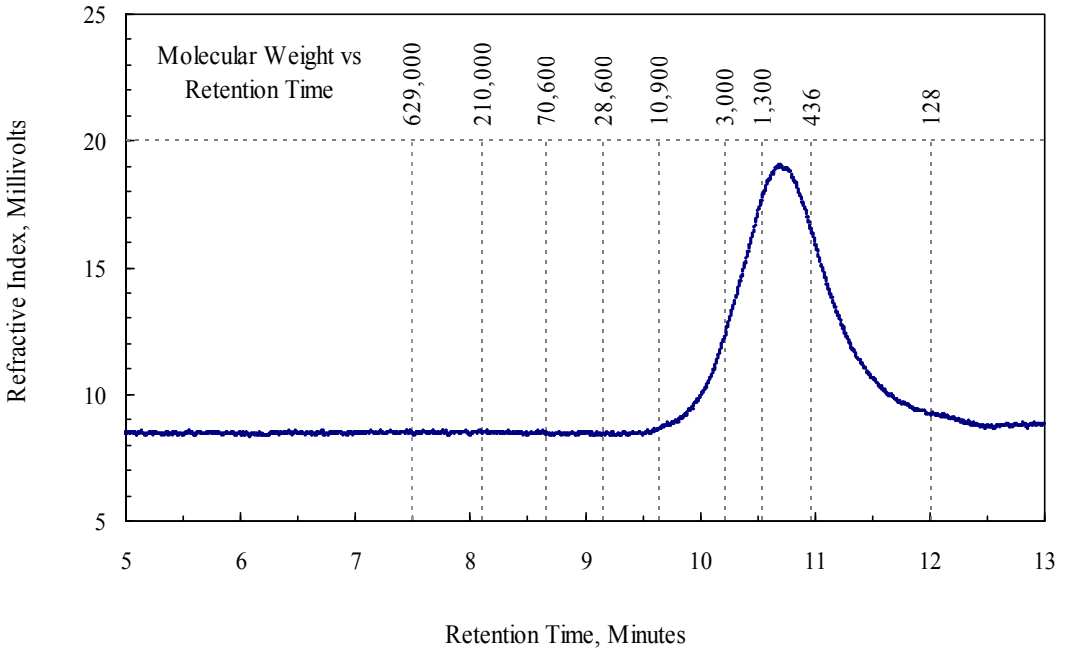



**FIGURE 1 Gel-permeation chromatographic results for PG 64-22 unmodified asphalt binder at a concentration of 0.2 percent in toluene.**

Figure 2a shows the test results for asphalt binder 2 at a 0.2 percent concentration in toluene. The data shows a small rise beginning at a retention time of 7.5 minutes, peaking at approximately 8.1, and dropping off at 9 minutes. A peak at retention time 8.1 is representative of a polymer having an average molecular weight of about 210,000. However, the low concentration of asphalt binder in toluene did not allow for a conclusive determination of the presence of polymer in this asphalt binder. Therefore, the asphalt binder was analyzed at a concentration of 7.0 percent asphalt binder in toluene. The results are shown in Figure 2b. At the higher concentration, there is a very pronounced peak in the data that can be associated with a polymer having an average molecular weight of 210,000. At the higher concentration, the evidence is conclusive that a polymer is present in asphalt binder 2. Asphalt binder 2 was polymer modified and was graded by the supplier as an AC-20P.

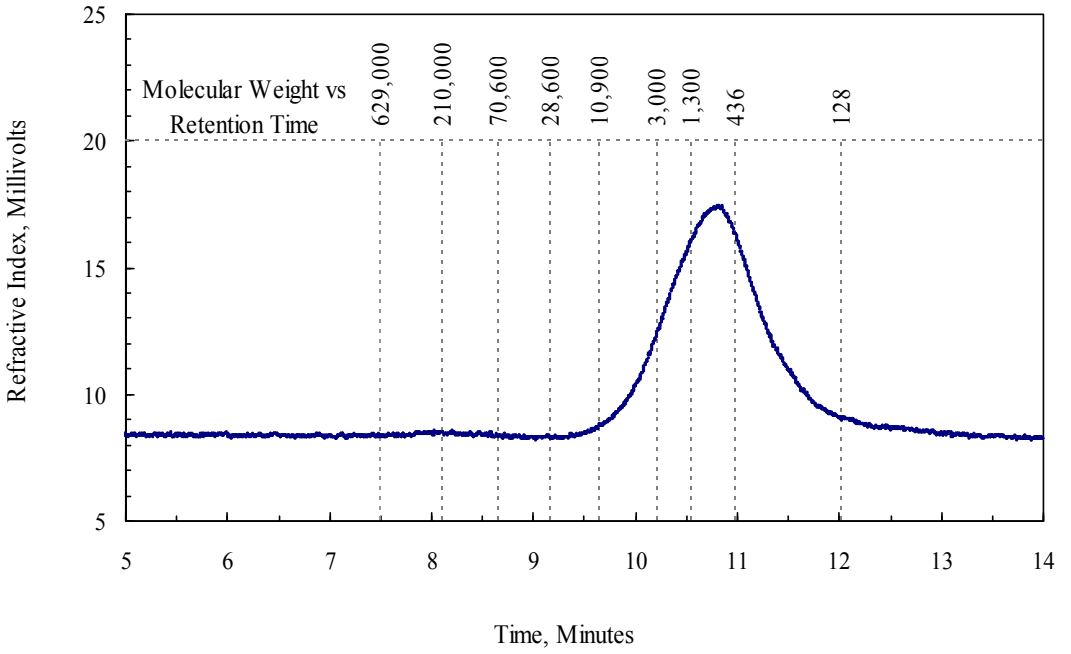



**FIGURE 2a Gel-permeation chromatographic results for AC-20P modified asphalt binder at a concentration of 0.2 percent in toluene (normal view).**

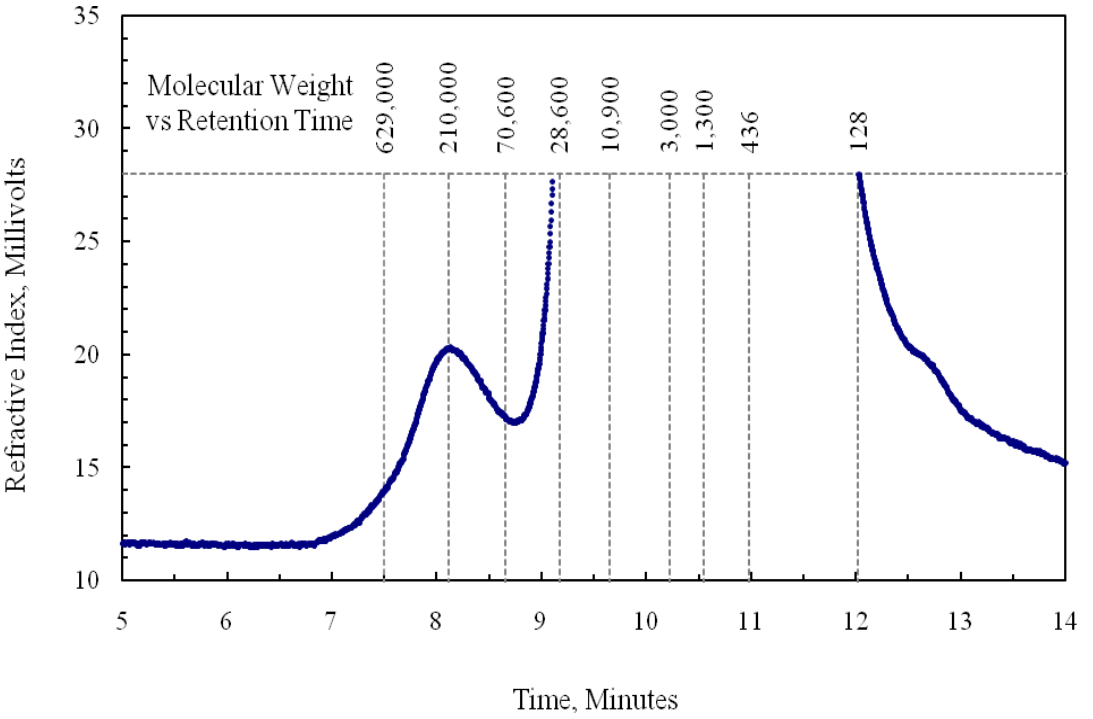



**FIGURE 2b Gel-permeation chromatographic results for AC-20P modified asphalt binder at a concentration of 7.0 percent in toluene (expanded view).**

In Figure 2a, the first sign of asphalt binder molecules did not begin to appear until after a retention time of 9.3 minutes. For the experimental configuration used

here, retention times in excess of 9.2 minutes are representative of molecular weights less than 28,600 and are characteristic of asphalt binders. In Figure 2c, the increase in refractive index at 8.8 minutes is representative of molecules having a molecular weight in excess of 50,000. Based on the data presented in Figure 2c, it should not be construed that the binder molecules grew in size, but rather is due to the occurrence of molecular association of polar molecules at the higher concentration. Thus, some of the asphalt binder molecules grouped together and traveled through the column as a larger associated molecule.

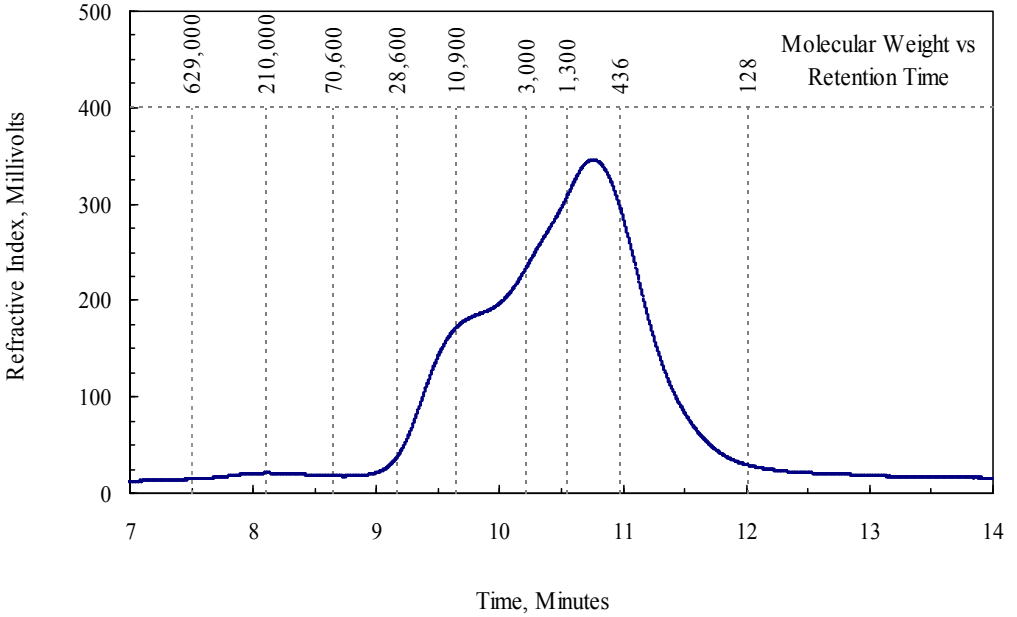



**FIGURE 2c Gel-permeation chromatographic results for AC-20P modified asphalt binder at a concentration of 7.0 percent in toluene (normal view).**

Asphalt binders 3, 4, and 5 were determined to be unmodified, PG 52-28, 64-10, and 76-16, respectively. Even at concentrations of 7.0 percent, no evidence of molecules having an average molecular weight larger than about 30,000 was observed for asphalt binders 3 and 4. Asphalt binder 5 was concluded to be unmodified after comparison of the appearance of the two chromatograms from the test results at the two concentrations of 7.0 and 0.2 percent. Asphalt binder 6, shown in Figure 3 at 7.0 percent concentration, is concluded to be a polymer-modified asphalt binder. The distinct rise in the data beginning at a retention time of about 7.8 minutes is indicative of a large polymer having an average molecular weight in excess of 210,000. At the end of the project it was revealed that this asphalt binder was polymer modified and had a PG of 70-28.

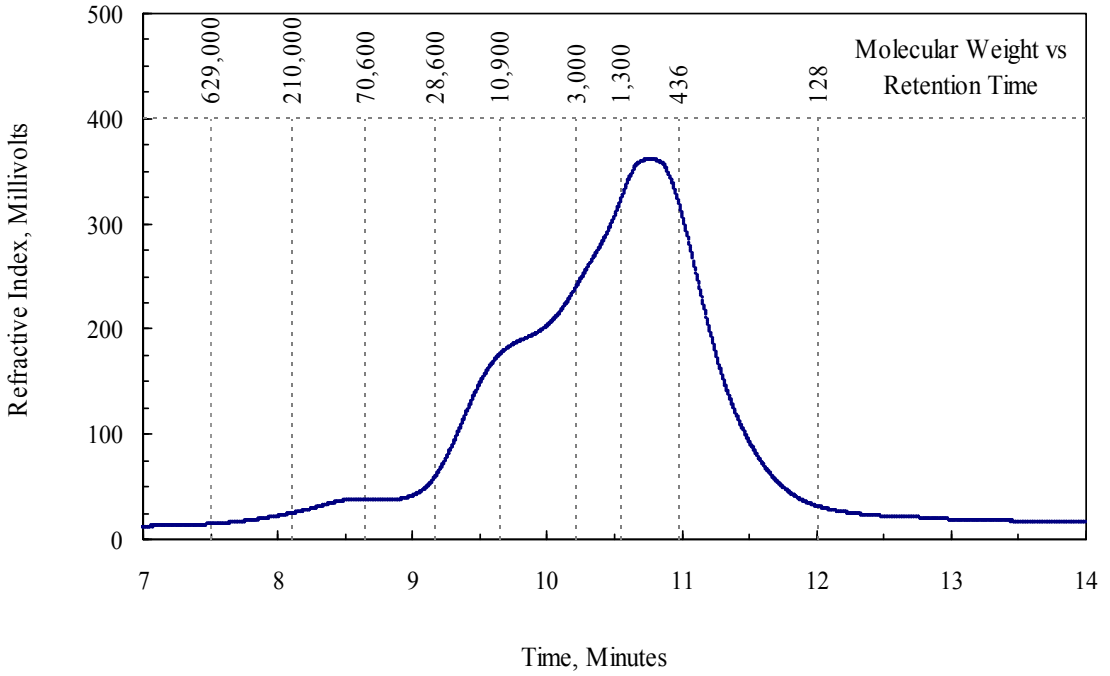



**FIGURE 3 Gel-permeation chromatographic results for PG 70-28 asphalt binder at a concentration of 7.0 percent in toluene.**

The GPC data for asphalt binder 7 is presented in Figure 4 at a concentration of 7.0 percent. The rise in the data beginning slightly before a retention time of 8 minutes is indicative of a polymer, but the appearance of the peak is not the same as was detected for asphalt binders 2 and 6. Specifically, the slope on the ascending side of the data is more gradual in comparison to the other two polymer-modified asphalt binders. A plot of asphalt binder 7 at a concentration of 0.2 percent did not reduce the slight but positive slope in the data for a retention time of 8 to 9 minutes. Therefore, this asphalt binder was concluded to have been modified with a variety of polymers having molecular weights ranging from greater than 210,000 down to that of the unassociated asphalt binder itself. After the project was completed, it was revealed that asphalt binder 7 actually consisted of a blend of polymers and other additives. The initial large molecules that were detected, at a retention time of 7.7 minutes, were styrene-butadiene-styrene (SBS) and ethyl-vinyl-acetate (EVA) polymers. In addition, a variety of other compounds were also present in this asphalt binder; a flux consisting of recycled motor oil, recycled polyethylenes, liquid rubber, and a cross linker. The actual PG of the asphalt binder was 64-34.

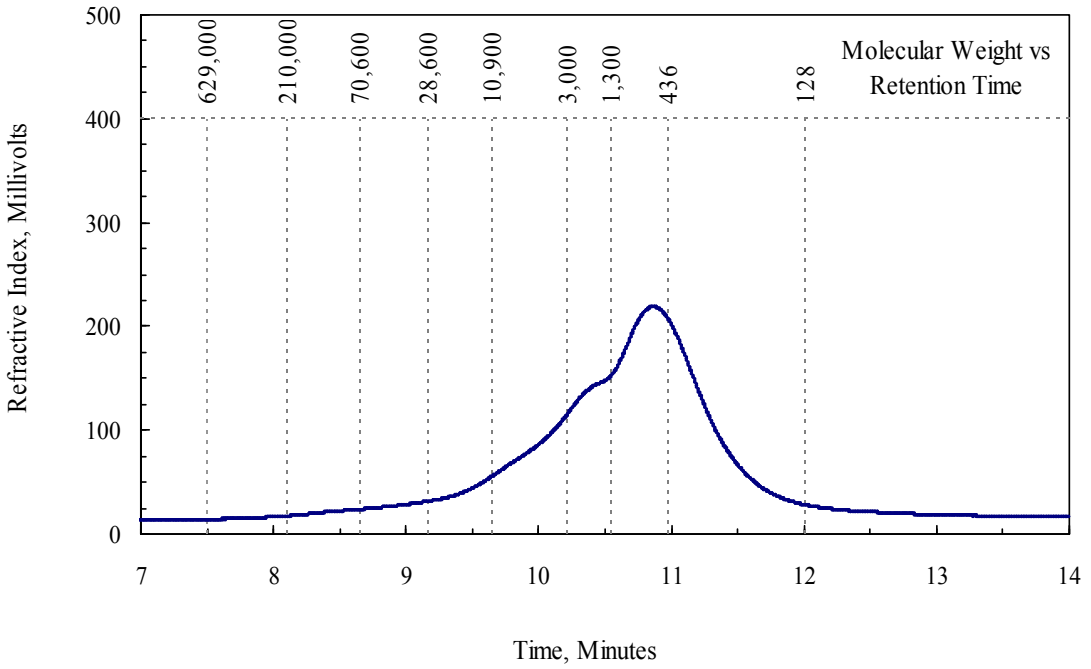



**FIGURE 4 Gel-permeation chromatographic results for PG 64-28 asphalt binder at a concentration of 7.0 percent in toluene.**

Asphalt binder 8 was concluded to be unmodified. The conclusions drawn from the data were substantiated by the FHWA and the PG was 64-28. Inspection of the chromatogram shown in Figure 5 reveals that there are two peaks for this asphalt binder. In this case, the concentration of the asphalt binder in toluene is 0.2 percent. The peak centered at the shorter retention time, about 10.4 minutes, may be the result of either intermolecular association of polar compounds in the asphalt binder or, may be more likely, the result of blending two asphalt binders of lighter and heavier grades to meet desired material characteristics. A summary of all the results presented in this report is shown in Table 2.

The GPC apparatus that was used for this work was configured to provide fast results at an average resolution. For better separation a high resolution GPC could be used. In part, a high resolution method relies on the selection of the spherical polyacrylamide beads, also known as the support, which is more appropriate for the average molecular weights of the solute that is to be separated. For example, since polymers and asphalt binders have widely divergent number average molecular weights, 100,000 versus 1,000, the support needs to be tailored for this specific range of molecular weights. In addition, only a refractive index detector was used to monitor the molecules as they emerged from the column. Other nondestructive detection techniques could be used to further define the chemical and physical nature of the solute. For example, an ultraviolet detector can be used to monitor for aromatic structures, an infrared detector can monitor specific wavelengths, a viscometer detector can monitor for changes in the viscosity, and an evaporative light scattering detector, which is considered to be a universal detector, can detect species that are not detected by the other detection techniques. The results

from the use of multiple detection systems could lead to the identification of not only the concentration of the polymer but also its chemical composition.

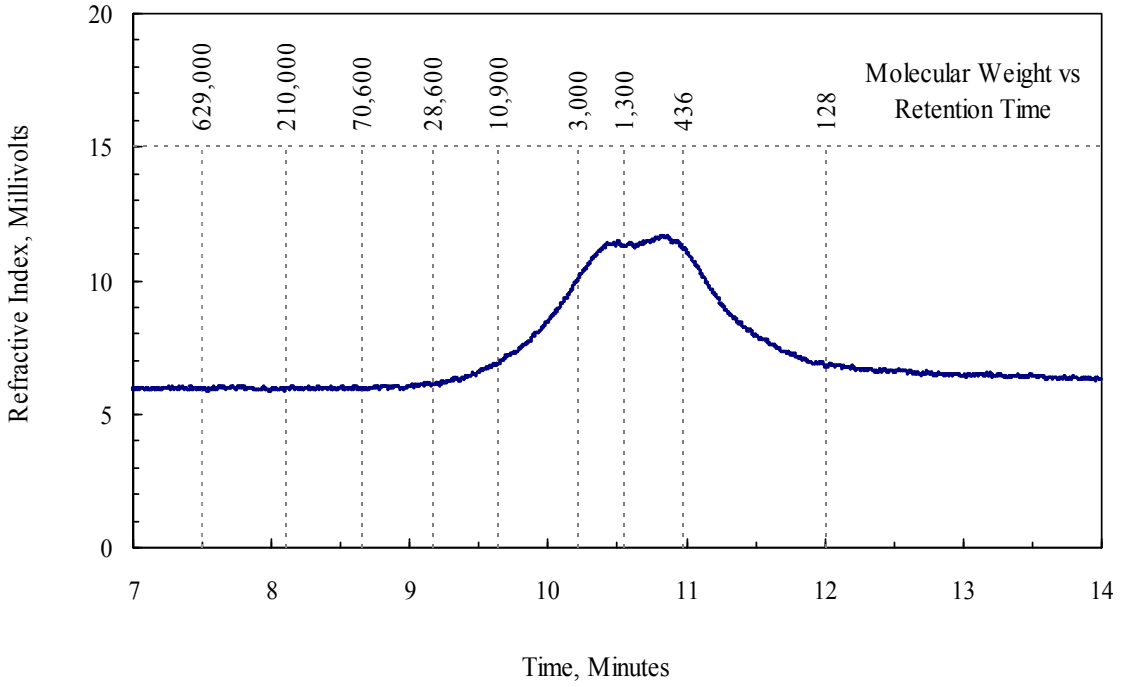



**FIGURE 5 Gel-permeation chromatographic results for PG 64-28 asphalt binder at a concentration of 0.2 percent in toluene.**

**Table 2. Results of the Gel-permeation Chromatographic Analysis of the Eight Asphalt Binders**

| Experimental Number of Asphalt Sample | Performance or Viscosity Grade of Asphalt | Conclusion Neat / Modified |
|---|---|---|
| 1 | PG 64-22 | Neat |
| 2 | AC-20P | Modified |
| 3 | PG 52-28 | Neat |
| 4 | PG 64-10 | Neat |
| 5 | PG 76-16 | Neat |
| 6 | PG 70-28 | Modified |
| 7 | PG 64-28 | Modified |
| 8 | PG 64-28 | Neat |

## CONCLUSIONS

The gel-permeation chromatographic analysis of eight different asphalt binders was conducted. It was determined that three of the binders contained a polymer, asphalt binders 2, 6, and 7. For the configuration of the apparatus used, the polymer was detected at retention times less than 9.2 minutes, while the asphalt binder itself was detected at retention times greater than 9.2 minutes. At a concentration of 7.0 percent asphalt binder in toluene, the occurrence of the intermolecular association of polar molecules in the different asphalt binders was evident.

The GPC instrumental method is described in sufficient detail such that other researchers can adapt the technique to their own use. The technique presented consisting of sample preparation, analysis of standards and samples (at two concentrations), and data reduction can be completed in about an 8-hour day or less depending on the number of samples, with the analysis time per sample being 15 minutes per concentration.

**Acknowledgements**

The authors gratefully acknowledge the Federal Highway Administration, U.S. Department of Transportation, for financial support of this project under Contract Number DTFH61-99C-00022. The authors gratefully appreciate Jackie Greaser for editing and formatting the document.

**Disclaimer**

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof. The contents of this report reflect the views of Western Research Institute, which is responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or the policy of the U.S. Department of Transportation. Mention of specific brand names is for information only and is not intended as an endorsement.

## REFERENCES

Altgetl, K.H., 1965, "Gel-permeation chromatography of asphalts and asphaltenes. I. Fractionation procedure", *Makromolekulare Chemie*, 88, 75-89.

Asi, I.M., Wahhab, H.I. Al-Abdul; Al-Dubabi, I.A.; and Ali, M.F., 1997, "Performance modeling of Arabian asphalt using HP-GPC", *Journal of Materials Engineering and Performance*, 6(4), 503-510.

Branthaver J.F., Petersen J.C., Robertson R.E., Duvall J.J., Kim S.S., Harnsberger P.M., Mill T., Ensley E.K., Barbour F.A., and Schabron J.F., 1993, "Binder Characterization and Evaluation, Volume 2: Chemistry", SHRP-A-368, SHRP, National Research Council, Washington, D.C., 287-320.

Brule, B., 1979, "Characterization of bituminous compounds by gel permeation chromatography (GPC)", *Journal of Liquid Chromatography*, 2(2), 165-192.

Donaldson, G.R., Hlavinka, M.W., Bullin, J.A., Glover, C.J., and Davison, R.R., 1988, "The use of toluene as a carrier solvent for gel permeation chromatography analysis of asphalt", *Journal of Liquid Chromatography*, 11(3), 749-765.

Garrick, N.W., and Biskur, R.R., 1990, "A classification system for interpreting GPC profiles of asphalts", *Asphalt Paving Technology*, 59, 33-53.

Garrick, N.W., and Wood, L.E., 1986, "Relationship between high-pressure gel permeation chromatography data and the rheological properties of asphalts", *Transportation Research Record 1096*, 35-41.

Herrington, P.R., and Ball, G.F.A., 1996, "Temperature dependence of asphalt oxidation mechanism", *Fuel*, 75(9), 1129-1131.

Huang, J., and Bertholf, D., 1997, "Effect of photooxidation on molecular weight distributions of Corbett asphalt fractions by gel permeation chromatography", *Petroleum Science and Technology*, 15(1 & 2), 37-49.

Huynh, H.K., Khong, T.D., Malhotra, S.L., and Blanchard, L.P., 1978, "Effect of molecular weight and composition on the glass transition temperature of asphalts", *Analytical Chemistry*, 50(7), 976-979.

Jennings, P.W., and Pribanic, J.A.S., 1989, "A perspective on asphalt chemistry research and the use of HP-GPC analysis", *Fuel Science and Technology International*, 7(9), 1269-1287.

Jennings, P.W., Pribanic, J.A.S., Mendes, T.M., and Smith, J.A., 1992, "High performance gel permeation chromatography in the characterization of self-assemblies in asphalt. I.", *Fuel Science and Technology International*, 10(4-6), 809-823.

Kim, K.W., Burati, J.L. Jr., and Amirkhanian, S.N., 1993, "Relation of HP-GPC profile with mechanical properties of AC mixtures", *Journal of Materials in Civil Engineering*, 5(4), 447-459.

Lee, Y.J., France, L.M., and Hawley, M.C., 1997, "Characterization of asphalt binders mixed with epoxy terminated ethylene terpolymer", *Polymeric Materials: Science and Engineering*, 76, 397-398.

Negulescu, I., Mohammad, L., Daly, W., Abadie, C., Cueto, R., Daranga, C, and Glover, I., 2006, "Chemical and rheological characterization of wet and dry aging of SBS copolymer modified asphalt cements", *Association of Asphalt Paving Technologists*, 75, 267-296.

Norton, J., 2003, "Unique Additives", *The Asphalt Contractor*®, September 2003, 17(10).

Price, R.P., and Burati, J.L. Jr., 1989, "A quantitative method using HP-GPC to predict laboratory results of asphalt cement tests", *Asphalt Paving Technology*, 58, 182-219.

Siddiqui, M.N., and Ali, M.F., 1999, "Investigation of chemical transformations by NMR and GPC during the laboratory aging of Arabian asphalt", *Fuel*, 78(12), 1407-1416.

Wei, J.B., Shull, J.C., Lee, Y.J., and Hawley, M.C., 1996, "Characterization of asphalt binders based on chemical and physical properties", *International Journal of Polymer Analysis and Characterization*, 3(1), 33-58.

# Effects of a Wax-Based Warm Mix Additive on Cohesive Strengths of Asphalt Binders

N. Wasiuddin[1], N. Saltibus[2] and L. Mohammad[3]

[1]Corresponding Author and Assistant Professor of Civil Engineering, College of Engineering and Science, Louisiana Tech University, 600 W. Arizona Ave., Ruston, Louisiana, USA; PH (318) 257-2392; FAX (318) 257-2306; email: wasi@latech.edu
[2]Master's Candidate, Civil Engineering, College of Engineering and Science, Louisiana Tech University, 600 W. Arizona Ave., Ruston, Louisiana, USA; PH (318) 957-9301; FAX (318) 257-2306; email: nes011@latech.edu
[3]Irma Louise Rush Stewart Distinguished Professor of Civil and Environmental Engineering and Director, Engineering Materials Characterization Research Facility, Louisiana Transportation Research Center, Louisiana State University, 4101 Gourrier Ave., Baton Rouge, Louisiana 70808, USA. Phone/Fax: (225) 767-9126/9179, Email: louaym@lsu.edu

## ABSTRACT

Despite the positive effects, the use of wax-based warm mix additive is very limited due to the concerns on cohesion and adhesion properties. The objective of this study is to evaluate the cohesive strengths of asphalt binders with and without Sasobit®, a wax-based warm mix additive using a pneumatic adhesion tensile testing instrument (PATTI) and using surface free energy method from optical advancing contact angle measurements. Two asphalt binders, PG 64-22 and PG 76-22M with and without 2% Sasobit® were coated in 2"X2"X1/4" glass plates. Cohesive strengths of asphalt binders were obtained in terms of pull-off tensile strength (POTS) using ½ in diameter pull-stubs. It was observed that the cohesive strengths of PG64-22 and PG76-22M are 1.964 MPa (284.9±21.3 psi) and 2.067 MPa (299.9±15.3 psi), respectively. Sasobit® reduced the cohesive strengths to 1.468 MPa (212.9±11.3 psi) and 1.503 MPa (218.0±17.7 psi), respectively for PG 64-22 and PG 76-22M. The cohesive strengths (free energy of cohesion) obtained from surface free energy measurements indicate that PG 64-22 has higher cohesive strengths and Sasobit® increases the cohesive strengths of both the binders. Overall, pull-off test was found to be simple and reproducible.

## INTRODUCTION

Warm mix asphalt (WMA), due to its environmental friendliness, is gaining strength in its popularity as a material used in the construction of roadways. Energy savings and a reduction in emission problems are benefits of the many warm mix technologies that are available for decreasing compaction and asphalt mixing temperatures. Organic additives, such as commercial wax, demonstrate significant promises from that of the available technologies. For these technologies, there was a noted improvement in compaction at temperatures as low as 88ºC (Hurley and Prowell, 2005). Lower plant temperatures, as results have shown, can lead to a

significant reduction, up to 30 percent, in energy consumption. Furthermore, WMA is an environmentally attractive material, as reduced production temperature leads to a reduction in emission. This reduction in emissions denotes a significant cost savings, taking into account that 30-50 percent of overhead costs at an asphalt plant are credited to emission control (Kristjansdottir et al., 2007). The effects of natural wax as an additive in asphalt binder and hot mix asphalt (HMA) has been under discussion for decades, suggesting both positive and negative effects. Natural wax in asphalt, however, is currently low in content and less probable to be harmful to binder or HMA properties (Edwards and Redelius, 2003). In order to gain certain positive effects, the use of commercial waxes as additives to asphalt binder in WMA, can be of significant interest. Commonly used commercial waxes in WMA are Fischer-Tropsch (FT), paraffin wax (Sasobit®), and montan wax (Asphaltan B®). Asphaltan B® is a product of Romonta, GmbH (Germany), while Sasobit® is a product of Sasol Wax GmbH (Germany). The production of Sasobit® is in FT Synthesis, where carbon monoxide converts into higher hydrocarbons in catalytic hydrogenation, which is followed by a distillation process. The end product of the aforementioned process consists of mainly fine crystalline long chain aliphatic polymethylene hydrocarbon chains with 40-100 carbon atoms. To compare, macrocrystalline bituminous paraffin waxes contain carbon chain lengths in the range of 25 to 50. Due to the longer carbon chains in the Sasobit®, there is a higher melting point, and the distribution of the wider wax molecules results in an increased plasticity span and a broader melting range. Natural asphalt wax is normally completely melted within a range of 60°C to 90°C, whereas when Sasobit® is mixed with asphalt, the melting temperatures are higher (approximately between 100°C and 130°C) (Butz et al., 2001). Many researchers have earlier studied the effects of Sasobit® in HMA and asphalt binders. The mixing and compaction temperatures of asphalt are known to be reduced by approximately 18-54°C with the help of Sasobit®. Also, Sasobit® has been reported to improve the resistance to deformation at high temperatures for asphalt binder and HMA (rutting) (Hurley et al., 2005). However, especially in the case of overdosing, the increase in creep stiffness and a reduction in creep rate (m) at low temperature grading may be of concern (Wasiuddin et al., 2007a; Edwards and Redelius, 2003). Moreover, incomplete drying of the aggregate can be caused by a lower mixing and compaction temperature. There is a potential for moisture damage, as a result of the water being trapped in the coated aggregate. Finally, reduced-aging properties of Sasobit® have also been reported by some researchers.

Significant research in the past has been undertaken on the effect of commercial wax such as Sasobit®, on HMA, asphalt binder. But an understanding of the effects that commercial wax has on the cohesive strengths of asphalt binders is still lacking. The test method for bond strength of coatings is specified in ASTM D 4541 *Pull-off Strength of Coatings using Portable Adhesion Testers*. Recently, a modified version has been used by several researchers (Copeland et al., 2006a; 2006b; Kanitpong and Bahia, 2003; 2005). Using a Pneumatic Adhesion Tensile Testing Instrument (PATTI), researchers observed that when tested dry, the failure is cohesive. In this study, a modified pull-off test has been used to measure the cohesive strengths of asphalt binders with and without Sasobit®.

Previously, the effects of amine anti-strip additives on the surface free energy (SFE) components (acids, base and non-polar) of asphalt binders for moisture induced damage potential of HMA were studied by Wasiuddin et al. (2007 b). In another study, the thermal degradation of anti-strip additives caused by aging was characterized by Wasiuddin et al. (2007c). Also, proposed by these researchers is the evaluation of wettability and adhesion, based on the SFE method, of polymer treated aggregates for moisture-induced damage potential of HMA (Wasiuddin et al., 2010). Wasiuddin et al. (2008) further evaluated wettability and adhesion characteristics of WMA with a wax-based and zeolite-based additive for moisture susceptibility analyses. In this study, cohesive strengths (free energy of cohesion) of asphalt binders with and without a wax-based WMA additive will be evaluated based on their SFE characteristics. Finally, the cohesive strengths obtained from the pull-off test will be compared with the cohesive strengths (free energy of adhesion) obtained from SFE analyses. A wax-based WMA additive, Sasobit®, will be used for this purpose.

## OBJECTIVES

The overall objective of this study is to determine the effects of a wax-based WMA additive on dry cohesive strengths of asphalt binders using both modified pull-off test and surface free energy measurements. The specific objectives are as follows:

1) Evaluate the dry cohesive strengths of asphalt binders with and without Sasobit® using pull-off test
2) Measure the surface free energy of asphalt binders with and without Sasobit® using an optical advancing contact angle analyzer
3) Calculate cohesive strengths (free energy of cohesion) from surface free energy values
4) Determine stiffness coefficient, $G^*/\sin\delta$ at 34°C and 70°C of asphalt binders with and without Sasobit®
5) Compare cohesive strengths of asphalt binders obtained from pull-off test and surface free energy measurements

## MATERIALS

Both the asphalt binders, PG 64-22 and polymer-modified PG 76-22M, used in this study were obtained from Ergon Asphalt and Emulsions, Inc., Jackson, Mississippi. The product source codes according to Louisiana Department of Transportation and Development (DOTD) are 41-BT and 41-AT, respectively. Sasobit® was obtained from Sasol Wax, South Africa.

## MODIFIED PULL-OFF TEST

### Modified Pull-Off Test Method and Sample Preparations

The modified pull-off test followed by Copeland et al. (2006a; 2006b) has been further modified in this study. Previously Copeland et al. (2006a; 2006b) used

pneumatic adhesion tensile testing instrument (PATTI) for cohesive and adhesive strengths of asphalt binders. The pull-off tensile strength of dry specimens was called a measure of cohesive strength whereas adhesive strength was determined after moisture conditioning in their study. Copeland et al. (2006a; 2006b) applied a porous ceramic stub to the pull-stub using two-part epoxy glue. The porous stub allows water to migrate consistently through to the asphalt film. A small sample of asphalt binder (< 10 grams) is mixed with one percent (by weight) 200 μm glass beads to ensure uniform film thickness. The sample is heated to about 100°C using a hot plate. The sample is then applied to the ceramic stub and the pull-stub is pressed onto a glass substrate by the test operator. The samples were then cured for 24 hrs at room temperature. Dry specimens were tested after curing.

In this study, 200 μm glass beads were not added to the asphalt binder. Also, no porous ceramic stub was adhered to the pull-stub. Instead, plain glass plates coated with asphalt binder were used and hereafter will be called ***asphalt samples*** in this paper. Plain glass plate coated with asphalt binder and a pull stub attached to that will be called ***pull-stub samples*** hereafter in this paper. The glass plates are 2 in. long, 2 in. wide and ¼ in. thick. Asphalt binders with or without Sasobit®, the glass plates and the sand blasted aluminum pull-stubs were heated for 2 hours at 167°C in an oven. The top face of the glass plate was coated and asphalt samples were prepared. In case of pull-stub samples, the top face of the glass plate and the bottom face of the pull stub were coated and adhered. The samples (Figure 1) were then rested overnight before pull-off testing using PATTI to obtain dry cohesive strength of asphalt binders at room temperature. Figure 2 describe the pull-off testing procedure using a PATTI.

## Pull-Off Test Results

Table 1 shows the cohesive strengths of asphalt binders with and without additives. Maximum gasket pressure needed to pull off is obtained in psi from the PATTI. A stop watch was used to count the time needed to pull off. Pressure rate was obtained by dividing the pull-off strength by the time required.

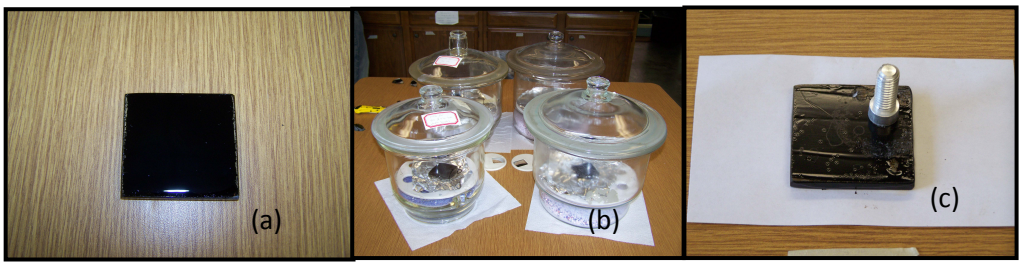



Figure 1. Sample Preparations for Dry Cohesive Strength:
(a) Asphalt Samples, (b) Asphalt Samples in Desiccators, and
(c) Pull-Stub Samples.

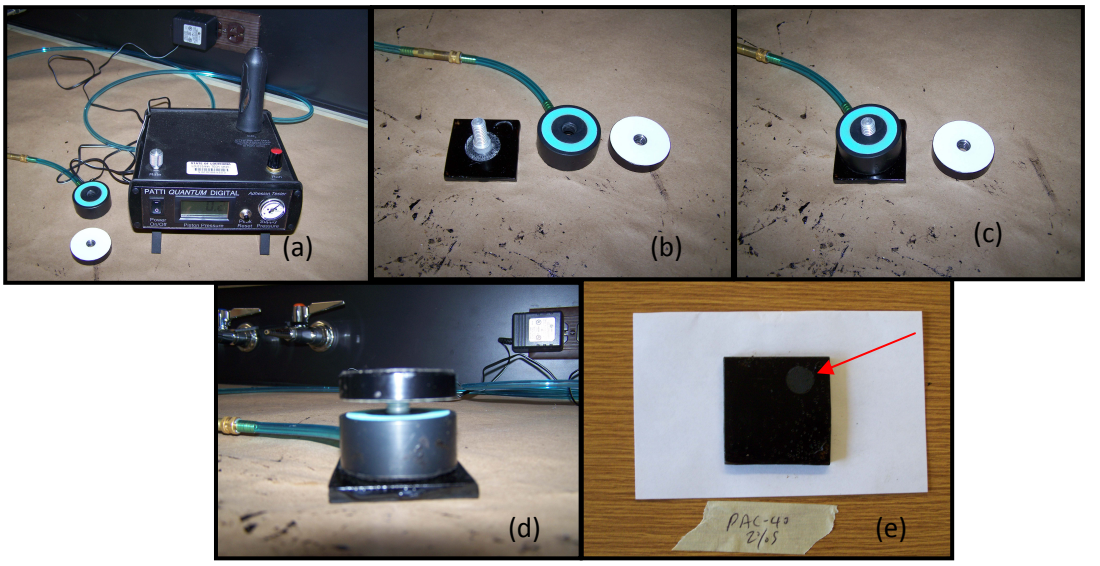



Figure 2. (a) PATTI, (b) Pull-Stub sample, Gasket (blue) and Reaction Plate (white), (c) Pull-Stub Sample Inside Gasket, (d) Pull-Stub Sample Inside Gasket and Reaction Plate, and (e) Cohesive Failure.

**Effects of Polymer on Dry Cohesive Strengths of Asphalt Binders**

Although polymer is used to improve performance grading of asphalt binders, it was observed in this study that the polymer does not increase the dry cohesive strength significantly. Table 1 shows that the pull-off strength of PG64-22 is 1.964 MPa (284.9 psi) with an standard deviation of 21.3 psi whereas, the pull-off strength of PG76-22M is 2.067 MPa (299.9 psi) with an standard deviation of 15.3 psi.

**Effects of Sasobit® on Dry Cohesive Strengths of Asphalt Binders**

Table 1 shows that Sasobit® reduces the dry cohesive strengths of both PG 64-22 and PG76-22M, a polymer modified binder. It can be seen that Sasobit® reduces the dry cohesive strengths of PG 64-22 from 1.964 MPa (284.9 psi) to 1.468 MPa (212.9 psi) while in case of PG76-22M, the reduction is from 2.067 MPa (299.9 psi) down to 1.503 MPa (218.0 psi). The standard deviations for these values are 21.3, 11.3, 15.3 and 17.7 psi respectively. This shows that the reductions are significant as compared to standard deviation values. In the case of PG 64-22, the reduction is 0.496 MPa (72.0) psi which is 25.3% while the reduction is 0.565 MPa (81.9 psi) for PG76-22M which is 27.3%.

**Effects of Asphalt Film Thickness on Dry Cohesive Strengths of Asphalt Binders**

In this study, an average film thickness of approximately 0.1mm was used. The film thickness was determined by obtaining the asphalt film volume from asphalt film weight and then dividing the volume by the glass plate area. Table 2 shows that

thickness has significant influence on dry cohesive strength. If the thickness is smaller than 0.1 mm, it will reduce the dry cohesive strength. As for example, with reduced thickness the dry cohesive strength of PG 76-22M reduces to 246.9 psi from 299.9 psi. Similar reduction in dry cohesive strength was observed for PG 64-22. In case of Sasobit® added asphalt binders, no significant reduction was observed with respect to thickness reduction for both the asphalt binders.

## SURFACE FREE ENERGY (SFE) METHOD

SFE of a solid (or liquid) is defined as the work required to increase a unit area of surface of that solid under vacuum. Consequently, the free energy of cohesion is the work done by a unit force acting along the surface of an asphalt binder at a right angle to any line of unit length against a cohesive force to create two interfaces from one (i.e., asphalt binder) under vacuum. The SFE of an asphalt binder mainly comprises of an apolar component (also called Lifshitz-van der Waals component) and an acid-base component and the acid-base term can be decomposed to a Lewis acidic surface parameter and a Lewis basic surface parameter as follows (Good, 1992):

$$\Gamma = \Gamma^{LW} + \Gamma^{AB}$$

$$\Gamma^{AB} = 2\sqrt{\Gamma^{+}\Gamma^{-}}$$

where, $\Gamma$ = SFE of the asphalt binder, $\Gamma^{LW}$ = Lifshitz-van der Waals component of the SFE, $\Gamma^{AB}$ = Acid-Base component of the SFE, $\Gamma^{+}$ = Lewis acid component of SFE, and $\Gamma^{-}$ = Lewis base component of SFE.

Dynamic contact angles for different liquids can be used, as employed in this study, to evaluate these SFE components. The Sessile Drop Method was used in this study to measure advancing contact angles.

### Advancing Contact Angle from Sessile Drop Method (SDM)

The advancing contact angle of a liquid on an asphalt film (in this case prepared asphalt samples) was measured by an optical contact angle analyzer called OCA 15 Plus from Future Digital Scientific Corp. The dynamic contact angle of a liquid drop was measured while releasing liquid at the rate of 0.33 µl/sec to a final drop size of 8 µl. A software called SCA20 available from Future Digital Scientific Corp. was used for measurements and data analyses.

### Calculation of Free Energy of Cohesion

Assuming that equilibrium film pressure is negligible for an asphalt binder, Young's equation and Dupre's equation can be combined with Good's postulate to obtain the so called Young-Dupre equation (Good, 1992). The resulting Young-Dupre equation can be expressed as follows:

$$\Gamma_L(1+\cos\theta) = 2\times\sqrt{\Gamma_S^{LW}\Gamma_L^{LW}} + 2\times\sqrt{\Gamma_S^{-}\Gamma_L^{+}} + 2\times\sqrt{\Gamma_S^{+}\Gamma_L^{-}}$$

where, $\Gamma_L^{LW}$, $\Gamma_L^{+}$, and $\Gamma_L^{-}$ = SFE components of liquid solvent, $\Gamma_S^{LW}$, $\Gamma_S^{+}$, and $\Gamma_S^{-}$ = SFE components of asphalt binder, and $\theta$ = Contact angle.

In the above equation, the SFE components of an asphalt binder are given by the three unknowns ($\Gamma_S^{LW}$, $\Gamma_S^+$, and $\Gamma_S^-$). To obtain these unknowns, dynamic contact angles must be measured in at least three different liquid solvents. The SFE characteristics of these liquid solvents must be known a priori. Water, formamide and diiodomethane were used here as liquid solvents because of their relatively large SFE, immiscibility with asphalt binder, and differing SFE components (Cheng et al., 2002). The free energy of cohesion (ΔG), as defined previously, can be calculated from the following equation:

$$\Delta G = 2\sqrt{\Gamma_{asphalt}^{LW}\Gamma_{asphalt}^{LW}} + 2\sqrt{\Gamma_{asphalt}^{+}\Gamma_{asphalt}^{-}} + 2\sqrt{\Gamma_{asphalt}^{-}\Gamma_{asphalt}^{+}} = 2\Gamma$$

where, ΔG = Free energy of cohesion.

**Surface Free Energy of Asphalt Binders**

Table 3 shows that the total SFE of PG 64-22 is 20.4 ergs/cm$^2$ while the total SFE of PG 76-22M is 15.2 ergs/cm$^2$. The major contributions to these energies come from non-polar SFE component ($\Gamma^{LW}$). Although the total SFE and the non-polar SFE of PG 64-22 are higher than PG 76-22M, the acid-base component of PG 76-22M is few times higher than that of PG 64-22.

Addition of 2% Sasobit® increases the total SFE of PG 64-22 to 30.2 ergs/cm$^2$. A similar increasing trend is observed in the case of PG 76-22M. Also, a general trend is that the additive increases the base component ($\Gamma^-$) of SFE for both the binders. A similar trend of increasing the base component by similar additives was observed previously by Wasiuddin et al. (2007b; 2007c; 2008).

**Dry Cohesive Strengths of Asphalt Binders from SFE Measurements**

Table 3 shows that the cohesive strengths (free energy of cohesion) of PG 64-22 and PG 76-22M are 40.8 ergs/cm$^2$ and 30.4 ergs/cm$^2$, respectively. Addition of 2% Sasobit® increases the cohesive strengths to 60.4 ergs/cm$^2$ and 58.5 ergs/cm$^2$, respectively for PG 64-22 and PG 76-22M.

Table 1. Dry Cohesive Strengths from Pull-Off Test

| PG 64-22 | | |
|---|---|---|
| | Dry Cohesive Strength (psi) ± Std. Dev. | Dry Cohesive Strength (MPa) |
| PG 64-22 | 284.9±21.3 | 1.964 |
| PG 64-22 + 2% Sasobit® | 212.9±11.3 | 1.468 |
| PG 76-22M | | |
| PG 76-22M | 299.9±15.3 | 2.067 |
| PG 76-22M + 2% Sasobit® | 218.0±17.7 | 1.503 |

Table 2. Effect of Film Thickness on Dry Cohesive Strengths

| | Pull-Off Strength (psi) | Pull-Off Strength (psi) |
|---|---|---|
| | Thin | Thick |
| PG 64-22 with No Additives | | |
| Average | 244.6 | 284.9 |
| St. Dev. | 38.4 | 21.3 |
| PG 64-22 with 2% Sasobit® | | |
| Average | 207.0 | 212.9 |
| St. Dev. | 3.5 | 11.3 |
| PG 76-22M with No Additives | | |
| Average | 246.9 | 299.9 |
| St. Dev. | 23.2 | 15.3 |
| PG 76-22M with 2% Sasobit® | | |
| Average | 210.1 | 218.0 |
| St. Dev. | 21.1 | 17.7 |

**Comparison of Dry Cohesive Strengths**

The dry pull-off strengths (dry cohesive strengths) from PATTI of PG 64-22 and PG 76-22M are 1.964 MPa (284.9 psi) and 2.067 MPa (299.9 psi), respectively whereas, the free energy of cohesion (dry cohesive strengths) from SFE measurements are 40.8 and 30.4 ergs/cm$^2$, respectively.

It is known that G*/sinδ which is an indictor of stiffness against rutting, of PG 76-22M is much higher than that of PG 64-22. Table 4 shows that the G*/sinδ of PG 64-22 and PG 76-22M are very different at 70°C whereas the values are comparable at lower temperatures such as 34°C. Therefore, the dry cohesive strengths obtained from pull-off test in this study where PG 76-22M shows a little higher cohesive strength than that of PG 64-22 can be justified. However, PG 64-22 has higher cohesive strength than that of PG 76-22M as obtained from SFE measurements can not be justified.

Table 3. Surface Free Energy (SFE) Components and Dry Cohesive Strengths (Free Energy of Cohesion)

| | Total SFE (ergs/cm$^2$) $\Gamma^L$ | Non-Polar SFE in (ergs/cm$^2$) $\Gamma^{LW}$ | Acid SFE (ergs/cm$^2$) $\Gamma^+$ | Base SFE (ergs/cm$^2$) $\Gamma^-$ | Acid-Base Polar SFE (ergs/cm$^2$) $\Gamma^{AB}$ | Cohesive Strength (ergs/cm$^2$) |
|---|---|---|---|---|---|---|
| PG 64-22 | 20.4 | 20.1 | 0.01 | 2.0 | 0.3 | 40.8 |
| PG 64-22+ 2% Sasobit® | 30.2 | 26.8 | 0.9 | 3.1 | 3.4 | 60.4 |
| PG 76-22M | 15.2 | 13.1 | 0.6 | 1.8 | 2.1 | 30.4 |
| PG 76-22M+ 2% Sasobit® | 29.3 | 26.8 | 0.7 | 2.2 | 2.4 | 58.5 |

Cohesive strengths from SFE measurements (Table 3) show that additives increase the cohesive strengths. Table 4 shows that G*/sinδ increases with addition of 2% Sasobit® at 34°C. However, on the other hand, the cohesive strengths obtained from pull-off test indicate that Sasobit® reduces the dry cohesive strengths of both the asphalt binders.

Table 4. G*/Sinδ of PG 64-22 and PG 76-22M at Different Temperatures

| G*/Sinδ (KPa) | | |
|---|---|---|
| Temperature | 34°C | 70°C |
| PG 64-22 | 227.9 | 1.43 |
| PG 76-22M | 247.9 | 4.32 |
| PG 64-22 with 2% Sasobit® | 618.0 | |
| PG 76-22M with 2% Sasobit® | 473.9 | |

## CONCLUSIONS

Several specific conclusions can be drawn from this study:

- The dry cohesive strengths of PG 64-22 and PG 76-22M obtained from pull-off test are comparable, 1.964 MPa (284.9 psi) and 2.067 (299.9 psi), respectively.
- Addition of 2% Sasobit® reduces the dry cohesive strengths of asphalt binders obtained from pull-off test, in this case reduces to 1.468 MPa (212.9 psi) and 1.503 MPa (218.0 psi), respectively for PG 64-22 and PG 76-22M.
- Stiffness against rutting, G*/sinδ at 34°C indicates that stiffness of PG 76-22M is only marginally higher than PG 64-22.
- Cohesive strengths (free energy of cohesion) of PG 64-22 and PG 76-22M obtained from surface free energy measurements are 40.8 and 30.4 ergs/cm$^2$, respectively which does not correlate well with pull-off test results.
- Surface free energy measurements indicate that 2% Sasobit® increases the cohesive strengths whereas, pull-off test indicates the reverse.
- Overall, pull-off test as followed in this study was found to be simple and reproducible.

## ACKNOWLEDGEMENTS

The authors want to sincerely thank Louisiana Transportation Research Center (LTRC) for providing funding for this project.

## REFERENCES

Butz, T., Rahimian, I., and Hildebrand, G. (2001). "Modifications of road bitumens with the fischer-tropsch paraffin Sasobit®." *Journal of Applied Asphalt Binder Technology*, 1(2), 70-86.

Cheng, D. X., Little, D. N., Lytton, R. L., and Holste, J. C. (2002) "Use of surface free energy properties of the asphalt-aggregate system to predict damage

potential." *Proceedings of Association of Asphalt Paving Technologists,* 71, 59-88.

Copeland, A., Kringos, N., Scarpas, A., Youtcheff, J., and Mahadevan, S. (2006). "Determination of bond strength as a function of moisture content at the aggregate-mastic Interface." *10th International Conference on Asphalt Pavements,* Quebec City, Quebec, August, 12-17.

Copeland, A.R. and Youtcheff, J. (2006a) "Moisture Sensitivity of Modified Asphalt Binders: Factors Influencing Bond Strength*." 10th International Conference on Asphalt Pavements*, Quebec City, Quebec, August, 12-17.

Edwards, Y., and Redelius, P. (2003). "Rheological effects of waxes in bitumen." *Energy and Fuels*, 17(3), 511-520.

Good, R. J. (1992) "Contact angle, wetting and adhesion: a critical review." *J. Adhes. Science and Tech.*, 6(12), 1269-1302.

Hurley, G. C., and Prowell, B. D. (2005a). "Evaluation of Sasobit® for use in warm mix asphalt." NCAT Report 05-06, *National Center for Asphalt Technology*.

Kanitpong, K. and Bahia, H.U. (2003). "Role of adhesion and thin film tackiness of asphalt binders in moisture damage of HMA." *J. Assoc. Asph. Paving Technol.*, 72, 502-528.

Kanitpong, K. and Bahia, H.U. (2005). "Relating adhesion and cohesion of asphalts to effect of moisture on asphalt mixtures' laboratory performance." *Presented at the 84th Annual Meeting of the Transportation Research Board*, *Washington, D.C.*

Kristjansdottir, O., Muench, S.T., Michael, L., and Burke, G. (2007). "Assessing the potential for warm mix asphalt technology adoption." *Presented at the 86th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Wasiuddin, N.M., Fogle, C.M., Zaman, M.M. and O'Rear, E.A. (2007b) "Effect of anti-strip additives on surface free energy characteristics of asphalt binders for moisture-induced damage potential." *J. Test. and Eval.*, 35(1), 36-44.

Wasiuddin, N.M., Fogle, C.M., Zaman, M.M. and O'Rear, E.A. (2007c). "Characterization of thermal degradation of liquid amine anti-strip additives in asphalt binders due to RTFO and PAV-Aging." *J. Test. Eval.*, 35(4).

Wasiuddin, N.M., Selvamohan, S., Zaman, M.M. and Guegan, M.L.T. (2007a). "A comparative laboratory study of Sasobit® and Aspha-Min® in warm mix asphalt." *Transportation Research Record. 1998*, Transportation Research Board, Washington, D.C., 82-88.

Wasiuddin, N.M., Zaman, M.M. and O'Rear, E.A. (2008). "Effect of Sasobit® and Aspha-Min® on wettability and adhesion between asphalt binders and aggregates." *Transportation Research Record. 2051*, Transportation Research Board, Washington, D.C., 80-89.

Wasiuddin, N.M., Zaman, M.M. and O'Rear, E.A. (2010). "Polymeric aggregate treatment using styrene-butadiene rubber (SBR) for moisture-induced damage potential." *Int. J. Pavement. Res. and Technol*., 3(1), 1-9.

# Comparison of Pavement Design Using AASHTO 1993 and NCHRP Mechanistic–Empirical Pavement Design Guides

Daba S. Gedafa, Ph.D., M.ASCE[1], James Mulandi, M.ASCE[2], Mustaque Hossain, Ph.D., P.E., F.ASCE[3] and Greg Schieber, P.E., M.ASCE[4]

**Abstract.** The new Mechanistic-Empirical Pavement Design Guide (MEPDG) provides methodologies for mechanistic-empirical pavement design as opposed to the empirical methodology used in the 1993 American Association of State Highway and Transportation Officials (AASHTO) pavement design guide. The objective of this study was to compare the pavement designs obtained using the 1993 AASHTO and the new MEPDG methods for typical Portland Cement Concrete (PCC) and Asphalt Concrete (AC) pavements in Kansas. Five in-service Jointed Plain Concrete Pavement (JPCP) projects were reanalyzed as equivalent JPCP and AC projects using both approaches at the same reliability level. The results show that the new MEPDG analysis yielded thinner AC sections for all projects than those obtained from the 1993 AASHTO design guide analysis. Four of the PCC sections, designed using the 1993 AASHTO design guide, were thicker than the sections obtained with MEPDG. The MEPDG analysis resulted in thicker PCC slab for the fifth project.

Effect of change in performance criteria on the thickness of AC and PCC sections has also been investigated. It has been found that AC sections are more sensitive to change in performance criteria as compared to PCC sections using MEPDG versions 1.0 and 1.1. In general, difference in thickness using both versions is not significant for all practical purposes.

## Introduction

The design procedure recommended by AASHTO is based on the results of the AASHO road test in the late 1950's and early 1960's. The first guide was published in 1961 and was subsequently revised in 1972 and 1981. The guide was revised again in 1986 and then another version was released in 1993. The 1986 and 1993 guides include some further modifications based on theory and experience. The AASHO road test was conducted under a given climatic setting with a specific set of pavement materials and subgrade soils under vehicle operating and loading characteristics that were typical at that time, but not representative of current conditions. Due to these limitations, a design guide based as fully as possible on the mechanistic principles was developed under the National Cooperative Highway Research Program (NCHRP 2004).

[1]University of Connecticut, Department of Civil & Environmental Engineering, 261 Glenbrook Rd Unit 2037, Storrs, CT 06269. E-mail:dgedafa@engr.uconn.edu

[2]Traffic Engineer, H.W. Lochner, Inc.,1245 East Brickyard Road, Suite 400, Salt Lake City, UT 84106. E-mail: jmulandi@hwlochner.com

[3]Department of Civil Engineering, 2118 Fiedler Hall, Kansas State University, Manhattan, KS 66506. E-mail: mustak@ksu.edu

[4]Bureau of Materials & Research, Kansas Department of Transportation, 2300 Van Buren, Topeka, KS 66611. E-mail: gregs@ksdot.org

## Objective

The main objective of this study was to do a comparative analysis of structural designs for typical asphalt and concrete pavement sections using both 1993 AASHTO Design Guide and MEPDG. The effect of change in performance criteria on the thickness of the pavement has also been investigated using MEPDG Software versions 1.0 and 1.1.

## Overview of the 1993 AASHTO Design Guide and MEPDG

The procedures developed in the 1986 and 1993 AASHTO guides for new construction or reconstructions are based on the algorithms originally developed from the AASHO road test. Some newer concepts such as the resilient modulus for pavement material characterization and reliability were introduced in these versions (AASHTO 1986, AASHTO 1993). Also, the layer coefficients for the various materials were defined in terms of resilient modulus as well as standard methods, such as the California Bearing Ratio (CBR). The environmental factors of moisture and temperature were objectively included in lieu of the previously used subjective regional factor. The reliability term was introduced to permit the designer to use the concept of risk analysis for various functional classes of highways.

Mechanistic-empirical (M-E) design combines the elements of mechanical modeling and performance observations in determining required pavement thickness for a given set of design inputs (Timm et al. 1998). Yoder and Witczak (1975) pointed out that for any pavement design procedure to be completely rational in nature, three elements must be fully considered: (i) the theory used to predict the assumed failure or distress parameter; (ii) the evaluation of the materials properties applicable to the selected theory; and (iii) the determination of the relationship between the magnitude of the parameter in question to the performance level desired. The newly developed M-E pavement design guide considered all three elements.

### *Design Approach in MEPDG*

In the MEPDG analysis, the designer first considers site conditions (traffic, climate, material and existing pavement condition, if applicable in case of rehabilitation) and construction conditions in proposing a trial design for a new pavement or rehabilitation. The trial design is then evaluated for adequacy against some predetermined failure criteria. Key distresses and smoothness are predicted from the computed structural responses of stress, strain, and deflection due to given traffic and environmental loads. If the design does not meet desired performance criteria at a preselected level of reliability, it is revised and the evaluation process is repeated as necessary (NCHRP 2004). This approach makes it possible to optimize the design and to fully ensure that specific distress types will not develop.

## Test Sections

Five in-service Jointed Plain Concrete Pavement (JPCP) projects in Kansas were reanalyzed as equivalent JPCP and asphalt concrete (AC) projects using the NCHRP MEPDG and the 1993 AASHTO Design Guide. The AASHTO design guide analysis was performed using the DARwin software. Four of the projects analyzed are located on Interstate Route 70 (I-70), and the other one on Kansas state route K-7. Table 1 tabulates the project features of these sections. Four

projects have a 15 cm lime-treated subgrade (LTSG) to reduce the plasticity and/or to control moisture susceptibility. The reported subgrade modulus in Table 1 was computed by the MEPDG software based on the correlation equation involving plasticity index and gradation. According to the Unified Soil Classification, all projects have silty clay (CL) soils.

**Table 1.** Project Features of the Study Sections

| Route | County | Mile Post Limit | Year Built | Subgrade Soil Type | Subgrade Modulus (MPa) | Initial AADT | % Truck |
|---|---|---|---|---|---|---|---|
| I-70 | Geary | 0-7 | 1990 | A-6 | 67 | 9,200 | 18 |
| I-70 | Dickinson | 20-22.6 | 1992 | A-6 | 48 | 11,970 | 22.3 |
| I-70 | Shawnee | 9-10 | 1993 | A-7-6 | 43 | 36,000 | 5 |
| K-7 | Johnson | 12-15 | 1995 | A-7-6 | 50 | 13,825 | 7 |
| I-70 | Wabaunsee | 0-5.2 | 2001 | A-6 | 49 | 18,000 | 20.5 |

## Design Inputs

### *1993 AASHTO*

In order to compare two procedures, it is extremely important that the inputs into the procedures are somewhat equivalent. Common inputs for AC and PCC pavements are design traffic, serviceability, reliability, design period, and subgrade properties. Specific design inputs for AC pavements are structural number based on layer coefficients and drainage coefficient, and then selection of layer thicknesses based on structural number. Specific design inputs for PCC pavements are effective modulus of concrete, concrete modulus of rupture, load transfer coefficient, and drainage coefficient.

### *MEPDG*

In the mechanistic-empirical (M-E) design, the key outputs are the individual distress quantities. For instance, for JPCP, MEPDG analysis predicts distresses, such as faulting, transverse cracking, and smoothness in terms of International Roughness Index (IRI). A reliability term has been incorporated in MEPDG for each predicted distress type to come up with an analytical solution, which allows the designer to design a pavement with an acceptable level of distress at the end of design life. The chosen failure criteria are associated with this design reliability. The failure criteria and design reliability are also required inputs for the MEPDG analysis although the designer and the agency have the control over these values. The design can fail if the predicted distress is greater than the allowable amount or if the predicted distresses are unacceptable. In this study, the design reliability used for all projects was 90. Other default and revised performance criteria for both AC and PCC are shown in Table 2.

Project specific input parameters for the MEPDG AC and PCC pavement analysis in this study are given in Tables 3 and 4, respectively. Some of the inputs used in this study were similar for both AC and PCC pavements except for the structural details. Important traffic inputs like monthly and hourly truck distribution, axle load spectra, and truck class distribution were similar for both pavement types and were derived from an analysis of Weigh-In-Motion (WIM) and/or Automatic Vehicle Classification (AVC) data in Kansas.

**Table 2.** Failure Criteria for AC and PCC Pavements in MEPDG Analysis

| Distress Type | MEPDG default | Revised MEPDG |
|---|---|---|
| ***AC Pavements*** | | |
| Longitudinal Cracking (m/km) | 190 | 95 |
| Alligator Cracking (%) | 25 | 10 |
| Transverse Cracking (m/km) | 190 | 95 |
| Fatigue Fracture (%) | 25 | 10 |
| Permanent Deformation (AC only, mm) | 6.35 | 6.35 |
| Permanent Deformation (Total Pavement, mm) | 19.05 | 19.05 |
| ***PCC Pavements*** | | |
| Transverse Cracking (% slabs cracked) | 15 | 10 |
| Mean Joint Faulting (mm) | 0.12 | 0.10 |

**Table 3.** Common Input Parameters in MEPDG for AC and PCC Pavement Design

| Input | Design Value | | | | |
|---|---|---|---|---|---|
| | (I-70 GE) | (I-70 DK) | (I-70 SN) | (K-7 JO) | (I-70 WB) |
| ***General Information*** | | | | | |
| Pavement construction date | Nov, 90 | Oct, 92 | Oct, 93 | Sep, 95 | May, 01 |
| ***Traffic*** | | | | | |
| Initial two-way AADTT | 1,656 | 2,790 | 1,800 | 968 | 3,690 |
| No. of lanes in design direction | 2 | 2 | 2 | 2 | 2 |
| Traffic growth factor (%) | 1.2 | 3.5 | 3 | 6.7 | 3.3 |
| Design lane width (m) | 3.7 | 4.3 | 3.7 | 3.7 | 3.7 |
| ***Treated Subgrade*** | | | | | |
| Subgrade type | N/A | LTSG | LTSG | LTSG | LTSG |
| Subgrade modulus (MPa) | N/A | 311 | 311 | 311 | 311 |
| Unit weight ($kN/m^3$) | N/A | 19.0 | 19.0 | 19.0 | 19.0 |
| Poisson's ratio | N/A | 0.4 | 0.4 | 0.4 | 0.4 |
| ***Compacted Subgrade*** | | | | | |
| Subgrade soil type | A-6 | A-6 | A-7-6 | A-7-6 | A-7-6 |
| Subgrade Modulus (MPa) | 67 | 48 | 43 | 50 | 49 |
| Plasticity index, PI | 15.8 | 26 | 25.7 | 19.9 | 20.15 |
| Percent passing %200 sieve | 71.8 | 78.1 | 93.3 | 94.3 | 96.7 |
| % passing #4 sieve | 100 | 100 | 100 | 100 | 100 |
| $D_{60}$ (mm) | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |

**Table 4.** Specific Input Parameters in MEPDG for AC and PCC Pavement

| Input | Design Value | | | | |
|---|---|---|---|---|---|
| | (I-70 GE) | (I-70 DK) | (I-70 SN) | (K-7 JO) | (I-70 WB) |
| ***AC Layer*** | | | | | |
| AC Layer thickness (cm) | 30 | 30 | 30 | 25 | 32 |
| Reference Temperature ($^oC$) | 20 | 20 | 20 | 20 | 20 |
| Total Unit Weight ($kN/m^3$) | 22.8 | 22.8 | 22.8 | 22.8 | 22.8 |
| Air Voids (%) | 4 | 4 | 4 | 4 | 4 |
| Poisson's ratio | 0.35 | 0.35 | 0.35 | 0.35 | 0.35 |
| ***PCC Layer*** | | | | | |
| PCC Layer thickness (cm) | 23 | 28 | 29 | 27 | 25 |
| Modulus of Rupture (MPa) | 4.8 | 4.3 | 3.3 | 3.7 | 4.7 |
| Material Unit Weight ($kN/m^3$) | 22.0 | 21.9 | 22.3 | 22.3 | 22.2 |
| Cement Type | I | I | I | I | I |
| Cement Content ($kg/m^3$) | 387 | 316 | 374 | 370 | 384 |
| Poisson's ratio | 0.20 | 0.20 | 0.20 | 0.20 | 0.2 |
| Aggregate Type | Limestone | Limestone | Limestone | Limestone | Limestone |
| Co-eff of thermal exp ($mm/mm/°Cx10^{-6}$) | 10 | 10 | 10 | 10 | 10 |
| Water-cement ratio (w/c) | 0.44 | 0.35 | 0.411 | 0.46 | 0.437 |

## Analysis and Results

### *Analysis Procedure*

Initial sections were designed using DARwin based on the 1993 AASHTO design guide. These sections were then reanalyzed using the new MEPDG software versions 1.0 and 1.1. No base layer was used. All sections were built directly on LTSG wherever applicable. The analysis was done at 90% reliability for all sections and the terminal IRI was 2.59 m/km. The terminal IRI value was chosen based on the limit imposed by the Kansas Department of Transportation (KDOT) Pavement Management System. The initial IRI was assumed to be 0.99 m/km and is the default value. The other default and revised performance criteria for the AC and PCC sections appear in Table 2 for both MEPDG versions. If the section passed all criteria for the smoothness (IRI) and other distresses, the thickness was reduced by 12.5 mm and the analysis was redone. This process was repeated until the section failed to pass in one of the failure criteria. The section that was eventually obtained was taken to be the equivalent MEPDG section. Tables 5 shows the results using MEPDG versions 1.0 and 1.1 as well as the AASHTO 1993 guide outputs.

**Table 5.** AC and PCC Thickness Using MEPDG Versions and AASHTO 1993 Guide

| | I-70 GE | | | I-70 DK | | | I-70 SN | | | K-7 JO | | | I-70 WB | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1.0 | 1.1 | | 1.0 | 1.1 | | 1.0 | 1.1 | | Def. | Rev. | | Def. | Rev. |
| | 1993* | ME[1]/ ME[2] | ME[1]/ ME[2] | 1993* | ME[1]/ ME[2] | ME[1]/ ME[2] | 1993* | ME[1]/ ME[2] | ME[1]/ ME[2] | 1993* | ME[1]/ ME[2] | ME[1]/ ME[2] | 1993* | ME[1]/ ME[2] | ME[1]/ ME[2] |
| AC Sections | | | | | | | | | | | | | | | |
| SM-9.5T (PG70-28) | 3.8 | 3.8/3.8 | 3.8/3.8 | 3.8 | 3.8/3.8 | 3.8/3.8 | 3.8 | 3.8/3.8 | 3.8/3.8 | 3.8 | 3.8/3.8 | 3.8/3.8 | 3.8 | 3.8/3.8 | 3.8/3.8 |
| SM-19A (PG70-28) | 6.4 | 6.4/6.4 | 6.4/6.4 | 6.4 | 6.4/6.4 | 6.4/6.4 | 6.4 | 6.4/6.4 | 6.4/6.4 | 6.4 | 6.4/6.4 | 6.4/6.4 | 6.4 | 6.4/6.4 | 6.4/6.4 |
| SM-19A (PG64-22) | 20.3 | 12.5/14 | 7.6/15 | 20.3 | 8.8/11.4 | 10/15 | 20.3 | 8.8/10 | 8.9/14 | 15.2 | 7.6/10 | 5/7.6 | 22.5 | 11/12.5 | 11.4/15 |
| TOTAL AC (cm) | 30.5 | 23/24 | 18/25 | 30.5 | 19/21.6 | 20.2/25 | 30.5 | 19/20.2 | 19.1/24 | 25.4 | 17.8/20.2 | 15.2/18 | 32.7 | 21.6/23 | 21.6/25 |
| LTSG | None | None | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| PCC Sections | | | | | | | | | | | | | | | |
| PCCP (cm) | 23 | 22/22 | 20/20 | 28 | 20.3/22 | 22/22 | 29.2 | 24.1/25 | 25/25 | 26.7 | 22/22 | 22/23 | 25 | 27/39 | 27/39 |
| LTSG (cm) | None | None | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |

* 1993 AASHTO [1] M-EPDG Default Criteria [2] M-EPDG Revised Criteria

The AC mixtures were designed following the superpave mix design procedure. The AC layer was assumed to be built in three distinct sublayers (surface, binder and base) for all projects. The designation of the mixtures, shown in Table 5, follows the KDOT nomenclature for the Superpave mixes. In Kansas, a superpave mix is designated as "SM." The numeric following SM indicates the nominal maximum aggregate size (NMAS) in the mix in mm. The alphabet "A" immediately after that specifies the aggregate gradation i.e. it indicates that the gradation passed above the maximum density line in the finer sand sizes. Thus the gradation is finer and it allows inclusion of more sandy materials in the mix. In general, the surface mixture is a 9.5 mm NMAS mixture, and the binder and base mixtures are 19 mm mixtures. For each sublayer, a distinct performance grade (PG) binder is used. The binder grades, shown in parentheses in Table 5, were selected using the LTPPBind software. In this analysis, the thickness of the base layer was altered mostly because of the lower temperature requirement for the PG binder in other layers.

All PCC sections in this study are JPCP with 4.6 m joint spacing. The dowel diameter was estimated as one eighth of the slab thickness. Sections where the PCC slab thickness was less than 25 cm, 32 mm diameter dowels were used. The concrete strength, in terms of 28-day modulus of rupture, ranged from 3.3 MPa to 4.8 MPa as shown in Table 4. All PCC sections have 3.7 m lane widths with tied concrete shoulders except the I-70 Dickinson county project, which is a specific pavement study (SPS) section of the Long Term Pavement Performance (LTPP) program. This SPS-2 section has a widened lane of 4.3 m with tied PCC shoulders. In Kansas, a 101.6 mm thick Portland Cement Treated Base (PCTB) is generally used under the PCC slab. In the 1993 AASHTO design guide method, Kansas assumes a k value of 110 MPa for this base on a lime-treated subgrade (LTSG). However, the resulting slab thickness remains unchanged. In this study, all sections were reanalyzed with a 101.6 mm PCTB using the MEPDG methodology to find the effect of this base layer. The modulus of PCTB was assumed to be 3,450 MPa and the unit weight of the PCTB layer materials was 21.2 kN/m$^3$. The PCC slab thickness was reduced by 12.7 to 25.4 mm because of the PCTB base.

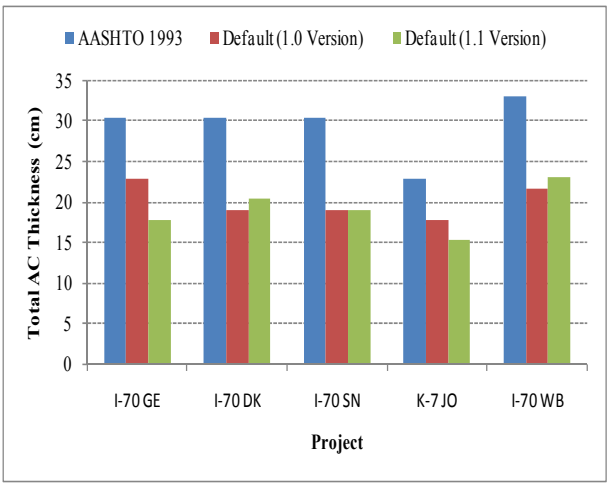
### *Comparison of AASHTO 1993 and MEPDG Software Versions Using Default Criteria*

Default performance criteria for both AC and PCC sections are given Table 2. The same default criteria have been used for both versions of the MEPDG software. Comparison has been made for the AC and PCC sections separately.

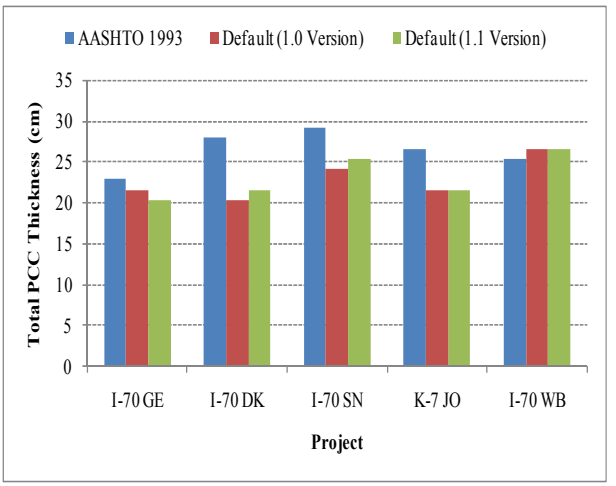
**AC Sections.** The results obtained using AASHTO 93 and two different versions of MEPDG based on default criteria are shown in Figure 1(a). Both versions of the software gave thinner AC when compared to the AASHTO 1993/DARwin thickness. Thickness reduction was 75 mm for the I-70 Geary County, 115 mm for the I-70 Dickinson, Shawnee, and Wabaunsee County projects and 50 mm for the K-7 Johnson County when MEPDG version 1.0 was used. Thickness reduction was usually higher with version 1.1. It was 127 mm for the I-70 Geary County project, 102 mm for the I-70 Dickinson County project, 114 mm for the I-70 Shawnee and Wabaunsee County projects, and 75 mm for the K-7 Johnson County project.

The least reduction in thickness was 50 mm for K-7 Johnson County while using version 1.0 and the highest reduction was 127 mm for I-70 Geary County for version 1.1. Thickness reduction is equal or higher using version 1.1 as compared to version 1.0 except for the I-70 Dickinson County, which showed a decrease in reduction from 114 mm to 102 mm. Both

versions showed considerable AC thickness reduction when compared to the 1993 AASHTO Guide/DARwin thickness.




(a) AC Sections




(b) PCC Sections

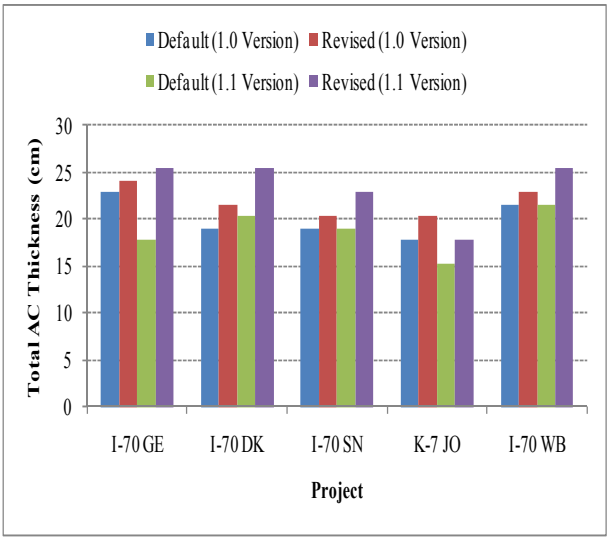
**Figure 1.** Comparison of total AC thickness using AASHTO 1993 and MEPDG software.

**PCC Sections**. Figure 1(b) shows the comparison of PCC slab thickness using the AASHTO 1993 Guide/DARwin and MEPDG versions 1.0 and 1.1 based on default criteria. The reduction in thickness is less as compared to the AC sections. The reduction in thickness was 12.5 mm for I-70 Geary County, 75 mm for I-70 Dickinson County, and 51 mm for I-70 Shawnee and K-7 Johnson Counties, respectively while using version 1.0 of the software. Reduction in thickness was 25 mm for I-70 Geary County, 64 mm for I-70 Dickinson County, 38 mm for I-70 Shawnee and 50 mm for K-7 Johnson County while using version 1.1. The highest reduction was 75 mm. On the other hand, there is an increase in thickness by 12.5 mm for the I-70 Wabaunsee County project while using both versions. In general, version 1.0 resulted in equal or lower thickness reduction as compared to version 1.1 except for the I-70 Geary County.

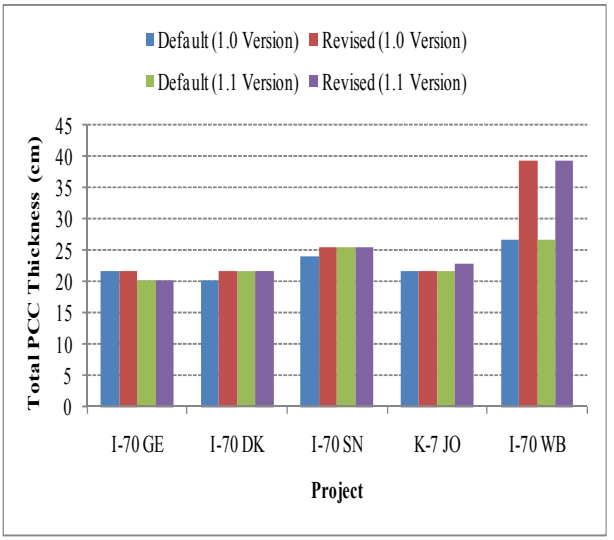
### *Effect of Changing Performance Criteria on AC and PCC Thicknesses*

**Effect on AC Thickness.** Figure 2(a) shows the analysis results using both versions of the MEPDG software based on performance criteria indicated in Table 2. As expected, there is an increase in AC thickness requirement due to change in performance criteria for both versions. The increase in thickness was 12.5 mm for I-70 Geary, Shawnee, and Wabaunsee Counties, and 25 mm for I-70 Dickinson and K-7 Johnson Counties. The highest increase in thickness was 75 mm for I-70 Geary County, and the least was 25 mm for K-7 Johnson County. The effect of changing performance criteria was more pronounced for version 1.1 of the MEPDG software.

**Effect on PCC Thickness.** The effect of changing performance criteria on the thickness of the PCC sections is illustrated in Figure 2(b). All sections have showed no or 12.5 mm increase in thickness except I-70 Wabaunsee, which has showed 127 mm increase in thickness for both software versions. The significant increase in thickness of I-70 Wabaunsee was due to faulting, which may be due to high truck traffic on this section compared to other sections. In general, the PCC sections showed less sensitivity to the changes in performance criteria.




(a) AC Sections




(b) PCC Sections

**Figure 2.** Effect of changing performance criteria on the thickness of AC sections

## Conclusions

Based on this study, the following conclusions can be made:

- For AC sections, the MEPDG procedure resulted in much thinner sections when compared to the sections obtained following the 1993 AASHTO design guide methodology.
- MEPDG software version 1.1 resulted in higher or equal AC thickness reduction as compared to version 1.0 except for I-70 Dickinson County.
- All PCC sections except I-70 Wabaunsee County have shown a reduction in thickness while using both versions of the MEPDG software as compared to the AASHTO 1993 design guide.
- MEPDG software version 1.0 resulted in equal or higher PCC slab thickness reduction except for the I-70 Geary County project.
- AC sections are more sensitive to changes in performance criteria as compared to the PCC sections.
- The effect of changing performance criteria is more pronounced for the AC sections while using version 1.1 of the MEPDG software.

## References

American Association of State Highway and Transportation Officials (AASHTO). *Guide for Design of Pavement Structures*. Washington, D.C., 1986.

American Association of State Highway and Transportation Officials (AASHTO). *Guide for Design of Pavement Structures*. Washington, D.C., 1993.

National Cooperative Highway Research Program (NCHRP). *Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures*. Final Report for Project 1-37A. Part 1, Chapter 1. Washington, D.C: NCHRP, Transportation Research Board, National Research Council, 2004.

Timm, D., B. Birgisson, and D. Newcomb. Development of mechanistic-empirical pavement design in Minnesota. In *Transportation Research Record, Journal of the Transportation Research Board, No. 1626*. TRB, National Research Council, Washington, D.C., 1998, pp. 181-188.

Yoder, E. J., and M.W. Witczak. *Principles of Pavement Design.* New York: John Wiley and Sons, 1975.

# The Impact of Weigh-in-Motion Measurement Error on Mechanistic-Empirical Pavement Design Guide Reliability

Syed Waqar Haider[1], and Ronald S. Harichandran[2]

[1]Assistant Professor, Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824; PH (517) 353-9782; FAX (517) 432-1827; email: syedwaqa@egr.msu.edu
[2]Professor and Chairperson, Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824:e-mail: harichan@egr.msu.edu

## ABSTRACT

Axle load spectra have a significant impact on predicted pavement performance. At the design stage, it is typically assumed that axle load spectra as measured by weigh-in-motion (WIM) systems are accurate. In fact, the quality of WIM-based data has inherent uncertainties due to inaccuracy and systematic bias in measurements. This paper investigates the impact of WIM measurement errors on axle load spectra (ALS) and quantifies the effects of these errors on design reliability. In the new Mechanistic-Empirical Pavement Design Guide (M-E PDG), the reliability procedure was developed based on the assumption that variability in performance prediction is approximately the same as the observed performance of the pavement sections used to calibrate the performance models. In this analysis, each distress type was approximated by a normal distribution and therefore, two parameters (mean and standard deviation) were determined to represent the expected value and associated variability by employing Monte Carlo simulation. The results show that the M-E PDG reliability analysis can compensate for negative axle load measurement bias for most of the distresses. However, a lower tolerance for negative bias needs to be enforced in order to ensure that both flexible and rigid pavements have the design reliability against cracking, especially for thinner pavements. While most of the findings further reinforce existing concepts, the study provides a systematic overview of WIM data accuracy and calibration needs and the impact of associated uncertainties on the pavement design process.

## INTRODUCTION

Traffic characterization plays a vital role in the analysis and design of pavements in the new Mechanistic-Empirical Pavement Design Guide (M-E PDG) mainly because of the associated high degree of uncertainty. In addition, the effects of traffic-related inputs are more pronounced because the M-E PDG directly uses traffic counts and axle load spectra to determine pavement damage and subsequent pavement performance prediction through transfer functions. The traffic data required by the M-E PDG includes: (a) traffic volume adjustment factors (hourly adjustment factor (HDF), monthly adjustment factor (MAF), vehicle class distribution, and traffic growth factor); (b) axle load spectra for various axle configurations, and; (c) general

traffic inputs (number of axles per truck, axle configuration, and wheelbase. etc.) (NCHRP Project 1-37A 2004; NCHRP Project 1-39 2005). Among the three traffic-related inputs, axle load spectra (ALS) require extensive efforts in data collection and several sensitivity studies have shown that ALS significantly affect pavement performance (Haider et al. 2008; Tran and Hall 2007). In practice, traffic data collection is carried out by a combination of sensor technologies which include weigh-in-motion (WIM) systems, automatic vehicle classifiers (AVC), and automated traffic recorders (ATRs). Each of these technologies has inherent uncertainty in the measurements. For axle load measurements, two approaches are commonly adopted: static and WIM scales. While static weighing has better accuracy, a number of practical issues and inefficiencies accompany this practice. The most important limitation of static weighing is the amount of axle load data that can be collected. On the other hand, since pavement design and analysis require a large and representative amount of traffic data, state-of-the-art WIM technology is popular because of its ability to collect traffic data continuously, although with less accuracy. In addition, WIM system errors introduced by the sensor technology, environmental effects, pavement condition, and other factors need to be considered to address concerns about measurement accuracy (Prozzi et al. 2008).

In most analysis and design scenarios, it is typically assumed that axle load spectra are measured accurately and extensively. However, measurement errors in the axle loads can seriously affect the pavement loading (FHWA 1998) and predicted performance (Haider et al. 2010). In order to compensate for several sources of variation, the M-E PDG uses the concept of reliability in pavement design. Even though mechanistic concepts provide a more accurate and realistic methodology for pavement analysis and design, a practical method to consider the uncertainties in design inputs and construction is needed so that a new or rehabilitated pavement could be designed for a desired reliability level. Within the context of the M-E PDG, the reliability of a given design is the probability that the performance of the pavement predicted for that design will be satisfactory over the design life (Darter et al. 2005). Because of the error associated with predicted pavement distresses and smoothness using transfer functions, the actual distress could be lower or higher than the average expected value. The error distribution for a distress measure is a function of the following sources of variation (Darter et al. 2005): (a) errors in estimation of traffic loadings, (b) variation in climatic conditions over time, (c) layer thicknesses and material property variations during construction, and (d) prediction model errors.

The M-E PDG reliability procedure was developed assuming that variability in performance prediction is approximately the same as the variability in performance prediction for the pavement sections used to calibrate the performance models (Darter et al. 2005). The procedure is based on analysis of the predicted versus measured distresses. In addition, after considerable analysis with various distributions, it was determined that the error in prediction of distresses is approximately normally distributed (NCHRP Project 1-37A 2003). Thus, the likely variation of the distress measures around the expected prediction can be represented by the mean and its standard deviation. The standard deviations of distresses are functions of the variability associated with the data used to calibrate the models. Therefore, it will be useful to investigate the effects of ALS and associated errors on the M-E PDG design

reliability for different structural and functional distresses in flexible and rigid pavements. The main objective of this paper is to investigate the effects of axle load measurement errors on the design reliability for different structural and functional distresses. The M-E PDG (version 1.0) software was used for all the analyses.

## WEIGH-IN-MOTION AXLE LOAD MEASUREMENT ERRORS

There are several concerns regarding the accuracy and calibration of WIM systems. Frequently, users of WIM systems infer static weights from dynamic measurements. The actual difference between static and instantaneous dynamic weight is often treated as an integral part of the WIM error (Davies and Sommerville 1987). WIM errors are made up of three components: (a) actual static and dynamic force differences; (b) dynamic force measurement errors; and (c) static load measurement errors. These WIM errors are mainly caused by three major factors (Prozzi et al. 2008): (a) roadway factors—smoothness and longitudinal/transverse profile; (b) vehicular factors—speed, acceleration, tire conditions, load and suspension type; and (c) environmental factors—wind, water and temperature. All of these factors can cause random error even if the WIM equipment is calibrated properly. In addition, a drift in calibration factors is likely to take place between periods of routine or recurring calibrations resulting in biased axle load measurements. The WIM system accuracy can be measured in terms of the relative difference between WIM and static weights. The relative WIM error can be expressed by (Bergan et al. 1995; Davies and Sommerville 1987).

$$\varepsilon = \frac{\text{WIM weight} - \text{static weight}}{\text{static weight}} \times 100 \tag{1}$$

where; WIM weight = load measured by a WIM scale for a given axle type, static weight = load measured on a static scale for the same axle type.

This relative error is commonly referred to as measurement error for a WIM scale and depends on its accuracy. Further, this accuracy will vary for different types of WIM sensor technologies. WIM measurement error generally follows a normal distribution (Bergan et al. 1995) with a zero mean (no bias) and a standard deviation $\sigma_\varepsilon$, i.e.,

$$\varepsilon = \frac{X' - X}{X} \sim N\left(0, \sigma_\varepsilon^2\right) \tag{2}$$

where

$X'$ = load measured on a WIM scale for an axle configuration
$X$ = load measured on a static scale for the same axle configuration
$\sigma_\varepsilon$ = standard deviation characterizing the accuracy of the WIM scale

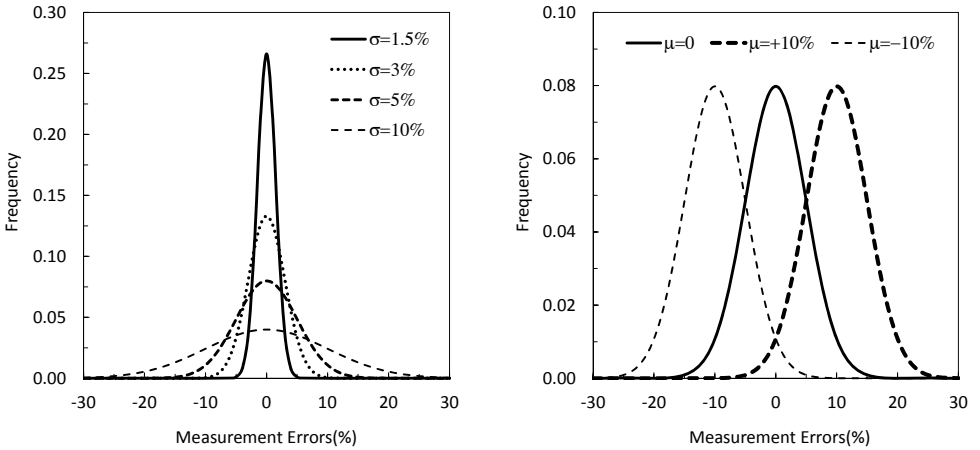
Typically, the axle load distribution measured by a WIM scale is bimodal for single and tandem axle configurations (Haider and Harichandran 2007), and can be represented by the following equation:

$$X' \approx f'(x) = \left( p_1 \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x-\mu_1)^2/2\sigma_1^2} + p_2 \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(x-\mu_2)^2/2\sigma_2^2} \right) \quad (3)$$

By knowing the distributions of axle loads from WIM scales and the associated error, the distribution of the corresponding axle loads that should be obtained from a static scale can be estimated from the following relationships:

$$X = \frac{X'}{\varepsilon + 1} \quad (4)$$

The measurement error ($\varepsilon$) entails two independent components according to the nature of error occurrence (Davies and Sommerville 1987): random error and systematic error. Both types of errors can be associated with internal measurement errors, or real static/dynamic weight discrepancies resulting from external factors (roadway, vehicle, and climate). More formally, the random error reflects stochastic fluctuations of a measurement (in either direction) from the true value, and is inherent in measurements because of the inability of the sensor to capture the true weight precisely. On the other hand, systematic errors persistently generate inaccuracies along one direction mainly because of inadequate calibration or drifts with time (Prozzi et al. 2008). For example, in Michigan, the WIM data showed that on average, the systematic error may vary between 5 to 10% in either direction within a year; however, this variation is site specific depending upon the sensor type and truck volume. Michigan Department of Transportation (MDOT) monitors the gross vehicle weight of class 9 trucks over time (300 days per year) to detect any such drift and performs sensor calibration every year. When the WIM scales are properly calibrated and installed in a sound and smooth pavement with normal traffic and environmental conditions, only random errors will occur and Equation (2) will hold. In this case the standard deviation of the distribution ( $\sigma_\varepsilon$ ) is a measure indicative of the WIM scale accuracy. As an example, Figure 1a illustrates the random error distributions of typical WIM sensors (single load cell, bending strain quartz and piezo) having varying accuracies (Bergan et al. 1995). Higher uncertainty is associated with less accurate sensor technology.




(a) Random error with no bias (b) Random and systematic error

**Figure 1. WIM error distributions**

A systematic bias in the measured loads may occur because of inaccurate initial calibration or drifts over time. When systematic error is present, the mean of the measured axle load distribution will shift to either side depending on whether the bias is positive or negative. Figure 1b shows the distribution shift because of negative and positive bias of the WIM scale by ±10% with a random error of $\sigma_\varepsilon$=5%. It should be noted that a systematic error will occur together with random error, while random error can arise alone if the WIM is properly calibrated.

Equation (4) can be used to estimate the true ALS given the WIM ALS (used in practice) and the expected error distribution [normal distribution with $\mu$=0 and $\sigma_\varepsilon$ when the WIM is calibrated properly (no bias), and normal distribution with $\mu$=$\alpha$ and $\sigma_\varepsilon$ when the WIM is not properly calibrated (with bias)]. The true ALS those were simulated in this study are presented next.

## AXLE LOAD SPECTRA WITH MEASUREMENT ERRORS

In this study, a typical ALS measured by a WIM scale was adopted as a reference from the specific pavement sections (SPS-1) of the long-term pavement performance (LTPP) study. The WIM scale was calibrated according to the LTPP protocols and the ALS data was based on continuous data collection over a full year. A bimodal distribution was fitted to the reference ALS. Once the parameters of the bimodal mixture distribution were estimated, a cumulative distribution function (CDF) was obtained through

$$F_{X'}(x) = P(X' < x) = \int_{-\infty}^{x} \left( p_1 \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x-\mu_1)^2/2\sigma_1^2} + p_2 \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(x-\mu_2)^2/2\sigma_2^2} \right) dx \quad (5)$$

Observations of $X'$were simulated by first generating uniform random numbers between 0 and 1, $u_i$ , and then transforming them through

$$x_i' = F_{x'}^{-1}(u_i) \quad (6)$$

The $x_i'$ will then have the CDF $F_{x'}(x)$. The error distribution, $N(\alpha, \sigma_\varepsilon)$, was simulated by assuming different values of $\alpha$ (bias) and $\sigma_\varepsilon$ (accuracy). The values of systematic error (bias) included 0, ±10%, ±20% and ±30% while random error ( $\sigma_\varepsilon$ ) value of and 10% were adopted (Bergan et al. 1995; FHWA 1998). Subsequently, the static (true) ALS were generated by using Equation (4). The random error was fixed at 10% while the systematic error was changed. Figure 2 shows ALS for realizations with various systematic errors. The numbers in the legend show the $\mu$ and $\sigma$ values of the error distributions used to simulate the ALS.

## INFLUENCE OF ALS ERRORS ON DESIGN RELIABILITY

As mentioned before, the M-E PDG reliability procedure was developed based on the assumption that variability in performance prediction is approximately the same as

the observed performance of the pavement sections used to calibrate the performance models. For example, the reliability analysis of bottom-up fatigue cracking (designated as alligator cracking) in flexible pavement is based on the empirically determined standard error of the measured (observed) cracking and can be estimated using (NCHRP Project 1-37A 2003):

$$SE_{AC} = 0.5 + \frac{12}{1 + e^{1.308 - 2.949 \times \log D}} \tag{7}$$

where

$SE_{AC}$ = standard error of estimate for measured alligator cracking

$D$ = predicted damage for bottom-up cracking

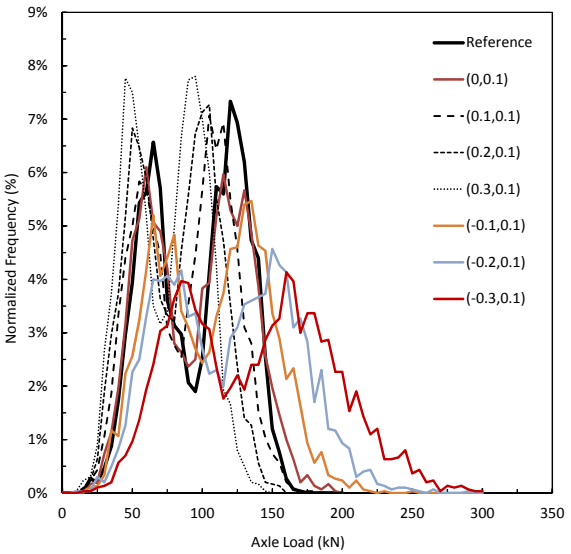



**Figure 2. Simulated axle load spectra with error**

The standard error of the estimate includes all sources of variation related to prediction, including (NCHRP Project 1-37A 2003): (a) errors associated with material characterization parameters, (b) errors related to assumed traffic and environmental conditions, and (c) model errors associated with the prediction equation

In the M-E PDG, Equation (7) is used to perform the reliability analysis for predicted bottom-up cracking in flexible pavements. It is assumed that the standard error determined using Equation (7) is the associated uncertainty to the expected cracking predicted by the M-E PDG calibrated model using mean inputs. The alligator cracking is estimated for the design level of reliability at each period by using:

$$Crack_{AC}^{R} = \overline{Crack}_{AC} + SE_{AC} \times Z_R \tag{8}$$

where

$Crack_{AC}^{R}$ = cracking level at reliability level R

$\overline{Crack}_{AC}$ = expected bottom-up cracking estimated using a deterministic model with average input values (corresponds to a 50% reliability level)

$SE_{AC}$ = standard error of estimate obtained from calibration

$Z_R$ = standard normal deviate (0 mean and 1 standard deviation) for selected reliability level R ($Z_R \approx 2$ for 95% reliability)

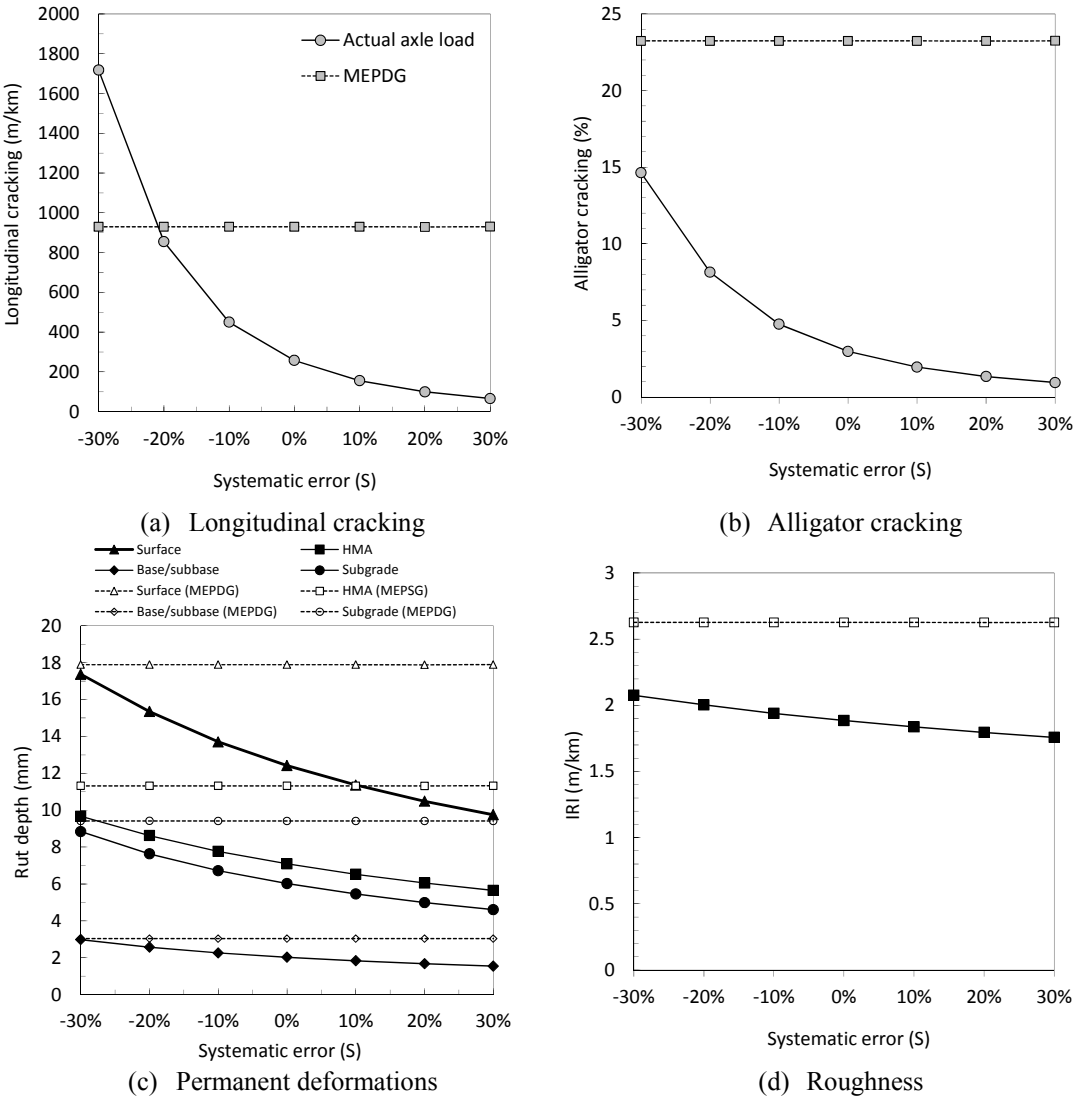
Similar procedures were adopted in the M-E PDG for reliability analyses of other distresses in both flexible and rigid pavements. In this study 500 ALS were simulated each for WIM (erroneous) and static (true) ALS at each combination of random and systematic errors. The simulation assisted in determining the distribution of each distress. Each distress was approximately normally distributed and therefore, two parameters (mean and standard deviation) were determined to represent the expected value and associated variability for each distress. The expected value of each distress was used to predict the standard error of the estimate (for example, Equation (7) was used for alligator cracking in flexible pavements) due to the reference WIM ALS. The 95% reliability using the M-E PDG procedure (NCHRP Project 1-37A 2003) for each distress was estimated using Equation (8) for alligator cracking in flexible pavements and similar equations for other distress types. In addition, 95$^{th}$ percentiles were estimated from the distribution of predicted distresses due to static ALS. In summary, each distress type was predicted at 95% reliability using the M-E PDG procedure for the reference ALS from WIM, and the 95$^{th}$ percentile distress levels also were estimated for all types of distresses predicted due to static ALS. A comparison of distress levels for each distress type at different systematic errors was made. The design reliability analyses results are discussed in the next section for each pavement type.

**Flexible pavements.** Figure 3 shows the magnitude of predicted distresses at 95% design reliability using the actual (static) ALS and the M-E PDG procedures for flexible pavement. The longitudinal cracking at the M-E PDG 95% design reliability will be exceeded if the negative bias is more than 20% (see Figure 3a). All other distress types (see Figures 3b through 3d) at the M-E PDG 95% reliability are conservative for the range of systematic errors considered in this study.

**Rigid pavements.** Figure 4 shows the magnitude of predicted distresses at 95% design reliability using the actual (static) ALS and the M-E PDG procedures for rigid pavement. The transverse cracking at the M-E PDG 95% design reliability level will be exceeded if the negative bias is more than about 15% (see Figure 4a). The faulting at the M-E PDG 95% reliability level is conservative for the range of systematic errors considered in this study (see Figure 4b). Also, in case of roughness, the M-E PDG reliability procedure under-predicted the roughness (IRI) when the negative bias is more than 20% (see Figure 4c).

The M-E PDG 95% design reliability distress values were determined for the reference WIM ALS. Therefore, the predicted distresses were not affected by the systematic error. In summary, the results of reliability analyses demonstrate that systematic error in WIM measurement can have a significant effect on pavement

design. The impact of erroneous WIM loads is much more serious on cracking distresses in both flexible and rigid pavements than the other distress types. The pavements can be significantly under-designed if the WIM measurements have high negative bias. It is therefore important to periodically calibrate WIM scales, and if drifts in calibration tend toward a negative bias between calibration times, then it is more conservative to calibrate the scales with a slightly positive bias.




(a) Longitudinal cracking

(b) Alligator cracking

(c) Permanent deformations

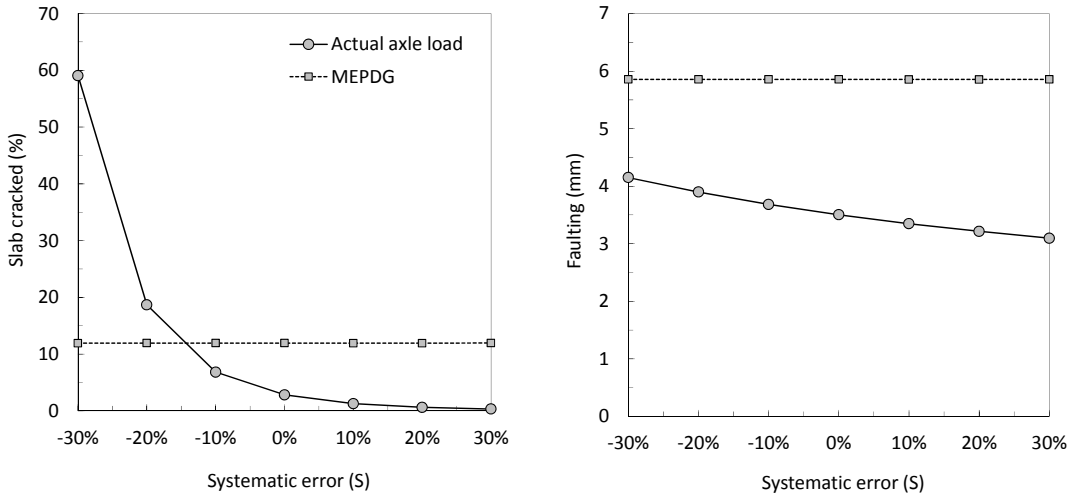
(d) Roughness

**Figure 3. Effect of WIM error on flexible pavement design reliability**

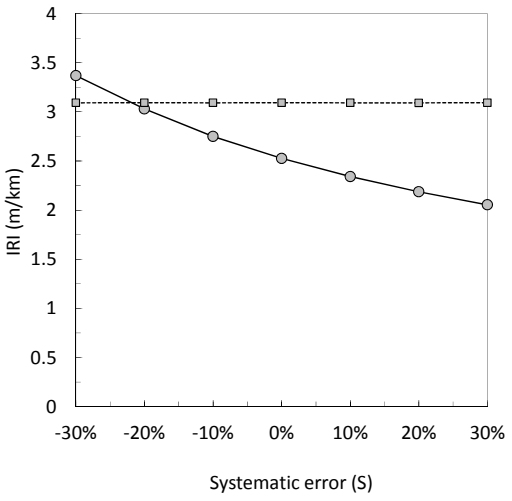
## CONCLUSIONS

Reliable and accurate axle load data are essential for pavement analysis and design. Errors in axle load measurements over time can have serious consequences on pavement design reliabilities. The results presented in this paper investigate the effect of axle load bias due to WIM measurement errors on the design reliability as used in

the M-E PDG analysis and design procedure. Typical axle loads are measured by WIM technology, which may yield two types of errors, random and systematic. Random error is a function of WIM accuracy (depending on the technology used), while systematic error is associated with equipment calibration. The effects of systematic error for a given random error (accuracy) on both flexible and rigid pavements design reliabilities were studied. The results show that cracking performance for both pavement types is significantly affected by negative measurement bias in axle loads. The M-E PDG reliability analysis can compensate for negative measurement bias for most of the distresses. However, a lower tolerance for negative bias needs to be enforced in order to ensure that both flexible and rigid pavements have the design reliability against cracking. WIM scales should be calibrated periodically to prevent a high negative bias.




(a) Transverse cracking (b) Joint faulting




(c) Roughness

**Figure 4. Effect of WIM error on rigid pavement design reliability**

## REFERENCES

Bergan, A., Berthelot, C., and Taylor, B. "Effect of Weigh In Motion Accuracy on Weight Enforcement Efficiency."

Darter, M., Khazanovich, L., Yu, T., and Mallela, J. (2005). "Reliability Analysis of Cracking and Faulting Prediction in the New Mechanistic-Empirical Pavement Design Procedure." *Transportation Research Record: Journal of the Transportation Research Board*, 1936(-1), 150-160.

Davies, P., and Sommerville, F. (1987). "Calibration and accuracy testing of weigh-in-motion systems." *Transportation Research Record*, 1123, 122-126.

FHWA. (1998). "WIM Scale Calibration: A Vital Activity for LTPP Sites " *TechBrief FHWA-RD-98-104*.

Haider, S. W., Buch, N., and Chatti, K. (2008). "Evaluation of M-E PDG for Rigid Pavements—Incorporating the State-of-the-Practice in Michigan." San Francisco, California, USA.

Haider, S. W., and Harichandran, R. S. (2007). "Relating Axle Load Spectra to Truck Gross Vehicle Weights and Volumes." *ASCE Journal of Transportation Engineering*, 133(12), 696-705.

Haider, S. W., Harichandran, R. S., and Dwaikat, M. B. (2010). "The Effect of Axle Load Measurement Errors on Pavement Performance and Design Reliabilities." *CD ROM, 89th Annual Meeting of Transportation Research Record*.

NCHRP Project 1-37A. (2003). "Appendix BB: Design Reliability." *Final Report: Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures*, ARA, inc., ERES division, 505 west University Avenue, Champaign, Illinois 61820.

NCHRP Project 1-37A. (2004). "Appendix AA: Traffic Loading." *Final Report: Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures*, ARA, inc., ERES division, 505 west University Avenue, Champaign, Illinois 61820.

NCHRP Project 1-39. (2005). "Traffic Data Collection, Analysis, and Forecasting for Mechanistic Pavement Design." *Report 538*, National Cooperative Highway Research Program, National Research Council, Transportation Research Board, Washington, D.C.

Prozzi, J., Hong, F., and Leung, A. (2008). "Effect of Traffic Load Measurement Bias on Pavement Life Prediction: A Mechanistic-Empirical Perspective." *Transportation Research Record: Journal of the Transportation Research Board*, 2087, 91-98.

Tran, N. H., and Hall, K. D. (2007). "Development and Influence of Statewide Axle Load Spectra on Flexible Pavement Performance." *Transportation Research Record*, 2037, 160-114.

# COMPARISON OF PREDICTED DISTRESSES BETWEEN DIFFERENT INPUT LEVELS USING M-E PDG FOR CRC PAVEMENT

Mohamed Elfino, Ph.D., P.E.[1], Harikrishnan Nair, Ph.D[1], Gary Crawford[2], Jagan Gudimettla, P.E[2]

1. Virginia Department of Transportation,1401 East Broad Street, Richmond, VA 23219
2. Federal Highway Administration, Office of Pavement Technology,1200 New Jersey Avenue SE, Washington DC 20590

## ABSTRACT

The Mechanistic Empirical Pavement Design Guide (M-E PDG) provides a number of new approaches for characterizing materials to be used in 21st Century pavement design. The M-E PDG software uses numerical models to analyze traffic, climate, subgrade and laboratory measurements of material properties to predict the performance of various pavement designs over their entire service life. A key element of the mechanistic design approach is the prediction of the response of the pavement materials, and thus of the pavement itself. The M-E PDG provides three hierarchical input levels: Level 1 is site/project specific with actual tests resulting in higher accuracy, Level 2 from less than optimal testing or by correlations and Level 3 from the agency database or user selected default values. Since obtaining Level 1 material inputs require considerable testing effort and cost, Level 1 data for M-E PDG analysis may not be possible all the times. In such situations, Level 2 or 3 inputs can still be used to perform M-E PDG analysis and predict design life of pavement.

The objective of this study is to compare M-E PDG predicted distresses of different input level data from a recently constructed Continuously Reinforced Concrete Pavement (CRCP) project at I-64 Battlefield blvd, Chesapeake, Virginia. The original design of this pavement was done using the 93 AAHSTO design method several years before the construction. This paper also documents the efforts associated with the laboratory and field procedures used for level 1 input characterization. The FHWA Mobile Concrete Laboratory (MCL) had performed all the concrete materials characterization for level 1 input on this project. M-E PDG software (version 1.1) is used to obtain the predicted distresses for the pavement structure using the same thickness obtained by AASHTO method. This paper also discusses the difference in predicted distresses when level 1 and level 3 traffic data was used. Results shows that there were significant differences in predicted distresses between Level 1 and Level 3 concrete properties inputs for M-E PDG analysis depending upon the type of Level 3 input used.

## INTRODUCTION

Despite potential higher initial cost, constructing highways using CRCP could provide a number of benefits such as lower long-term life cycle cost, low maintenance, reduced water penetration because of fewer joints and tighter transverse cracks, smoother pavements over time, ability to handle heavier truck loading and volumes, easier to rehab and can become a successful composite pavement because of the absence of joints. Structural design of CRC Pavement requires the determination of the amount of reinforcement to control transverse cracks and the slab thickness to resist traffic loads. In CRC pavement, no transverse joints are provided and volume changes in concrete due to environmental loading are controlled by transverse cracks, which are kept tight by longitudinal steel. One of the most important variables in reducing wheel load stress is crack width. The objective of CRCP design is to develop a pavement system that will keep crack width as small as possible. Transverse cracks with tight widths provide good load transfer and keep wheel load stresses low and thus extend pavement life. Tight crack widths also keep water from reaching steel reinforcement and prevent corrosion. Aggregate interlock will be lost if transverse cracks are not kept tight and further results in poor load transfer and higher wheel load stresses. Slab thickness, setting temperature, coefficient of thermal expansion (CTE), drying shrinkage of concrete, and longitudinal steel characteristics have effects on the widths of transverse cracks.

The AASHTO Guide for Design of Pavement Structures [1] is the primary document used to design new and rehabilitated highway pavements in the United States today. In AASHTO design guide, the overall serviceability of the pavement is quantified by the present serviceability index (PSI, range = 0 to 5), which is a composite performance measure combining cracking, patching, and other distresses. Roughness is the dominant factor governing PSI and is therefore the principal component of performance under this measure. The perceived deficiencies of the empirical design approach were the motivation for the mechanistic–empirical methodology developed in NCHRP Project 1-37A [2]. Structural responses (i.e., stresses, strains, and deflections) are mechanistically calculated (by using multilayer elastic theory or finite element methods) on the basis of material properties, environmental conditions, and loading characteristics. Thermal and moisture distributions are also mechanistically determined (with the enhanced Integrated Climate Model (EICM)). These responses are used as inputs in models to individually predict distresses. The models were calibrated by using data from the Long-Term Pavement Performance (LTPP) database for conditions representative of the entire United States.

## OBJECTIVES

The objectives of this study are to (i) Compare M-E PDG predicted distresses for different input levels of materials and traffic data from a recently constructed Continuously Reinforced Concrete Pavement (CRCP) project at I-64 Battlefield blvd, Chesapeake, Virginia and (ii) Documents the efforts associated with the laboratory and field procedures used for level 1 materials characterization and traffic.

## I-64 BATTLEFIELD BLVD PROJECT

The project was located on I-64 in the Chesapeake region of Virginia. The CRCP lanes are approximately 1.1 miles long. Four CRCP lanes in each direction are regular travel lanes and one CRCP lane is designated for High Occupancy Vehicles (HOV). Average annual daily traffic (AADT) for this location is 112,400 with 6% trucks. The existing Jointed Reinforced Concrete Pavement (JRCP) was recycled (Figure 1 (a), (b)) and cement treated to construct the base layer for the CRCP lanes. Open Graded Drainage Layer (OGDL) with 75 mm (3") thickness was paved above the Cement Treated Recycled Concrete (CTRC). The CRCP placed on top of the OGDL was 325 mm (13") thick and #7 rebar was used as longitudinal and #5 bar for the transverse direction reinforcement. Figure 2 shows CRCP project on I-64. During the construction of this project traffic was moved to the outside lanes, allowing for construction to take place in the center lanes. An on-site batch plant provided a steady supply of concrete, uninterrupted by traffic. The design included 0.7% steel, with #7 bars located at mid-slab. The original design of this pavement was done using the 93 AAHSTO design method several years before the construction. Structural layer properties obtained based on 93 AASHTO design are given in Table 1.


Figure 1(a). During JRCP Pavement breaking (b) JRCP Pavement removing

Table 1. Structural Layer Design Based on 93 AASHTO Method

| Layer | Type | Material | Thickness Inches (mm) |
|---|---|---|---|
| 1 | Concrete | CRCP | 13 (325) |
| 2 | Asphalt | Open Graded Drainage Layer (OGDL) | 3 (75) |
| 3 | Base | Cement Treated Recycled Concrete (CTRC) | 6 (150) |
| 4 | Subgrade | A-3 & A-2-4 | Semi infinite |

## MATERIAL CHARACTERIZATION

### Concrete Materials Characterization

The FHWA Mobile Concrete Laboratory (MCL) performed all the concrete materials characterization required for level 1 input on this project. Concrete samples for level 1 material input (for the CRC layer) were collected by the MCL on three days of production. Table 2 shows the fresh concrete properties from the three days of sampling. The results indicate that the fresh concrete properties were relatively consistent.


Figure 2. I-64 Battlefield Blvd Project construction

Table 2. Fresh Concrete Properties

| Properties | Day 1 | Day 2 | Day 3 |
|---|---|---|---|
| | | | |
| Fresh Concrete Properties | | | |
| Unit weight of Concrete, pcf | 146 | 147 | 147 |
| Slump, inches | 2 | 1.5 | 1.25 |
| Temperature of Concrete, $^{o}$ F | 85 | 83 | 86 |
| Air Content, % | 5.5 | 5.5 | 5 |

### Compressive Strength

Three cylinders were tested in compression according to the ASTM C 39 at 7, 14, 28 and 90 days with average values of 2813, 3863, 4366, and 4545 psi (Table 3).

Compressive strength is a level 2 input for CRCP in M-E PDG and could be used to assess the difference in designs for level 1 versus level 2 inputs. Virginia specifications require 3000 psi at 28 days.

Table 3. Compressive strength Results

| Properties | Day 1 | Day 2 | Day 3 |
|---|---|---|---|
| Average Compressive Strength, psi | | | |
| 7 Day | 2650 | 3041 | 2749 |
| 14 Day | 3624 | 3700 | 4265 |
| 28 Day | 3880 | 4240 | 4980 |
| 90 Day | 4007 | 3935 | 5693 |

1 psi = 6.9 k Pa, 1 pcf =16 Kg/$m^3$, 1inch = 25 mm, $^oC = (^oF-32)1.8$

**Modulus of Rupture (Flexural Strength)**

Third point flexural strength testing was performed according to ASTM C78 at 7, 14, 28, and 90 days. Results of flexural strength (based on three test data) were on the average of 598,708, 835, and 890 psi respectively (Table 4). These values are much higher than the 650 psi used in the 93 AASHTO design. The high flexural strength is possible due to the presence of high percentage of slag (50%) in the mix.

Table 4. Flexural Strength Results

| Properties | Day 1 | Day 2 | Day 3 |
|---|---|---|---|
| Average Modulus of Rupture, psi | | | |
| 7 Day | 602 | 677 | 516 |
| Standard Deviation | 28 | 67 | 16 |
| Coefficient of Variation | 5% | 10% | 3% |
| 14 day | 730 | 766 | 628 |
| Standard Deviation | 26 | 28 | 11 |
| Coefficient of Variation | 4% | 2% | 2% |
| 28 day | 919 | 785 | 802 |
| Standard Deviation | 27 | 17 | 36 |
| Coefficient of Variation | 3% | 2% | 5% |
| 90 day | 947 | 973 | 751 |
| Standard Deviation | 69 | 27 | 72 |
| Coefficient of Variation | 7% | 3% | 10% |

**Splitting Tensile Strength**

In M-E PDG, Splitting Tensile Strength is a level 1 input for CRCP for use in punchout models. Splitting tensile strength tests were performed according to ASTM C496 at 7, 14, 28 and 90 day with average values (based on three test data) of 307, 355, 401, and 481 psi respectively (Table 5).

Table 5. Splitting Tensile Strength results

| Properties | Day 1 | Day 2 | Day 3 |
|---|---|---|---|
| Average Split Tensile, psi | | | |
| 7 Day | 285 | 314 | 323 |
| Standard Deviation | 25 | 11 | 48 |
| Coefficient of Variation | 9% | 4% | 15% |
| 14 day | 341 | 345 | 397 |
| Standard Deviation | 50 | 17 | 40 |
| Coefficient of Variation | 15 | 5 | 10 |
| 28 day | 398 | 405 | 402 |
| Standard Deviation | 34 | 30 | 31 |
| Coefficient of Variation | 9% | 7% | 8% |
| 90 day | 470 | 485 | 490 |
| Standard Deviation | 3 | 23 | 13 |
| Coefficient of Variation | 1% | 5% | 3% |

**Modulus of Elasticity (MOE) and Poisson's Ratio**

Testing was performed at 14, 28, and 90 day old specimens according to ASTM C469 with average values (based on three test data) of 5.15, 5.38, and $5.75*10^6$ psi respectively (Table 6).

Table 6. Modulus of Elasticity Results

| Properties | Day 1 | Day 2 | Day 3 |
|---|---|---|---|
| Average Modulus of Elasticity, psi and Poisson's Ratio | | | |
| 14 Day | $5.13*10^6$ | $5.12*10^6$ | $5.20*10^6$ |
| 28 Day | $4.86*10^6$ | $5.41*10^6$ | $5.88*10^6$ |
| 90 Day | $5.77*10^6$ | $5.33*10^6$ | $6.15*10^6$ |
| Poisson Ratio | 0.20 | 0.19 | 0.21 |

**Coefficient of Thermal Expansion**

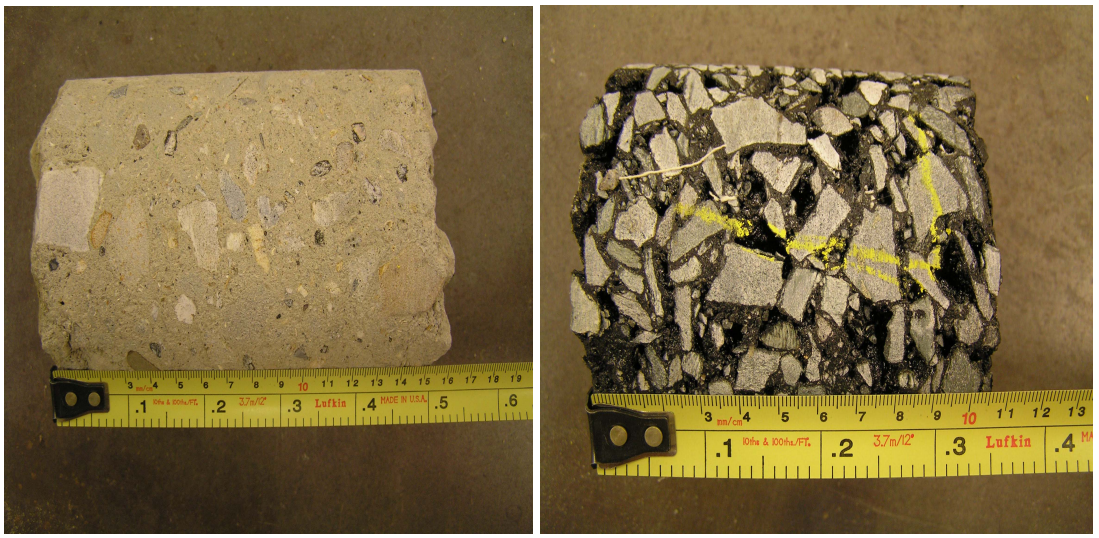
The coefficient of thermal expansion (CTE) of concrete is a property that can affect the performance of the concrete pavement and its service life and is one of the most important inputs in the M-E PDG. CTE affects the curling stresses and axial stresses and, as a consequence, affects the performance and serviceability of the pavement structure. The CTE influences early age cracking, fatigue cracking, faulting, and joint spalling (3). CTE value depends on the concrete composition, age and moisture state (4). The CTE can vary extensively among aggregates due to mineralogical differences. Even the same aggregate type can present different CTEs as a result of the differences in the mineralogical content (4). The CTE of specimens from three samples was determined after 90 days according to AASHTO TP 60. The average value was $5.77 * 10^{-6}$ in/in/ $^{o}$ F (Table 7).

Table 7. CTE Results

| Properties | Sample 1 | Sample 2 | Sample 3 |
|---|---|---|---|
| Coefficient of Thermal Expansion, CTE (90 day aged specimen) | | | |
| | $5.70 * 10^{-6}$ in/in/ $^{o}$ F | $5.82 * 10^{-6}$ in/in/ $^{o}$ F | $5.80 * 10^{-6}$ in/in/ $^{o}$ F |

## OGDL AND CTRC CHARACTERIZATION

In MEPDG, for level 1 characterization dynamic modulus testing is needed for OGDL. However, due to difficulty in testing cores (due to small height) dynamic modulus was calculated using the predictive model. The aggregate gradation used for OGDL is given in Figure 4. Performance Grade (PG 70-22) asphalt binder is used with 4.3 % ± 0.3 asphalt cement content. Cores were taken from both the asphalt treated Open Graded Drainage Layer (OGDL) and Cement Treated Recycled Concrete (CTRC) as shown in Figure 3.




**(A)** **(B)**

Figure 3 CTRC (Shown in A) and OGDL (shown in B) cores taken from I-64 Battlefield Blvd.

Elastic modulus testing is done on CTRC layer and the average modulus is found to be 700,800 psi. However, there was large variation in elastic modulus results possibly due to the sample variability (cement content variation) and testing difficulties due to the length of the core. Table 8 lists the variation in thickness and cement content for different samples from the field.

Table 8. Cement Treated Aggregate Data

| | Density lb/ft$^3$ (kg/m$^3$) | Thickness in (mm) | Cement (%) | Moisture Content (%) |
|---|---|---|---|---|
| Required (optimum) | 103.5 (1657) | 6 (152.4) | 4 | 14.7 |
| Average | 110.4 (1768) | 6.1 (154.94) | 8 | 13.6 |
| Maximum | 115.5 (1850) | 6.2 (157.48) | 9 | 16.1 |
| Minimum | 105.3 (1686) | 5.8 (147.32) | 4.9 | 12.3 |
| Standard Deviation | 2.4 | 0.1 | 1 | 0.8 |
| Coefficient of Variation | 2.2 | 1.8 | 12.9 | 6.1 |

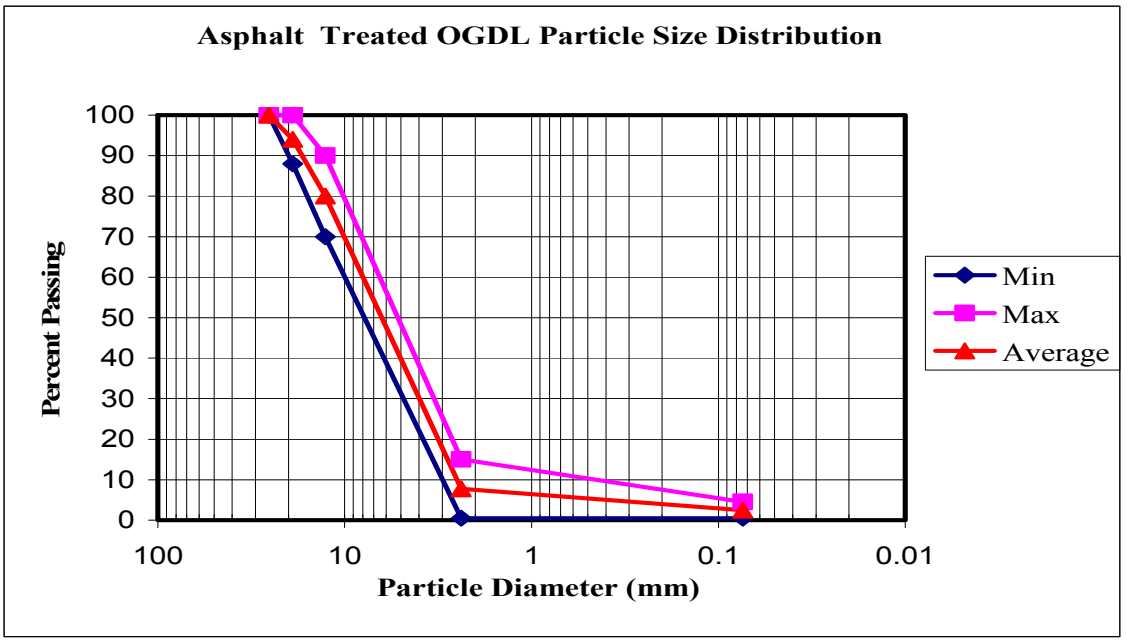



Figure 4 Open Graded Drainage Layer (OGDL) gradation.

## SUBGRADE CHARACTERIZATION

The results from resilient modulus (Mr) testing are required for Level 1 pavement design. In Level 1 input, the MEPDG allows for input of the regression constants of the constitutive model for a specific subgrade soil. This ensures a more accurate assessment of the modulus during the analysis over the design period including seasonal variation and varying stress conditions. The resilient modulus test was performed in accordance with AASHTO T 307-99, Standard Method of Testing for Determining the Resilient Modulus of Soils and Aggregate Materials by the soils lab at the VDOT Materials Division. Samples of 72.5 mm (2.9 inch) in diameter and 145 mm (5.8 inch) in height were compacted at Optimum Moisture Content (OMC) and maximum dry density

(MDD) using a static compactor. These sample were loaded in accordance with AASHTO T 307, and the recoverable strains were measured using two external linear variable differential transducers.

Soil samples were collected from different locations to characterize the soil for resilient modulus testing. It was found that majority of the soil type fall under A-2-4 and A-3. The average resilient moduli were 6500 psi for A-2-4 soil and 7000 psi for A-3 soil. Table 9 presents soil index properties and CBR test results. Average CBR value found to be 4% and using M-E PDG suggestion correlation (level 2 input), $M_r$= 2555 ($CBR^{0.64}$), resilient modulus is found to be 6200psi, which is very close to results obtained using RM testing. However, for input level 3, design guide [2] suggests resilient modulus value ranges of 28000-37500 for A-2-4 soil, which is found to be very high for this project.

Table 9. Soil Index Properties and CBR results

| Location | Depth (inch) | Soil Classification AASHTO | #10 | #40 | #200 | LL (%) | PL (%) | Soaked CBR (%) | OMC (%) | MDD (pcf) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1010+00 | 0"-17" | A-2-4 | 94.8 | 72.7 | 11.7 | n/a | NP | n/a | 9.6 | 116.2 |
| 978+50 | 3"-24" | A-4 (0) | 99 | 97.5 | 76.9 | n/a | NP | 2.5 | 34.4 | 71.9 |
| 987+50 | 6" - 24" | A-2-4 | 98.9 | 92.3 | 13.8 | n/a | NP | n/a | 12.5 | 107.9 |
| 1001+00 | 14.5"- 24" | A-4 (3) | 99.9 | 98.7 | 65.9 | 22.6 | 13.9 | 7.3 | 12.8 | 118.9 |
| 1001+00 | 0" - 14.5" | A-2-4 | 96.4 | 81.9 | 25 | n/a | NP | 3 | 10 | 124.3 |
| 1001+00 | 0" - 14.5" | A-2-4 | 96.4 | 81.9 | 25 | n/a | NP | 3.2 | 10 | 124.3 |

## TRAFFIC CHARACTERIZATION

Another improvement offered by the M-E PDG is the use of traffic input based on the number of axles by type and loading which is known as load spectra, while eliminating the use of ESAL's. VDOT collected traffic data on East bound I-64 for 4 days by establishing a traffic count meter 0.01 mi from SR 168 Battlefield Blvd. Information from this data was used as Level 1 input for the M-E PDG analysis. Vehicle class distribution required for Level 1 input is obtained and summary of this data is presented in Table 10.

Basic Traffic Inputs for I-64 Project

- Initial Two-Way AADTT -7750
- Number of Lanes in Design Direction – 5
- Percent of trucks in design direction (%) – 50
- Percent of trucks in design lane (%)-60,
- Operational speed (mph) – 60

Table 10. Traffic Data: Vehicle class distribution

| | All Classes | Class1 | Class2 | Class3 | Class4 | Class5 | Class6 | Class7 | Class8 | Class9 | Class 10 | Class 11 | Class 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Day 1 | 51727 | 186 | 36681 | 10869 | 341 | 501 | 537 | 94 | 261 | 1621 | 57 | 53 | 11 |
| Day 2 | 49520 | 175 | 35163 | 10418 | 242 | 394 | 429 | 151 | 248 | 1671 | 73 | 46 | 21 |
| Day 3 | 53912 | 204 | 38347 | 11215 | 328 | 467 | 674 | 137 | 265 | 1716 | 49 | 57 | 16 |
| Day 4 | 51498 | 101 | 37170 | 10656 | 351 | 441 | 446 | 36 | 222 | 1494 | 62 | 56 | 17 |
| Average | 51664 | 167 | 36840 | 10790 | 316 | 451 | 522 | 105 | 249 | 1626 | 60 | 53 | 16 |

## CLIMATIC INPUTS

Following inputs were used for as climate data

- Latitude (degrees, minutes) - 36.54
- Longitude (degrees, minutes) -76.11
- Elevation (ft) – 69
- Annual Average Depth of Water Table (ft) - 10
- Weather Station – Norfolk International airport

## PREDICTED DISTRESS DIFFERENCES USING DIFFERENT LEVELS OF INPUT FOR CONCRETE MATERIALS

The PCC strength and stiffness inputs in M-E PDG consist of the following: (i) Modulus of rupture (flexural strength), MR (ii) Modulus of elasticity, $E_{PCC}$ (iii) Compressive strength, $f'_c$ (iv) Split tensile strength (as measured by the indirect tensile strength test), $f_t$. Depending on the input level, different amount of information is required as follows:

**Level 1**- Laboratory values of MR, $f_t$, and $E_c$ at 7, 14, 28, and 90 days determined using appropriate testing procedures. The ratio of 20-yr to 28-day strength is also required.

**Level 2-** Laboratory-determined values of compressive strength ($f'_c$) at 7, 14, 28, and 90 days and the 20-yr to 28-day strength ratio. The remaining strength parameters (MR, $f_t$, and $E_c$) are estimated using well established strength-to-strength and strength-to-stiffness correlations.

**Level 3-** Estimated 28-day compressive strength or modulus of rupture from historical data or other sources. Other inputs are calculated based on the projected strength using the appropriate correlations.

Since Modulus of Rupture (MR) value has a significant effect on the fatigue cracking potential of the PCC slab, for level 2 and 3 inputs it is estimated as $MR = 9.5 * (f'c)^{1/2}$ psi. The splitting tensile strength value has a significant effect on the accumulation of damage and the development of punchouts in CRCP. For Level 2 and 3

inputs, splitting tensile strength is estimated from MR value by multiplied by a factor of 0.67. MR value is estimated from compressive strength using the equation above.

Table 11 shows the various strength test combinations that can be used for analysis using Level 3 inputs. As shown in the table, there are four options to perform analysis using Level 3 inputs. At a minimum, 28 day flexural strength (Modulus of Rupture) or 28 day compressive strength information is necessary for Level 3 input. Twenty eight day $E_{PCC}$ is an optional input. If measured $E_{PCC}$ cannot be obtained (not entered), then the software estimates $E_{PCC}$ based on the 28 day flexural strength or compressive strength information. All four options shown in Table 11 were used in the analysis. Data collected from Production day 3 was selected for the analysis.

Table 11. Concrete Strength inputs for M-E PDG Analysis

| Level 3 inputs | | | |
|---|---|---|---|
| Option 1 | Option 2 | Option 3 | Option 4 |
| 28 Day Modulus of Rupture | 28 Day Modulus of Rupture | 28 Day Compressive strength | 28 Day Compressive strength |
| 28 Day Modulus of Elasticity | - | 28 Day Modulus of Elasticity | - |

**MEPDG ANALYSIS RESULTS**

M-E PDG analysis was performed using national field calibration coefficients. All other inputs such as climate, traffic and material inputs for Open Graded Asphalt Layer, Cement Treated Layer and soil were the same. Table 13 shows the performance criteria and distress limits used for the analysis. Limits for minimum and maximum crack spacing were established based on Virginia Department of Transportation (VDOT)'s past experience with CRC pavements.

Table 13 Performance Criteria and Distress Limits

| Performance Criteria | | | |
|---|---|---|---|
| Item No. | MEPDG Information | Limit | Reliability |
| 1 | Initial IRI (in/mile) | 63 | |
| | Performance criteria/Rigid Pavement | | |
| 2 | Terminal IRI (in/mile) | 110 | 95 |
| 3 | CRCP Existing Punchouts | 5 | 95 |
| 4 | Maximum CRCP Crack Width (in) | 0.04 | |
| 5 | Crack Load Transfer Efficiency (LTE %), minimum | 75 | |
| 6 | Minimum Crack Spacing (ft) | 3.0 | |
| 7 | Maximum Crack Spacing (ft) | 6.0 | |

The M-E PDG output for rigid CRCP pavement analysis predicts the IRI, number of punchouts per mile, crack width, and load transfer efficiency. Punchout is caused by loss of aggregate interlock at closely spaced cracks; after repeated loading the

reinforcing steel will rupture and a piece of the concrete will punch downward into the lower pavement layers (5). The M-E PDG predicts the maximum crack width that is expected during the pavement design life. Cracks are expected to be wider during winter months when the pavement material contracts due to cooler weather. Load transfer efficiency (LTE) for a CRCP pavement refers to the ability to transfer loads at the cracked locations; it is reported as the percentage of the load that is transferred. The higher the LTE (up to 100%) indicates a better performing pavement and lower LTE values indicate a poorer performing pavement. The IRI and distresses considered by the M-E PDG for CRCP pavements are interrelated; a large crack width can cause poor LTE in a pavement that can in turn lead to punchouts and ultimately a rougher pavement.

Table 14 presents results of the analysis with different concrete properties inputs. From Table 14 it can be seen that with Level 1 inputs for concrete properties, predicted distresses did not exceed the corresponding targets. However, Level 3 analysis with compressive strength inputs (both with MOE estimated option and actual MOE option) and Flexural strength inputs, the predicted distresses (punchouts and load transfer efficiency) will exceed the limits before the design life of the pavement is reached. The difference in distress prediction between the Level 3 analysis options may be due to the prediction relations used in the M-E PDG. As mentioned previously, when compressive strength is used in Level 3 analysis, flexural strength and split tensile are estimated based on this compressive strength from existing relationships built in the software and this estimated flexural strength and split tensile strength are used in the models. For the I-64 CRC project, the PCC achieved a very high 28 day flexural strength (800 psi in this example (production day 3)). When this flexural strength is used as level 3 input, the predicted distresses were lower compared to other inputs. In the Level 3 compressive strength input options, flexural strength will be estimated from the compressive strength (which in this case is 4980 psi). This estimated flexural strength (670 psi) would be much lower than 800 psi. So due to the lower flexural strength used in the models, the predicted distresses were significantly higher.

Result shows that based on M-E PDG analysis, this pavement structure will last for 40 years as compared to 30 years from AASHTO method. When Flexural strength alone is used as a level 3 input, the predicted design life is found to be 29 years. Predicted design life is found to be different for other level 3 input options (Compressive strength only= 24 years, Flexural strength and modulus of elasticity =25 years, compressive strength and modulus of elasticity = 20 years) as shown in Table 14. This implies that, the predicted distress magnitudes and thereby design life is greatly influenced by the quality of the inputs.

Table 14. Summary of Level 1 and Level 3 M-E PDG Predicted Distresses

| **Performance Criteria - Level 1** | Distress Target | Reliability Target | Distress Predicted | Predicted Design Life | Reliability Predicted | Acceptable |
|---|---|---|---|---|---|---|
| Terminal IRI (in/mi) | 110 | 95 | 63.4 | 40 | 98.92 | Pass |
| CRCP Punchouts (per mi) | 5 | 95 | 0 | | 98.98 | Pass |
| Maximum CRCP Crack Width (in) | 0.04 | | 0.0185 | | | Pass |
| Minimum Crack Load Transfer Efficiency (LTE%) | 75 | | 97.3 | | | Pass |
| **Performance Criteria - Level 3 - Flexural Strength Only** | Distress Target | Reliability Target | Distress Predicted | | Reliability Predicted | Acceptable |
| Terminal IRI (in/mi) | 110 | 95 | 65.5 | 29 | 97.44 | Pass |
| CRCP Punchouts (per mi) | 5 | 95 | 0.1 | | 95.9 | Pass |
| Maximum CRCP Crack Width (in) | 0.04 | | 0.0279 | | | Pass |
| Minimum Crack Load Transfer Efficiency (LTE%) | 75 | | 68.2 | | | Fail |
| **Performance Criteria - Level 3 - Flexural Strength and $E_{PCC}$** | Distress Target | Reliability Target | Distress Predicted | | Reliability Predicted | Acceptable |
| Terminal IRI (in/mi) | 110 | 95 | 63.1 | 25 | 99.999 | Pass |
| CRCP Punchouts (per mi) | 5 | 95 | 2.7 | | 65.83 | Fail |
| Maximum CRCP Crack Width (in) | 0.04 | | 0.0295 | | | Pass |
| Minimum Crack Load Transfer Efficiency (LTE%) | 75 | | 47.5 | | | Fail |
| **Performance Criteria - Level 3 – Compressive Strength Only** | Distress Target | Reliability Target | Distress Predicted | | Reliability Predicted | Acceptable |
| Terminal IRI (in/mi) | 110 | 95 | 75.7 | 24 | 88.44 | Fail |
| CRCP Punchouts (per mi) | 5 | 95 | 6.3 | | 43.19 | Fail |
| Maximum CRCP Crack Width (in) | 0.04 | | 0.028 | | | Pass |
| Minimum Crack Load Transfer Efficiency (LTE%) | 75 | | 47.5 | | | Fail |
| **Performance Criteria - Level 3 - Compressive Strength and $E_{PCC}$** | Distress Target | Reliability Target | Distress Predicted | | Reliability Predicted | Acceptable |
| Terminal IRI (in/mi) | 110 | 95 | 93.6 | 20 | 67.87 | Fail |
| CRCP Punchouts (per mi) | 5 | 95 | 15.4 | | 16.21 | Fail |
| Maximum CRCP Crack Width (in) | 0.04 | | 0.0323 | | | Pass |
| Minimum Crack Load Transfer Efficiency (LTE%) | 75 | | 47.5 | | | Fail |

## PREDICTED DISTRESS DIFFERENCES USING DIFFERENT LEVELS OF TRAFFIC INPUT

One of the key new features in M-E PDG is more detailed and accurate traffic characterization. The traffic information required by M-E PDG includes (i) traffic volume adjustment factors, (ii) axle load distribution factors, and (iii) general traffic

inputs. Traffic volume adjustment factors involve monthly adjustment, vehicle class distribution, hourly distribution, and traffic growth factors. Axle load distribution factors include axle load spectra by axle type (single, tandem, tridem, and quadruple axles). General traffic inputs include the number of axles per truck, the axle configuration, wheelbase characteristic, and traffic wander.

Weigh-in-motion (WIM) sensors are useful for collecting the large amount of traffic data necessary to develop inputs for the M-E PDG design methodology. To evaluate the influence of traffic inputs, the predicted pavement distresses for this pavement structure were compared in the M-E PDG analysis using different traffic level inputs. Separate M-E PDG runs were performed for each group using Virginia state specific traffic input factors and default inputs. This allowed a separate comparison of each traffic input factor. Level 1 concrete inputs were used for this analysis.

(i) All default traffic inputs
(ii) Site specific vehicle class distribution and default values for remaining traffic inputs
(iii) Site specific vehicle class distribution and state wide averages for number of axles per truck and monthly adjustment factors

The traffic inputs developed for analysis within the M-E PDG include the following: axle-load spectra, monthly adjustment factors, vehicle class distribution factors, and number of axles per truck. Statewide average monthly adjustment factors and average number of axles per truck were determined based on the collection of statewide WIM data (6). The default M-E PDG inputs for Average Number of Axles per Truck are shown in Table 15 along with the statewide WIM average axles per truck. Monthly Adjustment Factors from Statewide WIM Data are shown in Table 16.

Table 17 presents results of the analysis with different traffic inputs. From Table 17, it can be seen that there is no significant variation in M-E PDG predicted distresses when different levels of traffic inputs used even though there are several notable differences between the default traffic data inputs included in the M-E PDG software and statewide average traffic inputs. Previous studies have investigated the traffic data inputs for use with the M-E PDG in other states (7, 8, and 9). The results are mixed, however, with some studies reporting similar predicted pavement distresses using default versus site-specific traffic inputs while other studies show a significant difference. Sensitivity of the predicted pavement performance to various traffic input parameters should be further investigated to obtain a better understanding of the impact of different traffic parameters on the predicted pavement performance.

Table 15. Default and Statewide Average Number of Axles per Truck M-E PDG Inputs (Reference 6)

| | MEPDG Default Axles per Truck | | | | Statewide WIM Average Axles per Truck | | | |
|---|---|---|---|---|---|---|---|---|
| **Vehicle Classification** | **Single** | **Tandem** | **Tridem** | **Quad** | **Single** | **Tandem** | **Tridem** | **Quad** |
| 4 | 1.62 | 0.39 | 0.00 | 0.00 | 1.91 | 0.09 | 0.00 | 0.00 |
| 5 | 2.00 | 0.00 | 0.00 | 0.00 | 2.05 | 0.00 | 0.00 | 0.00 |
| 6 | 1.02 | 0.99 | 0.00 | 0.00 | 1.05 | 0.97 | 0.00 | 0.00 |
| 7 | 1.00 | 0.26 | 0.83 | 0.00 | 1.25 | 0.04 | 0.41 | 0.55 |
| 8 | 2.38 | 0.67 | 0.00 | 0.00 | 2.21 | 0.72 | 0.00 | 0.00 |
| 9 | 1.13 | 1.93 | 0.00 | 0.00 | 1.23 | 1.87 | 0.00 | 0.00 |
| 10 | 1.19 | 1.09 | 0.89 | 0.00 | 1.05 | 0.92 | 0.87 | 0.10 |
| 11 | 4.29 | 0.26 | 0.06 | 0.00 | 5.00 | 0.00 | 0.00 | 0.00 |
| 12 | 3.52 | 1.14 | 0.06 | 0.00 | 4.00 | 1.00 | 0.00 | 0.00 |
| 13 | 2.15 | 2.13 | 0.35 | 0.00 | 1.57 | 2.61 | 0.07 | 0.00 |

Table 16. Monthly Adjustment Factors (MAF) from Statewide WIM Data (Reference 6)

| | Vehicle Classification | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Month** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** |
| January | 0.95 | 0.88 | 0.83 | 0.54 | 0.99 | 1.05 | 0.93 | 1.01 | 0.99 | 1.02 |
| February | 1.01 | 0.94 | 0.98 | 0.70 | 1.01 | 1.04 | 1.00 | 1.02 | 1.04 | 0.69 |
| March | 1.15 | 1.16 | 1.11 | 1.06 | 1.14 | 1.14 | 1.16 | 1.09 | 1.09 | 1.34 |
| April | 1.18 | 1.23 | 0.93 | 0.76 | 1.10 | 1.12 | 1.23 | 1.13 | 1.10 | 1.71 |
| May | 1.11 | 1.07 | 0.94 | 0.76 | 1.10 | 1.13 | 1.10 | 1.13 | 1.13 | 1.11 |
| June | 0.80 | 0.82 | 0.98 | 1.46 | 0.84 | 0.80 | 0.85 | 0.84 | 0.79 | 1.11 |
| July | 0.78 | 0.89 | 0.96 | 1.28 | 0.80 | 0.73 | 0.83 | 0.79 | 0.72 | 0.78 |
| August | 0.83 | 0.89 | 1.01 | 1.23 | 0.82 | 0.76 | 0.78 | 0.77 | 0.75 | 0.65 |
| September | 0.83 | 0.93 | 1.03 | 1.13 | 0.85 | 0.78 | 0.82 | 0.80 | 0.83 | 0.97 |
| October | 1.12 | 1.19 | 1.12 | 1.13 | 1.09 | 1.10 | 1.03 | 1.12 | 1.10 | 1.02 |
| November | 1.14 | 1.05 | 1.06 | 0.98 | 1.13 | 1.18 | 1.13 | 1.18 | 1.18 | 0.92 |
| December | 1.09 | 0.95 | 1.04 | 0.98 | 1.13 | 1.17 | 1.14 | 1.11 | 1.27 | 0.69 |

(The default MAF factors are 1.00 for all months and vehicle classes)

Table 17. Summary of M-E PDG Predicted Distresses using different traffic inputs

| **Performance Criteria – Default traffic Data from M-E PDG** | Distress Target | Reliability Target | Distress Predicted | Reliability Predicted | Acceptable |
|---|---|---|---|---|---|
| Terminal IRI (in/mi) | 110 | 95 | 63.4 | 98.89 | Pass |
| CRCP Punchouts (per mi) | 5 | 95 | 0 | 98.67 | Pass |
| Maximum CRCP Crack Width (in) | 0.04 | | 0.0185 | | Pass |
| Minimum Crack Load Transfer Efficiency (LTE%) | 75 | | 95.8 | | Pass |
| **Performance Criteria – level-1 site specific vehicle class distribution (other values default)** | Distress Target | Reliability Target | Distress Predicted | Reliability Predicted | Acceptable |
| Terminal IRI (in/mi) | 110 | 95 | 63.4 | 98.92 | Pass |
| CRCP Punchouts (per mi) | 5 | 95 | 0 | 98.98 | Pass |
| Maximum CRCP Crack Width (in) | 0.04 | | 0.0185 | | Pass |
| Minimum Crack Load Transfer Efficiency (LTE%) | 75 | | 97.3 | | Pass |
| **Performance Criteria – Level-1 site specific vehicle class distribution And level 2 values for Monthly distribution and average axles per truck** | Distress Target | Reliability Target | Distress Predicted | Reliability Predicted | Acceptable |
| Terminal IRI (in/mi) | 110 | 95 | 63.4 | 98.92 | Pass |
| CRCP Punchouts (per mi) | 5 | 95 | 0 | 99 | Pass |
| Maximum CRCP Crack Width (in) | 0.04 | | 0.0185 | | Pass |
| Minimum Crack Load Transfer Efficiency (LTE%) | 75 | | 97.3 | | Pass |

## SUMMARY AND CONLCUSIONS

The M-E PDG hierarchical input levels are provided as options to utilize designer's knowledge of the input parameter. The Design-Bid-Build (DBB) has less flexibility for using level 1 inputs compared to Design-Build (DB). This is because of the time difference between the design and construction being longer for DBB. Additionally, DB affords to make changes at later stage or just before construction without increase in cost; on the contrary the designer may save a lot of money through optimization of the design using level 1 input data.

To implement the M-E PDG fully, users will need to develop databases of the required inputs and determine whether or not their usage results in a significant difference in the predicted condition as compared to default values. Users will also have to determine how the significance is determined. The following specific conclusions are derived from this study.

(i) Level 1 design requires site-specific input data. However, since construction plans are prepared at least two years in advance, it is necessary to have close

cooperation between the designer, and the contract preparation team for getting level 1 data.

(ii) Results shows that there were significant differences in predicted distresses between Level 1 and Level 3 concrete properties inputs for M-E PDG analysis depending upon the type of Level 3 input used.

(iii) Result shows that based on M-E PDG analysis, this pavement structure will last for 40 years as compared to 30 years (from AASHTO method) if level 1 concrete material property were used. When Flexural strength alone is used as a level 3 input, the predicted design life is found to be 29 years. Predicted design life is found to be different (ranges from 20-29) for other level 3 concrete material input options. This implies that, the predicted distress magnitudes and thereby design life is greatly influenced by the quality of the inputs.

(iv) Results shows that there were no significant differences in predicted distresses when different traffic level inputs used. When differences in traffic inputs are considered between site specific conditions and a more general (e.g., statewide) distribution, it is critical to assess the practical significance in addition to the statistical significance since considerable resources and personnel are required to gather site-specific load spectra.

(v) Site-specific information should be used when readily available. Also, local knowledge and experience should help determine when site-specific data must be collected and used for design.

## REFERENCES

1. Guide for Design of Pavement Structures. American Association of State Highway and Transportation Officials, Washington, D.C., 1993
2. ARA, Inc., ERES Consultants Division. (2008). Guide for the Mechanistic-Empirical Design of New & Rehabilitated Pavement Structures. NCHRP Project 1-37A. Transportation Research Board, Washington, DC.
3. Tanesi, J, Kutay, M.E, Ala Abbas., and Meininger, R. Effect of Coefficient of Thermal Expansion Test Variability on Concrete Pavement Performance as Predicted by Mechanistic–Empirical Pavement Design Guide. In Transportation Research Record: Journal of the Transportation Research Board, No. 2020, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 40–44.
4. Won, Moon. Improvements of Testing Procedures for Concrete Coefficient of Thermal Expansion. In Transportation Research Record: Journal of the Transportation Research Board, No. 1919, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 23–28.
5. Huang, Y.H. (2004). Pavement *Analysis and Design.* 2nd ed. Pearson–Prentice Hall, Upper Saddle River, NJ.
6. Brian Diefenderfer and Bryan Smith. Analysis of Virginia-Specific Traffic Data Inputs for Use with the Mechanistic-Empirical Pavement Design Guide. VTRC 10-R19. Virginia Transportation Research Council, Charlottesville, VA.

7. Li, J., Hallenbeck, M. E., Pierce, L. M., and Uhlmeyer, J. (2009). "Sensitivity of Axle-load spectra in Mechanistic-Empirical Pavement Design Guide for Washington State Department of Transportation." Presented at Transportation Research Board 88th Annual Meeting, Transportation Research Board, Washington, DC.
8. Timm, D. H., Bower, J. M., and Turochy, R. E. (2006). "Effect of Load Spectra on Mechanistic-Empirical Flexible Pavement Design." Transportation Research Record. No. 1947, pp. 146-154. Transportation Research Board, Washington, DC.
9. Tran, N. H., and Hall, K. D. (2007). "Development and Influence of Statewide Axle-Load Spectra on Flexible Pavement Performance." Transportation Research Record. No. 2037, pp. 106-114. Transportation Research Board, Washington, DC.

# Evaluation of the MEPDG Dynamic Modulus Prediction Models for Asphalt Concrete Mixtures

S. M. El-Badawy[1], Ph. D. Research Fellow, A. Awed[2], Graduate Research Assistant, and F. M. Bayomy[3], Ph. D., P.E., Professor

[1]National Institute for Advanced Transportation Technology (NIATT), University of Idaho, P.O. Box 441022, Moscow, Idaho 83844-1022; PH (208) 885-6818; FAX (208) 885-6608; email: elbadawy@uidaho.edu , Permanent Address: Mansoura University, Public Works Eng., Dept., Mansoura, Egypt
[2]Department of Civil Engineering, University of Idaho, P.O. Box 441022, Moscow, Idaho 83844-1022; PH (208) 885-6784; FAX (208) 885-6818; email: awed8141@vandals.uidaho.edu
[3]Department of Civil Engineering, University of Idaho, P.O. Box 441022, Moscow, Idaho 83844-1022; PH (208) 885-6784; FAX (208) 885-6608; email: bayomy@uidaho.edu

## ABSTRACT

HMA dynamic modulus (E*) is one of key inputs to the Mechanistic-Empirical Pavement Design Guide (MEPDG). There are two different E* models in the MEPDG; the NCHRP 1-37A viscosity-based model, and the NCHRP 1-40D, which is based on the binder shear modulus. This paper focuses on evaluating the influence of the binder characterization input level on the predicted E* in MEPDG. Laboratory E* tests were conducted on samples from 15 different HMA plant-produced mixtures. The shear modulus (G*) and phase angle (δ) for each binder were also determined in the laboratory. Results showed that MEPDG levels 1and 3 binder characterization inputs with both E* predictive models yielded E* values that are in excellent to fair agreement with laboratory measured E*. However, the 1-37A model showed better results than the 1-40D model. On the other hand, high bias in E* values was observed when level 1 binder characterization data was used.

## INTRODUCTION

The most important hot-mix asphalt (HMA) property influencing the structural response of a flexible pavement is the HMA dynamic modulus (E*). It is the primary stiffness property for the characterization of asphalt concrete (AC) mixtures in all of the hierarchical input levels of MEPDG. Critical stresses, strains, and deflections in the AC layer(s) are calculated as a function of the AC modulus (E*) using the pavement response model incorporated in MEPDG software. In the current MEPDG software (version 1.10), level 3 E* values at different temperatures and loading frequencies can be estimated using two different predictive models according to the user's choice. The first model is the NCHRP1-37A viscosity-based E* (ARA, 2004). The second model is the NCHRP1-40D binder shear modulus G*-based model. The first model was implemented in the MEPDG in its first release (Version 0.7) while the second one was added to the MEPDG in version 1.0 (Witczak

et al. 2006). These two models predict E* as a function of mixture volumetric properties, gradation, loading frequency and binder stiffness. For binder characterization, MEPDG utilizes three hierarchical input levels. In level 1 (same as level 2 in this case), laboratory measured binder G* and phase angle (δ) at an angular frequency of 10 rad/s at different temperatures are required. In level 3, the only input required is the binder Superpave performance grade (in case of Superpave binders). On the other hand, the binder viscosity grade is required input for binders classified based on viscosity while the binder penetration grade is the required input for binders classified according to penetration.

The objective of this paper was to compare the predicted E* by these two models to the actual measured E* values and study the effect of binder characterization on such prediction. The predicted E* values using each model were conducted based on two different sets of binder characteristics. One set is based on level 1 G*-δ binder data at 10 rad/s angular loading frequency and at least three different temperatures as per MEPDG input requirements. The other set is based on MEPDG level 3 default binder properties based on superpave performance binder grade. A secondary objective of this paper is to compare the prediction accuracy of the two E* predictive models incorporated in MEPDG based on the data collected from 15 different HMA mixtures commonly used in Idaho.

## MEPDG E* PREDICTIVE MODELS

The two E* models adopted in MEPDG were developed by Witczak and his colleagues. These two models are the most commonly used when direct laboratory E* measurements are unavailable (Harran and Shalaby 2009; Ceylanet et al. 2009). Details of these models are presented next.

### NCHRP 1-37A viscosity-based E* model

This model was implemented in the first version of the MEPDG (version 0.7). The model was developed based on 2750 measured E* data points from 205 different HMA mixtures, including modified and unmodified binders, that have been periodically collected by Witczak and his colleagues since 1969 (Andrei et al. 1999). It predicts E* at different temperatures as a function of the mix aggregate gradation, volumetric properties, loading frequency and binder viscosity. The model is presented by Equation 1 (ARA 2004):

$$\log_{10} E^* = -1.249937 + 0.02923\rho_{200} - 0.001767(\rho_{200})^2 - 0.002841\rho_4 - 0.058097V_a - 0.82208\frac{V_{beff}}{V_{beff}+V_a} + \frac{3.871977 - 0.0021\rho_4 + 0.003958\rho_{38} - 0.000017(\rho_{38})^2 + 0.00547\rho_{34}}{1+e^{(-0.603313-0.313351\log f-0.393532\log\eta)}} \quad (1)$$

where E*=dynamic modulus, $10^5$ psi; $\rho_{34}$=cumulative % retained on the ¾ in sieve; $\rho_{38}$= cumulative % retained on the 3/8 in sieve; $\rho_4$=cumulative % retained on the No. 4 sieve; $\rho_{200}$=% passing the No. 200 sieve; $V_{beff}$=effective binder content, % by volume; $V_a$=air void content, %; f=frequency of loading, Hz; and η=viscosity at the age and temperature of interest, $10^6$ Poise.

## NCHRP 1-40D G*-based E* model

The main disadvantage of the 1-37A model presented above is that it characterizes the binder in terms of conventional viscosity rather than the dynamic shear modulus and phase angle of the binder (Garcia and Thompson 2007; Ceylan et al. 2009). The binder G* and δ are commonly used as a part of the superpave performance grade (PG) binder specification. Thus, to overcome this disadvantage, the MEPDG flexible pavement research team incorporated, in addition to the 1-37A model, another E* predictive model which characterizes the binder in terms of G* and δ. This was done as a part of the NCHRP 1-40D (02) project which is the Technical Assistance to NCHRP and NCHRP Project 1-40A: Versions 0.9 and 1.0 of the M-E Pavement Design Software. This model is a modified version of the Bari and Witczak's E* predictive model originally developed in 2005 (Witczak et al. 2007). It is implemented in the MPEDG since version 1.0. The E* database used in this model development contains 7400 data points from 346 mixtures. This database included the data used for the development of the 1-37A model. The model is presented below:

$$\log_{10} E^* = 0.02 + 0.758\left(|G_b^*|^{-0.0009}\right)\times\left(\begin{array}{l}6.8232 - 0.03274\rho_{200} + 0.00431\rho_{200}^{\ 2} + 0.0104\rho_4 - 0.00012\rho_4^{\ 2} \\ + 0.00678\rho_{38} - 0.00016\rho_{38}^{\ 2} - 0.0796V_a - 1.1689\left(\frac{V_{beff}}{V_a + V_{beff}}\right)\end{array}\right) \quad (2)$$
$$+\frac{1.437 + 0.03313V_a + 0.6926\left(\frac{V_{beff}}{V_a + V_{beff}}\right) + 0.00891\rho_{38} - 0.00007\rho_{38}^{\ 2} - 0.0081\rho_{34}}{1 + e^{(-4.5868 - 0.8176\log|G_b^*| + 3.2738\log\delta)}}$$

where $|G_b^*|$=dynamic shear modulus of binder (G*), psi; and δ= phase angle of the binder, degrees. All other variables are as previously defined in Equation 1.

The presented models (Eq. 1 & 2) follow the form of a segmoidal function. Both models relate E* to mix gradation parameters, volumetric properties, frequency of loading and binder properties. As mentioned earlier, the main difference between the two models, is the parameters included for binder characterization.

## MEPDG E* prediction methodology

As discussed before, both presented E* predictive models are function of the binder characteristics. There are two levels for binder input in MEPDG; level 1 and level 3 (level 2 is the same as level 1).

For level 1 binder characterization, MEPDG requires the G* and δ of the binder (aged at RTFO condition) at different temperatures and one angular frequency of 10 rad/s. The software then uses the following relationship to compute the viscosity at different temperatures as a function of G* and δ ( ARA 2004).

$$\eta = \frac{G^*}{10}\left(\frac{1}{\sin\delta}\right)^{4.8628} \quad (3)$$

where G* = binder complex shear modulus, Pa; δ =binder phase angle, degrees; and η= viscosity, Pa.s.

Once this step is completed, the ASTM viscosity-temperature relationship is established (Eq. 4). Then A and VTS parameters are determined by conducting linear regression using Equation 4 (ARA 2004).

$$\log\log\eta = A + VTS\log T_R \quad (4)$$

where η=viscosity, cP; $T_R$=temperature in Rankine at which the viscosity was estimated; and A= regression intercept VTS=regression slope of the Viscosity-Temperature Susceptibility.

The above relationship is then used directly to estimate the binder viscosity at the temperature of interest and then use the 1-37A model (Eq. 1) for E* computation. For the 1-40D Model, once the A and VTS are determined, the following equations (Eq. 5 to 10) are used to compute $G_b$* and $\delta_b$ at the temperature and frequency of interest in order to compute the E* at these temperatures and frequencies (Witczak et al. 2007):

$$\log\log\eta_{f_s,T} = A' + VTS'\log T_R \quad (5)$$

$$A' = 0.9699 * f_s^{-0.0527} * A \quad (6)$$

$$VTS' = 0.9668 * f_s^{-0.0575} * VTS \quad (7)$$

$$f_s = f_c/2\pi \quad (8)$$

$$\delta_b = 90 - 0.1785 * \log(\eta_{f_s,T})^{2.3814} * (f_s)^{(0.3507+0.0782VTS')} \quad (9)$$

$$|G_b^*| = 1.469 * 10^{-9} * \log(\eta_{f_s,T})^{12.0056} * f_s^{0.7418}(\sin\delta)^{0.6806} \quad (10)$$

where A′=adjusted A (adjusted for loading frequency); VTS′=adjusted VTS (adjusted for loading frequency); $f_s$= loading frequency in dynamic shear loading mode as used in the DSR test to measure $|G_b^*|$ and δ, Hz; $f_c$=loading frequency of a dynamic loading in "compression" mode (as used in the E* test of HMA mixtures), Hz; and $\eta_{fs,T}$=viscosity of the asphalt binder as a function of both loading frequency in DSR ($f_s$) and temperature ($T_R$), cP.

Equations 5 through 10 were developed based on asphalt binder properties database containing 8940 data points from 41 different virgin and modified asphalt binders (Bari&Witczak 2007; Witczak et al. 2007; Bari&Witcak, 2006). Finally, E* at any temperature and frequency can then be calculated using Equation 2.

It must be noted that the G*-based E* predictive model was developed based on estimated, rather than laboratory measured, G* and δ at the same temperature and frequency of E* from default A and VTS values (based on conventional binder characterization testing) (Far et al. 2009).

For Level 3, binder input, the program uses its internal default values of A and VTS for the selected binder grade. Then it follows the previous procedure explained for Level 1 to predicts E* either from Eq. 1 or Eq. 2 as selected by the user.

In summary, the above analysis indicates that the E* prediction methodology, in MEPDG using either the 1-37A or 1-40D E* predictive models, is based on the ASTM A-VTS regression parameters for binder characterization. The 1-37A model, estimates the binder viscosity as a function of temperature (no influence of frequency on binder viscosity) through Equation 4. On the other hand, the 1-40D model

estimates G* and δ at different temperatures and frequencies form A and VTS through a series of regression equations (Eq. 5 through 10).

It is important to note that the A and VTS used in the development of both E* predictive models are the default values in the MEDPG which were based on conventional viscosity binder testing data. Some researchers questioned the validity of the typical (default) A and VTS values in MEPDG to superpave performance grade binders, since the Superpave binders use DSR data (Harran and Shalaby 2009, Birgisson et al. 2005, Dongre′ et al. 2005).

## MATERIALS AND METHODS

### HMA mixtures tested

A total of 15 different plant produced asphalt concrete mixtures commonly used by Idaho Transportation Department (ITD) in pavement construction projects in the state of Idaho were used in the analysis. These mixtures contain five different superpave performance grade binder types, varied mix aggregate gradation, and volumetric properties.

### Specimen preparation for dynamic modulus testing

For each investigated mixture, two replicate samples from each loose mix were compacted in a Servopac gyratory compactor to achieve 150 mm diameter and 170 mm high cylindrical specimen. Dynamic modulus specimens were cored from the center of the gyratory specimens and the ends were sawed to obtain a cylindrical specimen of 100 mm diameter and 150 mm high in accordance to the AASHTO PP 60-09 procedures.

### HMA dynamic modulus testing

Dynamic modulus tests were performed in the University of Idaho Laboratory using the Simple Performance Tester (SPT). All tests were conducted according to AASHTO TP 62-07. All tests were run on two replicate samples at temperatures of 4.4, 21.1, 37.8, and 54.4 $^{o}$C and loading frequencies of 0.1, 0.5, 1, 5, 10, and 25 Hz at each temperature. Three on-specimen vertical LVDT's were used to monitor the axial deformation of each specimen. The total number of E* measurements for the 15 mixtures is 720 points.

### Binder dynamic shear rheometer testing

Dynamic Shear Rheometer (DSR) tests were conducted on the five PG binders used in the mixtures to determine binder G* and δ. The DSR tests were run according to AASHTO T315-06 (AASHTO 2006). All tests were performed at the same temperature and frequency sweep of the E* testing. Before testing, all binders were aged using Rolling Thin Film Oven (RTFO) to simulate mix aging during mixing and field compaction. The tested binder grades are PG 58-28, PG 64-28, PG 70-28, PG 64-34, and PG 76-28.

## DATA RESULTS AND ANALYSIS AND

### E* prediction based on level 3 binder data

For Level 3 binder data, typical default A and VTS values at RTFO condition for each investigated binder were taken from MEPDG. For the 1-37A model, Equation 3 was used to compute binder viscosity at different temperatures, then E* values were estimated at different temperatures and frequencies by Equation 1. For the 1-40D model, Equations 5 through 10 were used to compute $G_b$* and δ, at the same temperature and frequency values used in DSR and E* testing, based on the binder A and VTS. Then Equation 2 was used to estimate E* values at the same temperatures and frequencies.

### E* prediction based on level 1 binder data

For each investigated binder, equation 3 was used to calculate binder viscosity at different temperatures using data corresponding to angular loading frequency of 10 rad/s. Then the A and VTS parameters, for each binder, were determined by performing linear regression on Equation 4 as shown in Figure 1. Steps explained in the previous subsection were then followed to compute E* based on the 1-37A and 1-40D models. One may surmise from Figure 1 that the ASTM A-VTS linear relationship is not as accurate for the PG 64-28 and PG 70-28. Furthermore, the A and VTS parameters based on level 1 binder data were found to be significantly different from the typical MEPDG A and VTS for the investigated binders. This may be due to the fact that MEPDG default A and VTS values are based on conventional viscosity testing data rather than binder G* and δ from DSR testing.

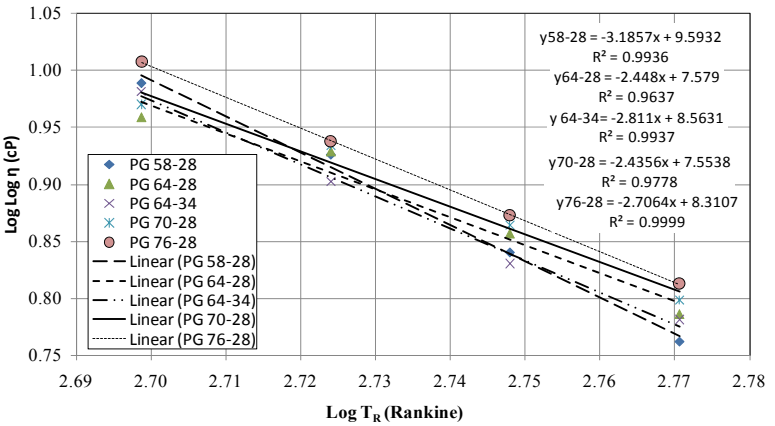
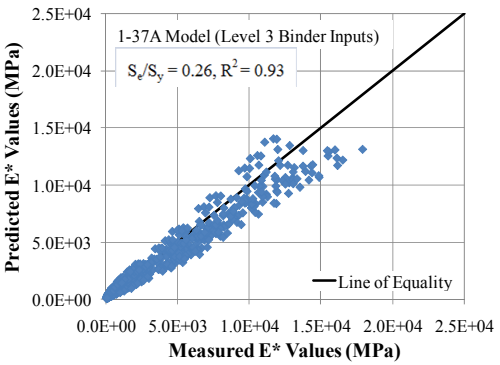


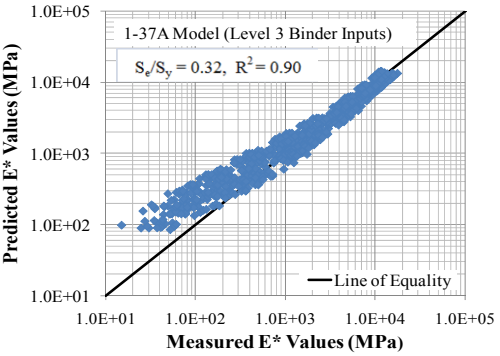

**Figure 1 A-VTS determination based on level 1 binder data**

### Accuracy and bias in MEPDG E* predictions

Laboratory measured E* values were compared to the predicted ones using the two predictive models incorporated in MEPDG. Figures 2 and 3 show a comparison between measured and predicted E* (based on MEDPG level 3 binder A-VTS) using the 1-37A and 1-40D models respectively. These figures are presented in both arithmetic and logarithmic scales. On the other hand, laboratory measured versus predicted E* values based on level 1 binder characterization inputs using the 1-37A and 1-40D models are shown in Figures 4 and 5 respectively.

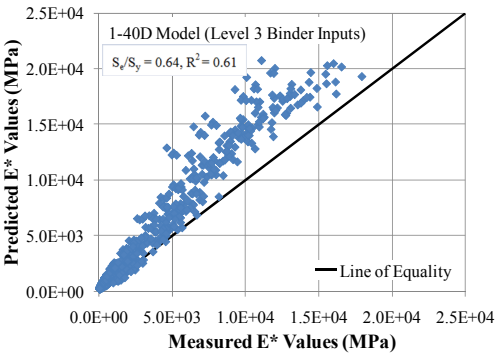



a) Arithmetic Scale

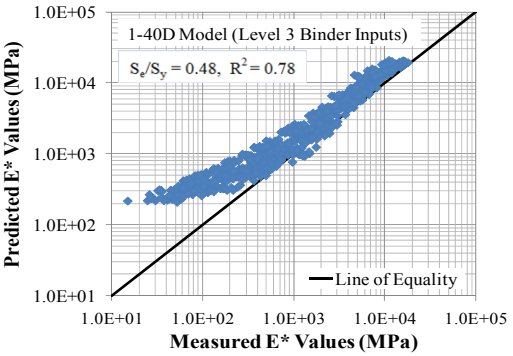



b) Logarithmic Scale

**Figure 2 Predicted versus Measured E* based on the 1-37A model (level 3 binder data)**

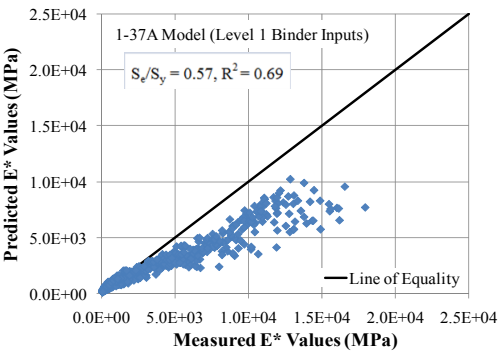



a) Arithmetic Scale

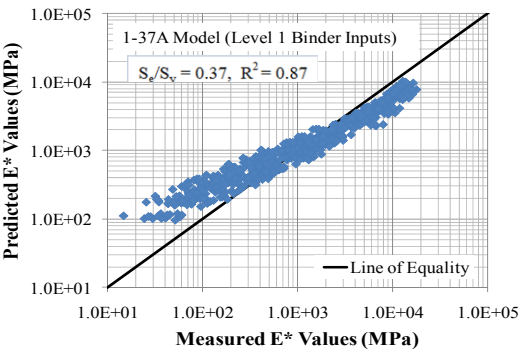



b) Logarithmic Scale

**Figure 3 Predicted versus measured E* based on the 1-40D model (level 3 binder data)**

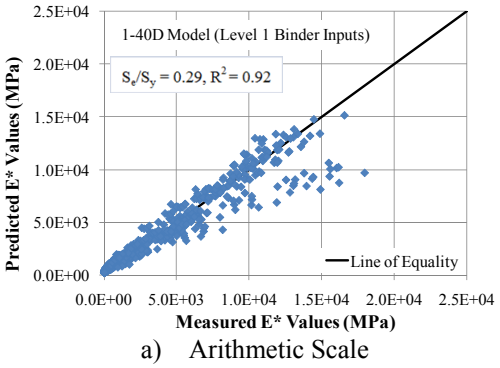



a) Arithmetic Scale

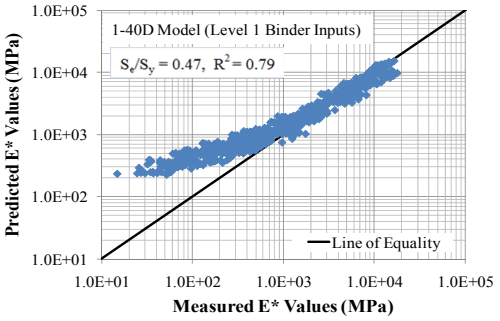



b) Logarithmic Scale

**Figure 4 Predicted versus Measured E* based on the 1-37A model (level 1 binder data)**




a) Arithmetic Scale




b) Logarithmic Scale

**Figure 5 Predicted versus measured E* based on the 1-40D model (level 1 binder data)**

The goodness-of-fit statistics shown in these figures were calculated with reference to the line of equality where $S_y$= standard deviation of the measured E* values about the mean measured; e=error between the predicted and measured E* values; $S_e$=standard error (i.e., standard deviation of error); $R^2$=coefficient of determination.

Both $S_e/S_y$ and $R^2$ are measures of the model accuracy (degree of scatter with reference to the line of equality) while Se and the average error ($e_{avg}$) are measures of the bias in predictions. The line of equality is a line with slope constrained to unity and intercept constrained to zero. Thus, the overall bias in prediction can also be assessed by measuring how close are the slope and intercept of the unconstrained linear regression between measured and predicted E* values to 1 and 0, respectively (Ceylan et al. 2009). Figure 6 compares the overall bias measures for each model. All bias parameters shown in this figure are calculated based on the arithmetic values of E* rather than the logarithmic values. The model with the lowest overall bias is the one with all parameters shown in Figure 6 are close to zero.

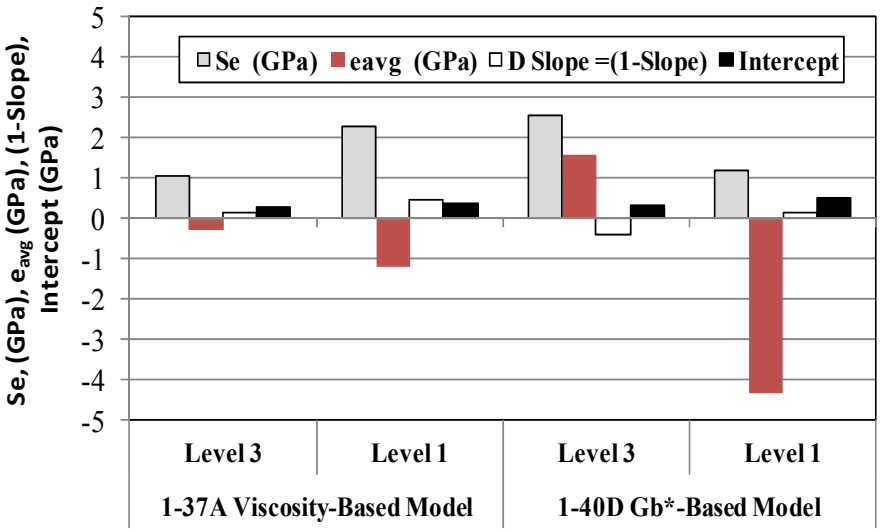



**Figure 6 Bias in MEPDG E* prediction models for Idaho mixtures**

Figures 2 through 6 along with the goodness-of-fit statistics shown in the graphs reveal that the 1-37A model with level 3 binder characterization show the most accurate ($R^2$=0.93, $S_e/S_y$=0.26 in arithmetic scale) and least biased E* estimates. However, this model shows some bias and high scatter in the predictions at the high temperature values. The observed bias and scatter at the high temperature regime agree with several literature studies (Ceylan et al. 2009; Harran and Shalaby 2009; Schwartz 2005). On the other hand, the 1-40D model with levels 1 binder characterization resulted in E* estimates that are in excellent agreement with the measured data ($R^2$=0.92, $S_e/S_y$=0.29) while the 1-40D model with level 3 binder characterization showed lower accuracy ($R^2$=0.61, $S_e/S_y$=0.64). Despite the high accuracy of the 1-40D model with level 3 binder data, it showed vey high bias at the higher temperatures compared to the 1-37A model with level 3 binder data. It significantly over predicts E* at the high temperatures.

With level 1 binder characterization, both E* models showed significant overall bias as shown in Figures 4, 5, and 6. The main reason for the high bias in the predicted E* can be attributed to the estimated A and VTS values from $G_b$* and δ calculated for the ASTM viscosity temperature relationship. The 1-37A model along with level 1 binder data was found to under predict E* values at low temperature and over estimates E* at higher temperatures. The E* estimates based on this model along with level 1 data showed the highest overall bias among all predictions as shown in Figure 6. It is suspected that the high bias in the estimated E* using both 1-37A and 1-40D models along with level 1 binder data and the low accuracy of the 1-37A model with level 1 binder data attributed to the fact that both models were calibrated based on the typical A and VTS values from viscosity testing of conventional binders rather than A and VTS estimated from superpave DSR testing.

## CONCLUSIONS

A comparison between the MEPDG 1-37A η-based and $G_b$*-based dynamic modulus prediction models along with level 1 and level 3 binder characterization inputs have been completed for typical ITD HMA mixtures. Based on the analysis of E* data from 15 ITD Superpave mixtures with five different binder grades, mix aggregate gradation, and volumetric properties, the following conclusions were found:

1. The 1-37A η-based E* predictive model along with level 3 binder characterization produces reasonably accurate estimates of the dynamic modulus of ITD HMA mixtures although some bias in the E* estimates was found at the higher temperature values.
2. The 1-37A η-based E* predictive model along with level 1 binder characterization produced highly biased and less accurate estimates of the dynamic modulus of ITD mixes. This model was found to underestimate E* at low temperatures and overestimate E* values at high temperatures.
3. The 1-40D $G_b$*-based E* predictive model along with level 1 binder characterization produces excellent E* estimates at the lower temperatures. However, at the higher temperatures, this model shows a highly significant biased E* estimates for ITD HMA mixtures compared to the 1-37A model with level 3 binder data.
4. The 1-40D model along with level 3 binder data was found to overestimates E* especially at the high temperatures.

The results of this study suggest that the 1-37A η-based E* model along level 3 binder characterization input data can be used to characterize ITD HMA mixes for MEDPG. It is not recommended to use MEPDG level 1 binder characterization for ITD mixes with level 3 E* prediction. Finally, MEPDG default A and VTS values for superpave PG grades should be revised based on binder G* and δ from DSR testing. If this is done, MEPDG E* predictive models should be recalibrated.

## ACKNOWLEDGEMENT

This study is partially funded by the National Science Foundation, Award No. 0612630. AC mixes were provided by the Idaho Transportation Department (ITD). Authors would like to acknowledge the technical and financial support of ITD and NSF.

## REFERENCES

Andrei, D., Witczak, M., & Mirza, W. (1999). "Development of a Revised Predictive Model for the Dynamic (Complex) Modulus of Asphalt Mixtures." Appendix CC-4, NCHRP 1-37A, National Research Council, Washington, DC.

ARA, Inc., ERES Consultants Division. (2004). "Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures." NCHRP 1-37A Final Report, Transportation Research Board, National Research Council, Washington, DC.

Bari, J., and Witczak, M. W. (2006) "Development of a New Revised Version of the Witczak E* Predictive Models for Hot Mix Asphalt Mixtures." Journal of the Association of Asphalt Paving Technologies," Vol. 75, pp. 381-417.

Bari, J., and Witczak, M. W. (2007). "New Predictive Models for the Viscosity and Complex Shear Modulus of Asphalt Binders for Use with the Mechanistic-Empirical Pavement Design Guide." In Transportation Research Record, Journal of the Transportation Research Board, No. 2001, Washington, DC, pp. 9-19.

Birgisson, B., Sholar, G., and Roque, R. (2005). "Evaluation of Predicted Dynamic Modulus for Florida Mixtures." In Transportation Research Record, Journal of the Transportation Research Board, No. 1929, Washington, DC, pp. 200-207.

Ceylan, H., Schwartz, C. W., Kim, S., & Gopalakrishnan, K. (2009). "Accuracy of predictive Models for Dynamic Modulus of Hot-Mix Asphalt." Journal of Materials in Civil Engineering, pp. 286-293.

Dongre′, R., Myres, L., D'Angelo J., Paugh, C., and Gudimettla, J. (2005) "Field Evaluation of Witczak and Hirsh Models for Predicting Dynamic modulus of Hot-Mix Asphalt." Journal of the Association of Asphalt Paving Technologies," Vol. 74, pp. 381-434.

Far, M., Underwood, B., Ranjithan, S., Kim, R., and Jackson, N. (2009). "Application of Artificial Neural Networks for Estimating Dynamic Modulus of Asphalt Concrete." In Transportation Research Record, Journal of the Transportation Research Board, No. 2127, Washington, DC, pp 173-183.

Harran, G., and Shalaby, A. (2009). "Improving the Prediction of the Dynamic Modulus of Fine-Graded Asphalt Concrete Mixtures at High Temperatures." Canadian Journal of Civil Engineering , 36 (2), pp. 180-190.

Schwartz, C. W. (2005). "Evaluation of the Witczak Dynamic Modulus Prediction Model." Proc., 84th Annual Meeting of The Transportation Research Board (CD-ROM), Transportation Research Board, Washington, D.C.

Witczak, M., El-Basyouny, M., & El-Badawy, S. (2007). "Incorporation of the New (2005) E* Predictive Model in the MEPDG." NCHRP 1-40D Final Report, Arizona State University, Tempe, AZ.

# Comparison of Idaho Pavement Design Procedure with AASHTO 1993 and MEPDG Methods

S. M. El-Badawy[1], PhD; F. M. Bayomy[2], PhD, PE; M. Santi[3], PE; C.W. Clawson[4], PE.

[1]Research Fellow, National Institute for Advanced Transportation Technology (NIATT), University of Idaho, P.O. Box 441022, Moscow, Idaho 83844-1022; PH (208) 885-6818; FAX (208) 885-6608; email: elbadawy@uidaho.edu; Permanent address: Public Works Engineering Dept., Faculty of Engineering, Mansoura University, Mansoura, Egypt.

[2]Professor, Department of Civil Engineering, University of Idaho, P.O. Box 441022, Moscow, Idaho 83844-1022; PH (208) 885-6784; FAX (208) 885-6608; email: bayomy@uidaho.edu

[3]Assistant Materials Engineer, Idaho Transportation Department, P.O. Box 7129, Boise, Idaho 83707-1129, PH (208) 334-8450; FAX (208) 334-4411; email: mike.santi@itd.idaho.gov

[4]District 2 Materials Engineer, District 2, Idaho Transportation Department, P.O. Box 837, Lewiston, Idaho 83501, PH (208) 799-5090; FAX (208) 799-4301; email: chad.clawson@itd.idaho.gov

## ABSTRACT

Several in service pavements located in different regions of Idaho that have been designed according to the ITD design method were redesigned using the AASHTO 1993 as well as the Mechanistic-Empirical Pavement Design Guide (MEPDG) procedures. All designs were conducted at a 50% reliability level. The nationally calibrated MEPDG (version 1.1) was used to predict the performance of the three design methods. Level 2 subgrade material characterization inputs were used in the MEPDG analysis. All other MEPDG inputs were level 3. Performance indicators predicted using MEPDG related to the three design methods were compared to each other. Results showed that, relative to AASHTO 1993 and MEPDG procedures, ITD design method significantly overestimates the thickness of the pavement structure, and particularly the thickness(s) of the unbound layer(s). On the other hand, the AASHTO 1993 and MEPDG guides show reasonable agreement on the resulting pavement structure.

## INTRODUCTION

The current flexible pavement design method adopted by the Idaho Transportation Department (ITD) is an empirical procedure based on the R-value of the subgrade (ITD 2009). It incorporates traffic in terms of Traffic Index which is a function of the design 18-Kip Equivalent-Single Axle Load (ESAL) (Juolo and Kilchoer 1980; PMS 1986). It also incorporates a Climatic Factor (CF) to reflect the

various geographical regions in the State of Idaho. The ITD procedure determines the pavement thickness as a Gravel Equivalence (GE) based on an empirical equation adopted from the California Department of Transportation (Caltrans). The GE is then transferred to various layer thicknesses through a gravel factor ($G_f$) for each type of material. Observations in Idaho showed that many existing roadways that have been designed by the ITD design method have performed beyond their selected design lives, and still perform adequately. There is increased emphasis to optimize the thickness design for the selected design period to achieve the most cost effective design section.

The current design practice of flexible pavements in the U.S. includes various methods. Results of a survey in 2007 (Gary 2009) showed that, based on 50 State Highway Agencies (SHAs) responders, 63% of the state DOTs are using the AASHTO 1993 Pavement Design Guide, 12% are still using the 1972 AASHTO Interim Design Guide, 13% are using individual state design procedure, 8% are using a combination of AASHTO and state procedure, while the rest are using other design procedures. In addition, the same survey results showed that 80% of the SHAs including ITD have implementation plans for the MEPDG (Gary 2009).

Although the various versions of the AASHTO design guides are based on empirical models drawn from field performance data measured at the AASHO road test in the late 1950's, yet they have served well for several decades (Carvalho and Schwartz 2006). However, the AASHTO 1993 method has many limitations and deficiencies which motivated the development of the MEPDG (ARA. 2004).

The MEPDG software is an advanced tool for the analysis/design of new and rehabilitated flexible and rigid pavement structures based on mechanistic-empirical principles. The software mechanistically calculates the structural responses within the flexible pavement system, using the JULEA multi-layer elastic analysis and the finite element analysis (ARA 2004). Moisture and temperature variations within the pavement structure are also calculated internally using the Enhanced Integrated Climatic Model (EICM). The EICM, in MEPDG, utilizes a comprehensive database from 851 weather stations throughout the United States. Pavement distresses and roughness are predicted, from the mechanistically calculated strains and deformations, using statistical transfer functions. In the current software version of the MEPDG (version 1.1), these transfer functions are nationally calibrated based on field data from 94 LTPP sections distributed all over the US. The software also allows users to input user defined calibration coefficients to reflect local conditions.

## OBJECTIVES

The main objective of this study is to compare design results by the current ITD empirical design method to the most commonly used design methods in the US, especially those used in the Northwest region; the AASHTO 1993 and the MEDPG.

## PROJECTS SELECTED

Different in-service roadway sections designed according to ITD method were identified in different regions in the State of Idaho. The projects selected are sections

from SH-62, SH-3, US-95, and US-2. The US-2 project has three different sections with the same design inputs expect for the subgrade strength. The US-2 sections allow for direct evaluation of the influence of the subgrade strength on the designed pavement structure using the three pavement design methods considered in this study. Table 1 shows performance indicators of these projects as of 2008. The data shows good performance for the US-95 pavement section even though it is near the end of its service life.

**Table 1 Performance Indicators of the Selected Pavement Sections**

| Design Project | SH-62 | US-95 | SH-3 | US-2(a) | US-2(b) | US-2(c) |
|---|---|---|---|---|---|---|
| Years in Service | New | 18 | 10 | 6 | 6 | 6 |
| RI[1] | - | 3.6 | 2.4 | 3.4 | 4.7 | 4.1 |
| CI[2] | - | 4.5 | 2.5 | 4.7 | 5 | 5 |
| SN[3] | | 47 | 52 | 44 | 44 | 44 |
| Rut Depth, in | - | 0.19 | 0.22 | 0.16 | 0.23 | 0.24 |

[1]RI (Roughness Index): ranges from 0.0 to 5.0 with 0.0 being extremely rough and 5.0 being perfectly smooth.
[2]CI (Crack Index): ranges from 0.0 to 5.0 with 0.0 representing extreme cracking, and 5.0 corresponding to little or no cracking.
[3]SN (Skid Number): ranges from 20 to 100, threshold value SN < 35.

## INPUT DATA FOR EACH DESIGN METHOD

Input data vary significantly among the three design methods. Data required was collected from several sources. Some of the data sources beside ITD design phase reports are Google Earth, ITD Pavement Management System (PMS), ITD Highway Information through ITD official website, and U.S. Geological Survey website. The design inputs for the three methods are listed in Tables 2, 3 and 4 for ITD, AASHTO 1993 and MEPDG respectively.

**Table 2 Design Input Data for ITD Design Method**

| Design Input | Design Project | | | | | |
|---|---|---|---|---|---|---|
| | SH-62 | US-95 | SH-3 | US-2(a) | US-2(b) | US-2(c) |
| Design Life, years | 20 | 20 | 20 | 20 | 20 | 20 |
| Traffic Index | 8.78 | 9.9 | 10.5 | 11.51 | 11.51 | 11.51 |
| Climatic Factor | 1.1 | 1.0 | 1.0 | 1.1 | 1.1 | 1.1 |
| Base R-Value | 75 | 80 | 80 | 80 | 80 | 80 |
| Subbase R-Value | - | 65 | - | 65 | 65 | 65 |
| Subgrade R-Value* | 8 | 5 | 25 | 50 | 20 | 15 |

* Laboratory measured values

**Table 3 Design Inputs for the AASHTO 1993 Design Guide**

| Design Project | SH-62 | US-95 | SH-3 | US-2(a) | US-2(b) | US-2(c) |
|---|---|---|---|---|---|---|
| Design Reliability, % | 50 | 50 | 50 | 50 | 50 | 50 |
| Design ESLAs | 81,6000 | 2,240,000 | 3,696,000 | 7,920,000 | 7,920,000 | 7,920,000 |
| ΔPSI (Loss of Service-Survivability) | 2 | 2 | 2 | 2 | 2 | 2 |
| Structural layer Coefficient ($a_i$) | | | | | | |
| $a_1$ | 0.44 | 0.44 | 0.44 | 0.44 | 0.44 | 0.44 |
| $a_2$ | 0.14 | 0.17 | 0.17 | 0.14 | 0.14 | 0.14 |
| $a_3$ | - | 0.19 | - | - | - | - |
| Drainage Coefficients ($m_i$) | | | | | | |
| $m_2$ | 1.2 | 1 | 1.2 | 1.2 | 1.2 | 1.2 |
| $m_3$ | - | 1.2 | - | - | - | - |
| Mr (psi), Base* | 30,000 | 40,000 | 40,000 | 40,000 | 40,000 | 40,000 |
| Mr (psi), Subbase* | | 35,000 | | 30,000 | 30,000 | 30,000 |
| Mr (psi), Subgrade* | 5,595 | 3,930 | 15,030 | 28,905 | 12,255 | 9,480 |

* Estimated Values

Reasonable assumptions were made for some of the input parameters where information was not available. Most of the inputs used in MEPDG for the investigated projects are considered level 3 inputs. For the subgrade material characterization, MEPDG level 2 was used since lab data for the R-value of the subgrade was available. The appropriate weather station data for each location was chosen based on the Latitude, Longitude and Elevation of the project.

For the traffic data inputs, MEPDG uses the axle load spectra while ITD uses the Traffic Index (function of the classical ESALs) and AASHTO 1993 design method uses the classical ESALs as the only traffic input. In order to maintain consistent traffic inputs for the three methods, the traffic load spectra analysis in the MEPDG was modified to reflect the design ESALs. The use of ESALs in the MEPDG was conducted by: 1) using 100% FHWA Truck Class 5 (the only truck class with 2-axles) and 0% for all other truck classes, 2) using 100% of 18-kip single axle load and 0% for all other single axle loads, 3) using 100% of trucks on design direction, and finally, 4) using 100% trucks in the design lane. In addition, the design ESALs expected to use the pavement was assumed to be uniformly distributed over the deign life.

Furthermore, the MEPDG Enhanced Integrated Climatic Model (EICM) was deactivated in all runs. Instead, the representative modulus module was used in all MEPDG runs.

The above procedures were employed in order to simulate the ITD and AASHTO 1993 design methods which use one representative value for strength characterization for each of the unbound granular and subgrade layers.

**Table 4 Design Inputs for the MEPDG Design Guide**

| State Highway | SH-62 | US-95 | SH-3 | US-2 | US-2 | US-2 |
|---|---|---|---|---|---|---|
| Location: | | | | | | |
| County | Lewis | Washington | Nez Perce | Bonner | Bonner | Bonner |
| Latitude, Deg.Min. | 46.14 | 44.19 | 46.28 | 48.15 | 48.15 | 48.15 |
| Longitude, Deg.Min. | -116.14 | -116.55 | -116.45 | -116.39 | -116.39 | -116.39 |
| Elevation, ft | 3215 | 2490 | 850 | 2085 | 2085 | 2085 |
| Main Traffic Inputs: | | | | | | |
| Design Life, years | 20 | 20 | 20 | 20 | 20 | 20 |
| Speed, mph | 25 | 60 | 55 | 60 | 60 | 60 |
| AADTT (design lane) | 56 | 153 | 253 | 542 | 542 | 542 |
| HMA Properties: | | | | | | |
| Binder Type | PG58-28 | AC-10 | PG70-28 | PG58-28 | PG58-28 | PG58-28 |
| Cumulative, % Retained 3/4" Sieve | 0 | 0 | 3 | 1 | 1 | 1 |
| Cumulative, % Retained 3/8" Sieve | 15 | 26 | 23 | 27 | 27 | 27 |
| Cumulative, % Retained #4 Sieve | 45 | 49 | 54 | 53 | 53 | 53 |
| % Passing #200 Sieve | 8.2 | 4.9 | 4 | 6 | 6 | 6 |
| % $V_{beff}$* | 11.01 | 9.38 | 11.6 | 9.95 | 9.95 | 9.95 |
| % Air Voids | 7 | 7 | 7 | 7 | 7 | 7 |
| Unbound Granular Base Course Properties: | | | | | | |
| Material Type | A-1-a | A-1-a | A-1-a | A-1-a | A-1-a | A-1-a |
| Mr, psi | 30,000 | 40,000 | 40,000 | 40,000 | 40,000 | 40,000 |
| Unbound Granular Subbase Course Properties: | | | | | | |
| Material Type | - | Permeable Aggregate | - | Permeable Aggregate | Permeable Aggregate | Permeable Aggregate |
| Mr, psi | - | 35,000 | - | 30,000 | 30,000 | 30,000 |
| Subgrade Properties: | | | | | | |
| Material Type | CL | MH | SM | SP-SM | ML | CL |
| Mr, psi | 5595 | 3930 | 15030 | 28905 | 12255 | 9480 |
| GWT Depth, ft | 6 | 7 | 45 | 10 | 10 | 10 |

*$V_{beff}$ = effective binder content by volume

## PAVEMENT STRUCTURE DESIGN

Three designs were considered for the selected pavement sections. The ITD design was used from ITD design phase reports and it was also checked. The pavements were also designed using both the AASHTO 1993 guide and the MEPDG method using the inputs presented previously. In order to find the design structure for each project, the MEPDG software was run several times based on the criteria for

primary roadways presented in Table 5. Finally, MEPDG was also used to evaluate the performance of the pavement structures deigned by the ITD and AASHTO 1993 methods.

**Table 5 MEPDG Recommended Design Criteria for Primary Roadways (AASHTO, 2008)**

| Distress | Threshold Value at 50% Reliability |
|---|---|
| Terminal IRI (in/mile) | 200 |
| AC Alligator Cracking (% lane area) | 20 |
| Transverse Cracking (ft/mile) | 700 |
| Total Rutting (in) | 0.50 |

## RESULTS AND ANALYSIS

A comparison of the computed pavement structures using the three design methods is shown in Table 6. Based on these results, it can be concluded that, ITD design method always yields a pavement structure that is significantly thicker compared to the other two methods. A reasonable agreement was found between the three design methods regarding the design thickness of the asphalt layer. However, the unbound granular layer(s) thickness(s) was found to be much thicker in the ITD design method. On the other hand, for all practical purposes, the AASHTO 1993 and MEPDG methods yielded reasonably similar pavement structures.

Recall that the US-2 project has three sections with the same inputs except the subgrade strength. The designed pavement structures using the three different methods for the three US-2 sections showed that the subgrade strength is overemphasized by the ITD design method compared to the other two design methods. This is shown in Table 6 by the dramatic increase in the subbase thickness with the decrease in the subgrade strength based on ITD design.

The MEPDG predicted distresses using pavement structures resulted from the three design methods for the pavement sections are shown in Figures 1 through 3. No thermal cracking was predicted for any of the pavement structures resulted by any of the investigated design methodologies. Based on the results shown in Figure 1, one can surmise that, there is no significant difference in the predicted AC rutting for all pavement structures resulted from the three design methods. The primary reason for this observation is that, generally, there was no significant difference in the design AC thickness resulted from the different design methods. In addition, the AC rutting is only a function of the AC layer properties, traffic, and environment and it is not affected by the foundation properties. This observation agrees with other studies (Witczak et al. 2004; El-Basyouny et al. 2005). Figure 1 also shows that, the predicted subgrade rutting for the pavement structures designed by the ITD method is the lowest. The main reason for this observation is that ITD design always yields thicker pavement structure. It is obvious that if the pavement structure above the subgrade is thicker, the compressive strain at the top as well as any point within the subgrade decreases, hence subgrade rutting decreases. It can be also concluded that as

the thickness of the unbound base/subbase layers increases, the rutting within this layer(s) also increases. This is obvious in the structures designed with the ITD method (see Table 6 and Figure 1). Furthermore, for all practical purposes, Figure 1 also shows no significant difference in the predicted total rutting for the structures designed with AASHTO 1993 and MEPDG as these two methods generally yielded similar pavement structures. For only the US-95 section, the ITD and AASHTO 1993 designs yielded a structure with a total rutting slightly in excess of the MEPDG recommended threshold value. All other structures conform to the MEPDG recommended design criteria.

**Table 6 Comparison of the Recommended Pavement Structure by the Investigated Design Methods**

| Design Project | Layer Thickness, in | ITD | AASHTO 1993 | MEPDG |
|---|---|---|---|---|
| SH-62 | AC | 4.2 | 4.0 | 4.0 |
| | Granular Base | 22.2 | 8.0 | 6.0 |
| SH-3 | AC | 5.3 | 4.5 | 4.5 |
| | Granular Base | 23 | 6.0* | 6.0 |
| US-95 | AC | 3.6 | 4.0 | 5.0 |
| | Granular Base | 6.0 | 6.0 | 6.0 |
| | Granular Subbase | 30 | 6.0 | 8.0 |
| US-2(a) | AC | 6.0 | 5.5 | 4.5 |
| | Granular Base | 6.0 | 6.0* | 8.0 |
| | Granular Subbase | 9.0 | 0.0 | 0.0 |
| US-2(b) | AC | 6.0 | 5.5 | 5.5 |
| | Granular Base | 6.0 | 6.0* | 6.0 |
| | Granular Subbase | 26.4 | 0.0 | 0.0 |
| US-2(c) | AC | 6.0 | 5.5 | 5.5 |
| | Granular Base | 6.0 | 7.0 | 6.0 |
| | Granular Subbase | 30 | 0.0 | 0.0 |

* Minimum Recommended Thickness by the AASHTO 1993 Design Method

A comparison of Figures 2, 3, and 4 along with the MEPDG recommended design criteria reveals that MEPDG deigns were governed by the total pavement rutting not by cracking or International Roughness Index (IRI). Figure 1 shows that among the three design methods, ITD design method always yielded pavement sections with the least amount of fatigue cracking. This is obviously because ITD method always yields thicker (2.5 to 6 times as thick) unbound granular layers which results in a stronger support for the AC layer, hence less tensile strain at the bottom of the AC layer occurs.

In comparing the actual distress to that predicted by MEPDG, a comparison was not possible since ITD distress measurement system is inconsistent with the MEPDG requirement. For example, the current ITD method of pavement evaluation combines thermal cracking, bottom-up alligator fatigue cracking, and top-down longitudinal cracking into one cracking index (CI). On the other hand, the MEPDG calibration requires the value of each one of these distresses separately.

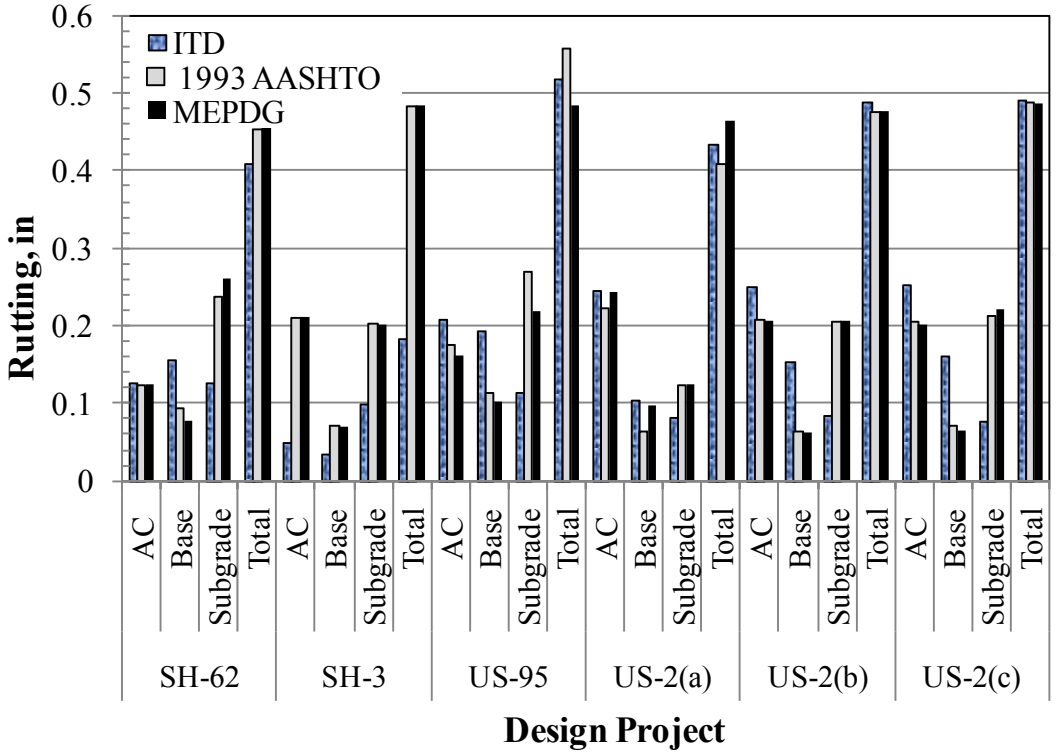



**Figure 1 Comparison of MEPDG predicted rutting from the three design methods**

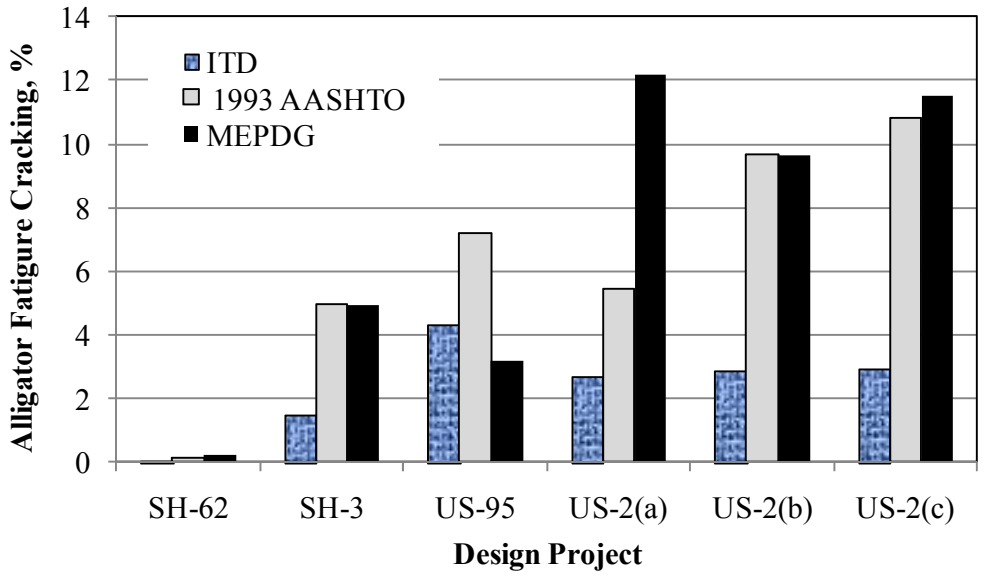



**Figure 2 Comparison of MEPDG predicted fatigue cracking from the three design methods**

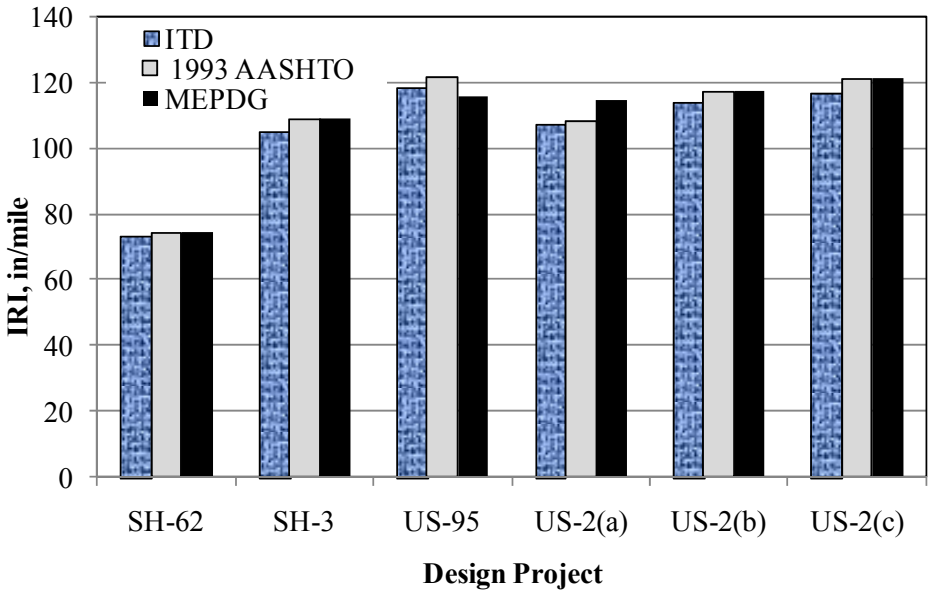



**Figure 3 Comparison of MEPDG predicted IRI from the three design methods**

## CONCLUSIONS

A comparison between ITD, AASHTO 1993, and MEPDG design methods have been completed based on six different existing roads located in different regions in the State of Idaho. The nationally calibrated MEPDG software (version 1.1) was used to predict performance of the resulted pavement structure from each design method. The three design methods were compared and the following conclusions are drawn:

1. In general, using the design criteria recommended by MEPDG, both ITD and AASHTO 1993 methods yielded pavement structures that conform to the criteria. However, ITD design method was found to yield highly conservative pavement structures. The unbound granular layer(s) thickness(s) resulted from the ITD design methods were much thicker compared to the 1993 and MEPDG designs.
2. The three design methods yielded reasonably similar thickness for the asphalt layers.
3. The influence of the subgrade strength on the designed pavement structure is overemphasized by the ITD design method compared to both the AASHTO 1993 and MEPDG methods.
4. A reasonable agreement was found between the AASHTO 1993 and MEPDG design methods regarding the resulting pavement structure.
5. No thermal cracking was predicted for any of the studied pavement sections.
6. The resulting structural design for each of the investigated pavement sections, using MEPDG method, was found to be governed by the predicted total pavement rutting. Predicted alligator fatigue cracking, thermal cracking and IRI were much less than the MEPDG recommended threshold values for the investigated sections.

Using MEPDG, it is anticipated that considerable savings will be obtained for designing pavement structures for ITD. It is recommended that ITD proceed with the MEPDG implementation and calibration. This will require ITD to change their current distress measurement system and update their PMS system as well to comply with the MEPDG requirements.

## ACKNOWLEDGEMENT

This study is partially funded by the Idaho Transportation Department (ITD). Authors would like to acknowledge the technical and financial support of ITD.

## REFERENCES

ITD. (2009). "Materials Manual." Idaho Transportation Department, State of Idaho.

Juola, R. C., and kilchoer, R. (1984). "Load Distribution Factors for Idaho Highways – The Role of Load Analysis in Pavement Design." Internal Report, Idaho Department of Highways.

PMS. (1986). "The Feasibility of Updating the Idaho Department of Transportation's Pavement Design procedure." Pavement Management Systems Limited, Cambridge, Ontario.

Crawford, G. (2009). "National Update of MEPDG Activities." 88th Annual TRB Meeting, Power Point Presentation, Washington D.C.

Carvalho, R., and Schwartz, C. (2006). "Comparison of Flexible Pavement Designs: AASHTO Empirical Versus NCHRP Project 1-37A Mechanistic-Empirical." In Transportation Research Record, Journal of the Transportation Research Board, No. 1947, Washington, DC, pp. 167-174.

AASHTO. (1993). "Guide for Design of Pavement Structures." American Association of State Highway and Transportation Officials, Washington, D.C.

ARA, Inc., ERES Consultants Division. (2004). "Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures." NCHRP 1-37 A Final Report, Transportation Research Board, National Research Council, Washington, DC.

AASHTO. (2008). "Mechanistic-Empirical Pavement Design Guide: A Manual of Practice." Interim Edition, American Association of Highways and Transportation Officials.

El-Basyouny M., Witczak, M. W., and El-Badawy, S. (2005). "Verification of the Calibrated Permanent Deformation Models for the 2002 Design Guide." Journal of the Association of Asphalt Paving Technologists, Vol 74, pp. 601-652.

Witczak, M., El-Basyouny M., and El-Badawy S. (2004) "Appendix GG-2: Sensitivity Analysis for Permanent Deformation for Flexible Pavements." Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures, NCHRP 1-37A.

# Integrating Traditional Characterization Techniques in Mechanistic Pavement Design Approaches

A.A. Araya[1], M. Huurman[2] and A.A.A. Molenaar[3]

[1]Road and Railway Engineering, Delft University of Technology, P.O. Box 5048, 2600 GA, Delft, Tel: +31 15 278 4008, Fax: +31 15 278 3443, email: a.a.araya@tudelft.nl
[2]Road and Railway Engineering, Delft University of Technology, P.O. Box 5048, 2600 GA, Delft, Tel: +31 15 278 1525, Fax: +31 15 278 3443, email: m.huurman@tudelft.nl
[3]Road and Railway Engineering, Delft University of Technology, P.O. Box 5048, 2600 GA, Delft, Tel: +31 15 278 4812, Fax: +31 15 278 3443, email: a.a.a.molenaar@tudelft.nl

## ABSTRACT

Although widely applicable and useful, the traditional CBR test does not provide the mechanical behaviors such as resilient and permanent deformation characteristics of granular road materials. A relatively simple testing technique is developed to characterize the resilient modulus of granular materials based on the traditional CBR test using repeated load cycles. The Finite Element Method (FEM) analysis has been attempted for the purpose of modeling the Repeated Load CBR and derives an equivalent resilient modulus of the sample as a bulk. Strain gauges were used to measure the lateral deformation of the CBR mould from which the confining stress can be estimated to determine a stress dependent resilient modulus. Furthermore, a large scale cyclic load triaxial test was carried out to validate the result of the repeated load CBR on coarse granular materials from South Africa and Ethiopia. The repeated load CBR test is quite useful to estimate the resilient modulus of unbound granular materials that can be used as an input in mechanistic pavement design analysis in the absence of triaxial testing facilities.

## INTRODUCTION

An examination of the history of pavement design reveals an evolutionary process that began with rule-of-thumb procedures and gradually evolved into empirical design equations based on experience and road test pavement performance studies. As Elliott and Thompson (1985), Monismith (2004), de Beer (1990) stated this evolution and transformation has been accompanied by the development of an understanding of material behavior, load-pavement distress relationships and environment interactions. Through the years, much of the development has been hampered by the complexity of the pavement structural system both in terms of its indeterminate nature and in terms of the changing and variable conditions to which it is subjected.

The advent of the powerful digital computers and their penetration even to the remote places has created, these days, the possibility of the practical use of analytical solutions to determine stresses and strains in pavements. Today, much effort is spent on further developing of mechanistic design procedures, both

improving the existing analytical tools for determination of pavement responses and by performing extensive long-term pavement performance studies. Many countries, particularly developing countries, however still rely on empirical design methods, realizing that more sophisticated mechanistic design procedures often require too many assumptions regarding material behavior and too complicated material testing techniques to be of direct practical use (Molenaar 2004).

In flexible pavements, especially when the surface is a thin asphalt layer or chip seal, the role of granular layers is very important in the general performance of the structure since the load is distributed to the subgrade mainly through these layers. Over the last four decades, many researchers have been investigating the resilient behavior of granular materials as the shift from the empirical to the mechanistic design of pavement gained popularity. Many factors were determined to have an effect on the resilient modulus of granular materials. The state of stress was found to have the most influence on the resilient behavior. Hicks (1970); Smith and Nair (1973); Sweere (1990) and Huurman (1997) have shown that the resilient modulus increases with an increase in confining pressure. On the other hand, Brown (1974) reported a significant effect of the deviator stress, especially at high stress levels. Many other factors such as density, moisture content, stress history, number of load cycles, particle shape and mineralogy, and load characteristics (frequency, and load sequence) have been found to impact the resilient behavior of granular materials. A detailed discussion of these parameters can be found in LeKarp et al. (2000).

Most of these investigations have been carried out using facilities such as the Repeated Load Triaxial (RLT) test and Hollow Cylinder Apparatus (HCA) etc. However such specialist tests are considered too complex and unaffordable to implement in routine road construction projects particularly in developing countries. Even in the industrialized countries such tests are implemented mainly for research purpose but hardly used in day to day engineering practice. The existence of the gap between research and industry based practice reveals the absence of appropriate characterization techniques. These characterization techniques developed for research purposes have economical and practical limitations that prevent their widespread use. Edward (2007) reports these barriers include level of complexity, skills or trainings required prior to use, availability and affordability. On the other hand despite their worldwide acceptance and existence for long time index testing such as California Bearing Ratio (CBR), being too empirical, have technical limitations to be used in the Mechanistic design methods.

A new characterization technique, a repeated load CBR (RL-CBR) test, is developed in this research to characterize mechanical properties such as resilient modulus of unbound granular road materials. This intermediate testing mechanism, which is based on the traditional CBR testing equipment, can bring mechanistic design method into practice. The advantage of integrating the traditional CBR characterization technique in to mechanistic pavement deign approach through the development of the repeated load CBR testing is that the RL-CBR can be carried out in most standard non-sophisticated road engineering laboratories. This paper presents the characterization techniques of the RL-CBR and the results for the South African ferricrete granular subbase material compared to the well accepted triaxial test results.

## THE REPEATED LOAD CBR

The principle of the RL-CBR test is similar to the standard CBR test but repeated loads are applied. Upon multiple repetitions of the same magnitude of loading granular materials comes to a state in which almost all strain under a load application is recoverable. The permanent (plastic) strain ceases out to exist or becomes negligible and the material behaves basically elastic i.e. with stable recoverable deformation (Araya et al. 2010). From the applied stress and the measured strain an elastic modulus can be estimated. By recording the load and displacement and plot in x-y axes, similar to Figure 1 are obtained from which load levels and total, resilient (elastic) and permanent (plastic) deformations under the penetration plunger can be determined.

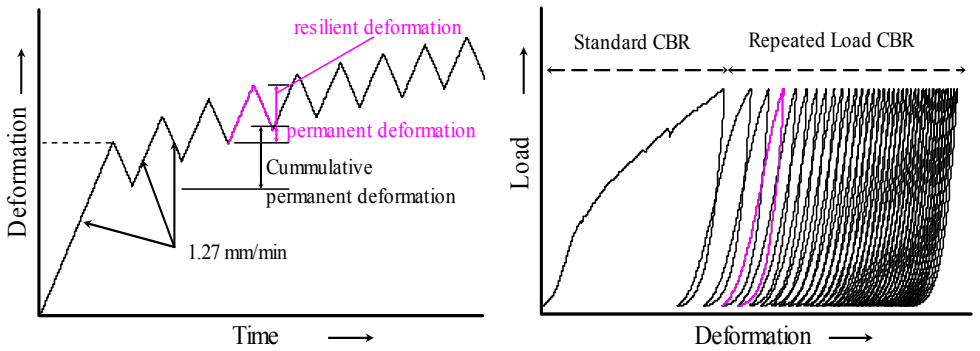



**Figure 1. Repeated load CBR test principle and load-deformation pattern**

**Test setup.** The intention of the repeated load CBR tests setup is to estimate the resilient modulus from a standard CBR testing facility by repeating the loading and unloading cycle. The tests in this research were performed using large mould having a diameter of 250 mm and a height 200 mm to accommodate the full 0/45 gradation. Proportionally a bigger penetration plunger of 81.5 mm diameter is used instead of the standard CBR 49.64 mm diameter plunger.

To simulate the repeated load application in the standard CBR the test is performed in the displacement controlled mode at a constant displacement rate of 1.27 mm/min (0.05 inch/min) = 0.021 mm/sec for both loading and unloading. The specimen is first loaded at the rate of 1.27 mm/min to a predetermined deformation (for e.g. 2.54 mm) or a target load level. The load is recorded and unloaded with the same rate (1.27 mm/min) to a minimum contact load of 0.5 to 1 KN (0.1 to 0.2 MPa) to keep the plunger in contact with the specimen. The loading and unloading cycles are generally repeated for about 50 – 100 load cycles at which the permanent deformation due to the last 5 loading cycles will be less than 2% of the total permanent deformation at that point.

The RL-CBR test is carried out with strain gauges. Strain gauges were used to measure the lateral deformation of the CBR mould to get the feeling of the confining stress developed by the steel mould. Four strain gauges capable of measuring in micro-strains are glued at the external surface of the mould which measures the mould lateral deformation during the loading and unloading cycles. The

four strain gauges are positioned in such a way that the variation in lateral deformation (if any) along the height and circumference of the mould can be observed. Two of the strain gauges are position at the mid-height of the mould in opposite diametric side of the mould and the other two near to the top edge of the mould (40 mm below the top edge). The schematic diagram of the RL-CBR mould and the strain gauges are shown in Figure 2.

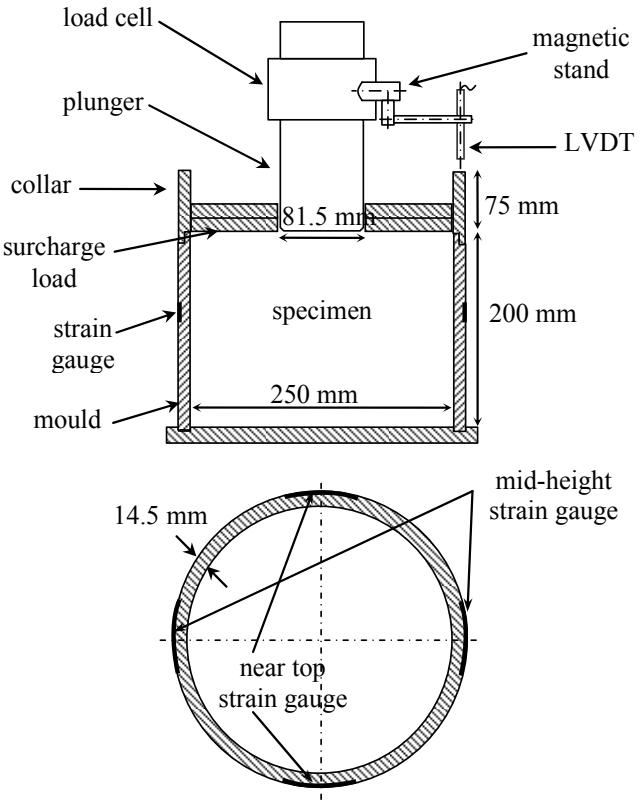



**Figure 2. Repeated load CBR test with strain gauges schematic diagram**

**Material.** The material tested was a South African subbase granular material ferricrete (FC). The ferricrete are natural gravel obtained from a borrow pit in South Africa. It is natural coarse aggregate relatively week to crushing where its particles are characterized with porous spherical shape and rough surface texture with grading 0/45, Table 1 shows the wet sieve gradation of the material. The iron-rich (sub) tropical ferricrete is characterized as a mineral conglomerate consisting of surficial sand and gravel cemented into a hard mass by iron oxide derived from the oxidation of percolating solutions of iron salts. The RL-CBR test is carried out at varying moisture content (MC) and degree of compaction (DOC). The molding moisture varies as dry (5%), moderate (7.5%), wet (9.5%) and the DOC varies from 95% - 100% of the maximum modified Proctor dry density (MPDD).

**Table 1 Gradation of the ferricrete**

| Sieve size | 45 | 31.5 | 22.4 | 16 | 8 | 4 | 2 | 0.5 | 0.18 | 0.063 |
|---|---|---|---|---|---|---|---|---|---|---|
| % passing | 100 | 99.1 | 88.8 | 75.2 | 50.0 | 33.3 | 27.3 | 23.0 | 16.9 | 10.7 |

**Finite Element Method.** A Finite Element (FE) analysis has been attempted to model the RL-CBR test using ABAQUS. As the loading and boundary conditions of the test are symmetric with respect to the central axis an axisymmetric approach is adopted in the modeling. A linear elastic material property is assumed for the steel mould and the granular material with 210 GPa elastic modulus, E, and 0.2 Poisson's ratio, ν, for the steel mould and varying E and ν for the granular material as shown in Table 2. The plunger is assumed as a rigid body and a contact surface property was defined between the plunger and the granular as hard contact; and the mould and the granular with exponential pressure-overclosure relationship defined in ABAQUS.

**Table 2 Granular material properties used in the FE analysis**

| | |
|---|---|
| Poisson's ratio ν [-] | 0.15, 0.25, 0.35, 0.45 |
| Elastic modulus E [MPa] | 100, 200, 300, 400, 500, 600, 800, 1000 |

For a given material property of the granular material a strain controlled is used to simulate the test i.e. a vertical displacement is applied on the rigid plunger. Stresses, strains and deformations through out the granular material and steel mould; and total load on the plunger has been recorded for each granular material property combination and applied vertical plunger displacement. This data set can be recorded for each node or element of the mesh shown in Figure 3.

By using non linear multidimensional least square regression fitting on the FE analysis data transfer functions have been developed that relate material properties and the bulk stresses components of the specimen with the laboratory measurable test parameters i.e. the plunger load, plunger deformation and the lateral or tangential mould exterior strain at mid height. The vertical and radial stresses of the bulk sample are approximated by weighted average of vertical and radial stresses along the central axis (the axisymmetry) using the vertical deformation in each element along the depth of the sample as weighing factor, equation1. This is based on the assumption that the granular material under the plunger is carrying the load most, thus the stress and strains along the central axis are considered as representative of the bulk. An illustrative stress and strain distribution along the depth of the specimen at the axis of symmetry is shown in Figure 4.

$$\sigma_V = \frac{\sum \sigma_{v,i} u_{v,i}}{\sum u_{v,i}} \qquad \sigma_h = \frac{\sum \sigma_{h,i} u_{v,i}}{\sum u_{v,i}} \tag{1}$$

Where $\sigma_V$ = vertical stress of the sample as a bulk [kPa]

$\sigma_h$ = horizontal or radial stress of the sample as a bulk [kPa]

$\sigma_{v,i}$ = vertical stress of each element along the axisymmetry [kPa]

$\sigma_{h,i}$ = horizontal stress of each element along the axisymmetry [kPa]

$u_{v,i}$ = vertical deformation of each element along the axisymmetry [mm]

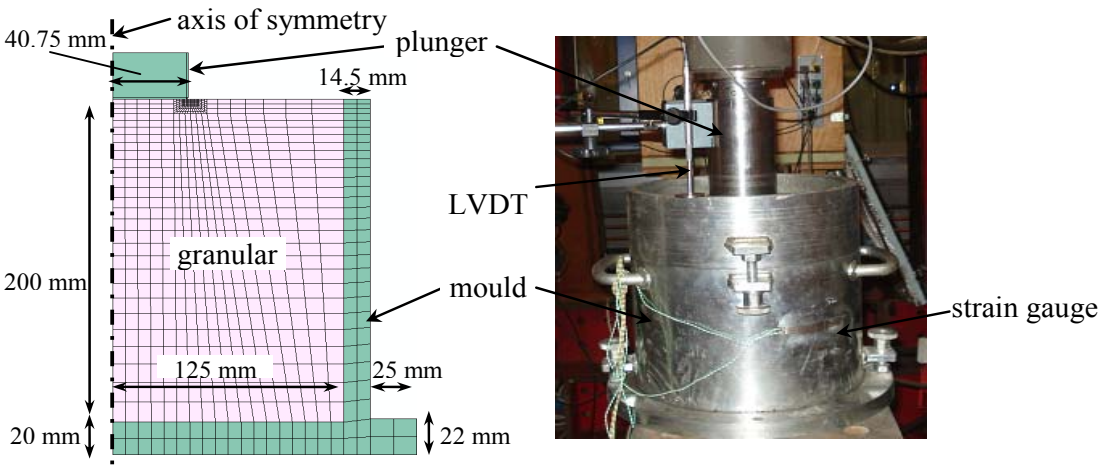



**Figure 3. RL-CBR Finite Element model mesh (left) and its test setup (right)**

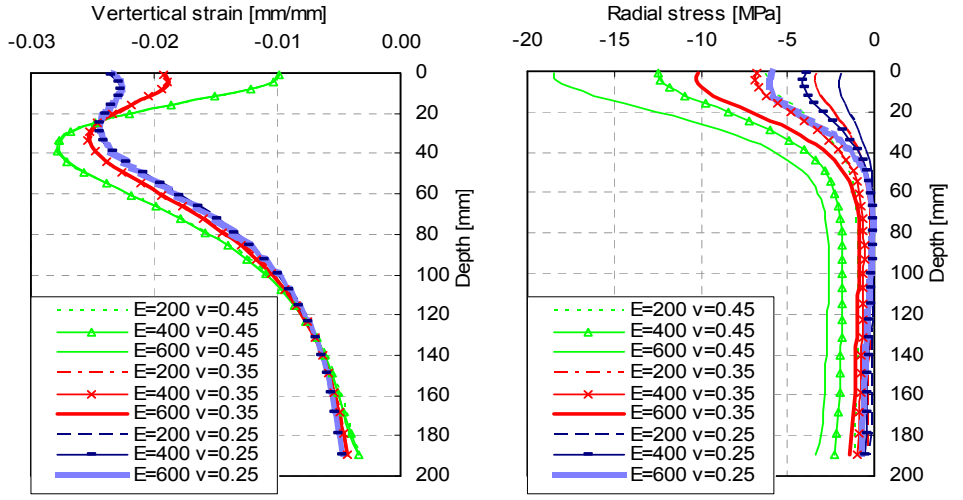



**Figure 4. Illustrative stress and strain distribution along the axisymmetry for a 2.5 mm plunger deformation**

Based on linear elastic theory for axisymmetric condition the following transfer functions were developed from the regression for the vertical and radial bulk stresses, the Poisson's ratio and elastic modulus. Note that in equation 2 below common soil mechanics sign convention is adopted with compressions as positive and tensile negative.

$$\sigma_V = k_1 \sigma_p \qquad \sigma_h = k_3 \varepsilon_{tt} \exp\left(k_4 / \nu\right)$$

$$\nu = k_2 \left(\frac{\varepsilon_{tt}}{\sigma_p}\right) \qquad E = \frac{k_5 \left(\sigma_V - 2\nu\sigma_h\right)}{u_v} \tag{2}$$

Where $\sigma_p$ = vertical plunger stress = total plunger load/ plunger area [kPa]

$\nu$ = Poisson's ratio [-]

E = Elastic or stiffness modulus [MPa]

$\varepsilon_{tt}$ = tangential strain at mid height of mould exterior [micro-strain]

$u_v$ = vertical plunger deformation [mm]

$k_1$ to $k_5$ = model parameters where: $k_1 = 0.368$ [-] $k_2 = -120.927$ [kPa]
$k_3 = 43.898$ [kPa] $k_4 = -0.072$ [-] $k_5 = 0.144$ [mm]

The regressions for the above four relations in equation 2 show a good fit with determination of correlation $r^2 > 0.99$ see Figure 5; this fit is of course an indication of relations of the parameters presented in the Finite Element model through the stiffness matrix (force and displacement relation), the kinematic compatibility (strain and displacement relation) etc under the given boundary conditions.

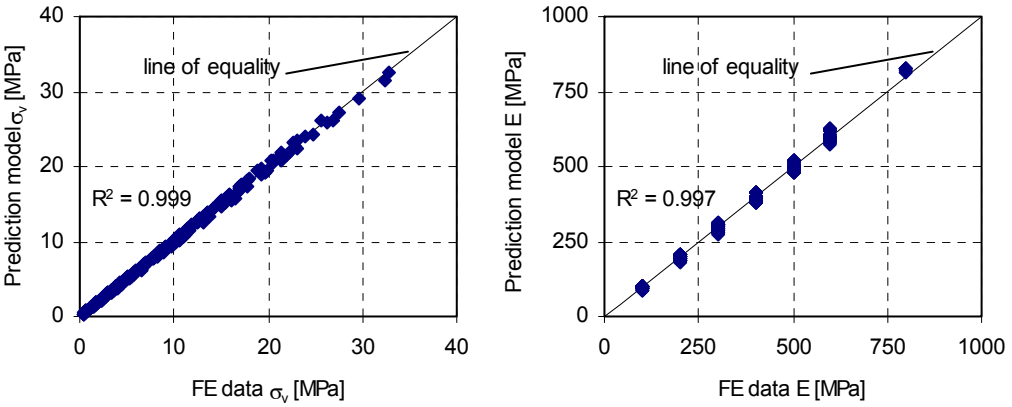



**Figure 5. Model prediction fit for illustrative transfer functions $\sigma_v$ and E**

**RL-CBR Equivalent modulus.** The equivalent stiffness modulus, $E_{equ}$, of the sample as a bulk is estimated, the same as the elastic modulus of the finite element analysis, as a function of the vertical (axial) stress, horizontal (radial) stress, the Poisson's ratio and exterior strain at mould mid-height as shown in equation 2. The advantage of such a model is that it is based from basic theory of elasticity Hooke's law where the model parameters are related to boundary conditions. Moreover it is express as stress dependent modulus similar to resilient modulus of cyclic load triaxial test models such as the $M_r$-$\Theta$ models, thus can be compared and validated with such model of triaxial results. For such purpose a cyclic load triaxial testing has been carried out for the same material.

## CYCLIC LOAD TRIAXIAL TEST

A large scale triaxial setup with a diameter of 300 mm and a height of 600 mm specimen was used in the study for testing the full 0/45 mm coarse material. The triaxial apparatus is equipped with a hydraulic loading system actuator and MTS controller capable of cycling the axial stress and with a partial vacuum constant confining pressure (CCP). The test is carried out according to the European Standard, EN13286-7 (CEN 2004) test protocol. The cyclic load signals used are a haversine at a loading frequency of 10 Hz for the first 20,000 load cycles of conditioning phase

and 1 Hz for the series of short loadings 100 cycles each. The stress range used is a ratio of axial stress to their respective failure axial stress, $\sigma_1/\sigma_{1,f}$ = 0.05 to 0.6, where the monotonic shear failure triaxial tests are carried out prior to the cyclic load triaxial tests.

The objective of the cyclic conditioning is to stabilize the permanent strains of the material and attain a practically elastic behavior. Generally the conditioning is performed with a stress level corresponding to the maximum cyclic and confining stresses applied in the test. The triaxial cell is equipped with transducers measuring the axial and radial strains on the middle third, 200 mm, of the specimen as shown in Figure 6. The resilient modulus ($M_r$) is defined as the ratio of the cyclic deviatoric stress ($\sigma_d$) to the recovered strain ($\varepsilon_r$):

$$M_r = \frac{\sigma_d}{\varepsilon_r} \tag{3}$$

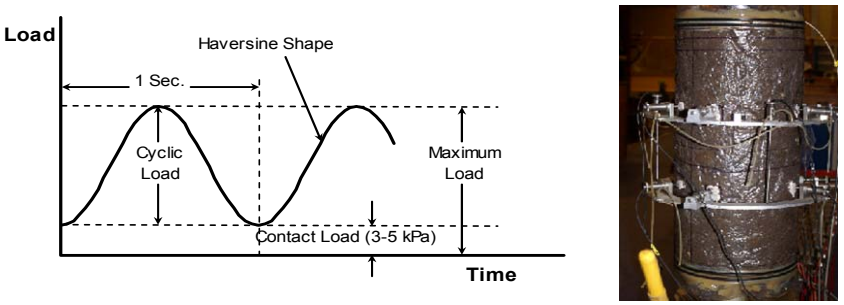



**Figure 6. Haversine triaxial cyclic loading curve (left) and instrumented triaxial specimen ready for testing (right)**

Similar to the RL-CBR test cyclic load triaxial testing were carried out for the ferricrete subbase material with varying the moisture content (MC) and degree of compaction (DOC as % MPDD) conditions. However for comparison with the RL-CBR test with strain gauges the cyclic load triaxial test result of the ferricrete material compacted with moderate (7%) MC at 98% DOC will be presented.

## RESULT AND DISCUSSION

**Cyclic triaxial test result.** The stress dependency of the resilient modulus was analyzed using the simple and well known isotropic non-linear $M_r$ – θ model for comparison with the result of the RL-CBR tests with strain gauge. For the ferricrete compacted at moderate MC and 98% DOC the $M_r$-θ in log-log scale is presented in Figure 7.

$$M_r = k_1\theta^{k_2} \tag{4}$$

Where $M_r$ = resilient modulus [MPa]
θ = bulk stresses = $\sigma_1+\sigma_2+\sigma_3$ [kPa]
$k_1$ & $k_2$ = model parameters

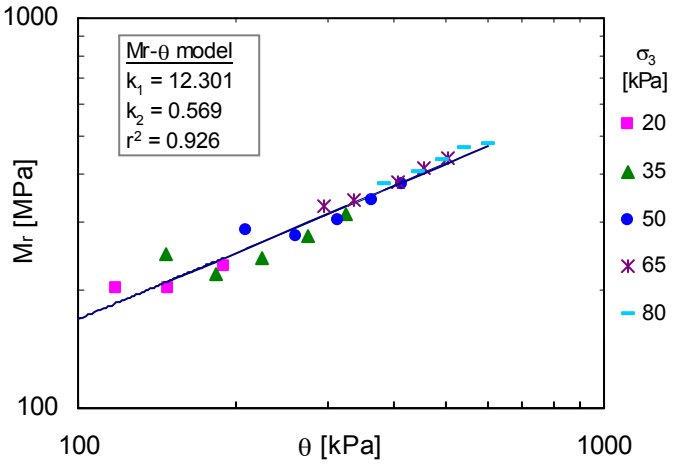



**Figure 7. Resilient modulus variation with bulk, θ, and confining stress, $\sigma_3$**

**RL-CBR test result.** To obtain stress dependent behavior from the RL-CBR test with strain gauges, large numbers of tests have been carried out at various plunger load levels. For the ferricrete with similar compaction conditions eight tests at different load level is conducted. From the RL-CBR laboratory test setup described earlier three parameters were measured the average plunger stress, $\sigma_p$, the plunger deformation, $u_v$, and the exterior strain at mould mid-height, $\varepsilon_{tt}$. The equivalent modulus ($E_{equ}$) is then computed using the transfer functions in equation 2, developed from the finite element analysis, with deviator values of $\sigma_p$, $u_v$ and $\varepsilon_{tt}$ between the maximum of loading and minimum of unloading of the last 5 cycles of the 100 cycles. The equivalent modulus is plotted verses the bulk stress θ, i.e. $\sigma_v + 2\sigma_h$ where $\sigma_v$ and $\sigma_h$ are in absolute values of a stress state of a specimen under testing. Although mould strain measurements have been recorded at mid-height and near the top of the mould at different locations as shown in Figure 2, the average of the two mid-height strain gauges is considered for the computation of the transfer function. The hoop strain at the mid-height gauges is higher than the gauges near the top of the mould which is in agreement with the result obtained from the FE analysis too. On the other hand, small variation strain measurement is observed among the gauges at the same height but in different vertical direction. However an average of the strain measurement is considered to compute the stress state for the entire bulk sample.

In Figure 8A the $E_{equ}$-θ model is presented along with the $M_r$-θ model of the triaxial test result from Figure 7. It can be recognized that the equivalent modulus from the RL-CBR test with strain gauge provides the stress dependent resilient behavior of the material. The stress state of a RL-CBR test specimen is uncontrolled and generally at a very high stress level due to the high confinement from the steel mould which results higher equivalent modulus compared to the resilient modulus of the triaxial. In addition the $E_{equ}$-θ model shows less stress dependent, or a gentle slope, than the $M_r$-θ. However, when granular materials loaded at much higher stress levels close to failure, the resilient modulus tends to decrease its stress dependency.

Figure 8B shows the equivalent modulus of the RL-CBR test appears as a continuation of the resilient modulus of the triaxial data with more scatter and at high stress level. This indicates that the RL-CBR test is a more complex form of a triaxial test that can provide a good estimate of the stiffness modulus.

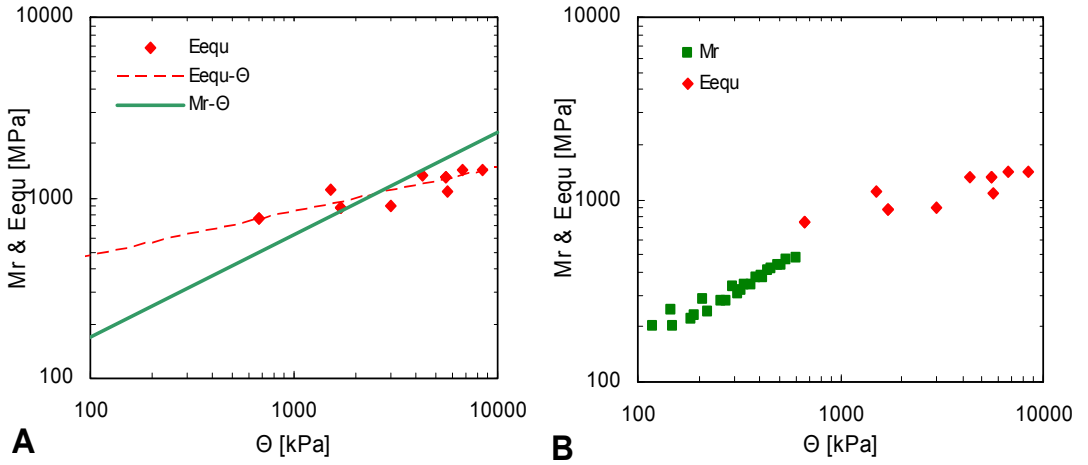



**Figure 8. Cyclic triaxial test resilient modulus, $M_r$, and RL-CBR test with strain gauges equivalent modulus, $E_{equ}$, as a function of bulk stress, θ.**

## CONCLUSIONS

In this paper, an intermediate (between the fundamental cyclic triaxial test and the traditional standard CBR test) method of characterization of unbound granular materials, the RL-CBR test, was discussed. It was shown that a good estimate of stress dependent modulus of the ferricrete granular subbase can be obtained with this characterization technique. The method can thus be used to characterize and estimate the resilient modulus of unbound granular materials, which can be used as an input in mechanistic design procedures, in the absence of fundamental test results.

It should be clear that non-linear finite element models may be developed, as unbound granular materials have a non-linearity behavior. However the intension of the study is to investigate an alternative simple way of characterization technique that can be implemented in common road engineering laboratories than assessing and modeling advanced material behaviors. Indications are that such advanced models will strongly contribute to the understanding of the stress-strain development in the complex CBR specimen and may provide a more fundamental material behavior.

It is to be noted that in the RL-CBR the granular arrangement or grain pattern in specimen preparation of the coarse particles, the big ratio between the bigger particle size and plunger diameter have a significant effect on the test result. Moreover the very slow rate of load application in the RL-CBR testing, comparing to wheel loading and the cyclic triaxial load, will have an effect on the relationship with the triaxial modulus and the field practice.

## REFERENCES

Araya, A. A., Molenaar, A. A. A., and Houben, L. J. M. "Characterization of Unbound Granular Materials Using Repeated Load CBR and Triaxial Testing." *GeoShanghai 2010 International Conference*, Shanghai, China, 355-363.

Brown, S. F. (1974). "Repeated load testing of a granular material." *Journal of the geotechnical engineering division*, 100(7), 825-841.

CEN (2004). "Unbound and hydraulically bound mixtures - Part 7: Cyclic load triaxial test for unbound mixtures." EN 13286-7, CEN, ed., European Committee for Standardization (CEN), Brussels.

De Beer, M. (1990). "Aspects of the design and behavior of road structures incorporating lightly cementious layers," Ph D thesis, University of Pretoria, Pretoria, South Africa

Edwards, J. P. (2007). "Laboratory Characterization of Pavement Foundation Materials," Loughborough University, Loughborough.

Elliott, R. P., and Thompson, M. R. (1985). "Mechanistic design concepts for conventional flexible pavements." University of Illinois, Urbana, Illinois.

Hicks, R. G. (1970). "Factors Influencing the Resilient Response of Granular Materials," University of California at Berkeley, Berkeley.

Huurman, M. (1997). "Permanent deformation in concrete block pavements," Delft University of Technology, Delft.

Lekarp, F., Isacsson, U., and Dawson, A. (2000). "State of the Art. I: Resilient Response of Unbound Aggregates." *Journal of Transportation Engineering, ASCE*, 126(1), 66-75.

Molenaar, A. A. A. "Are there any lessons to be learned from Pavement Research?" *Proceedings of the 8th Conference on Asphalt Pavements for Southern Africa*, Sun City, South Africa.

Monismith, C. L. (2004). "Evaluation of Long-Lasting Asphalt Pavement design Methodology: A Perspective." Distinguished Lecture International Society for Asphalt Pavements; Presented at International Symposium on Design and Construction of Long Lasting Asphalt Pavements, Auburn University, Alabama.

Smith, W. S., and Nair, K. (1973). "Development of Procedures for Characterization of Untreated Granular Base Coarse and Asphalt Treated Base Course Materials." *Rep. No. FHWA-RD-74-61, Federal Highway Administration, Washington, DC*.

Sweere, G. T. H. (1990). "Unbound Granular Bases for Roads," Dissertation, Delft University of Technology, Delft.

# Effects of Asphalt Pavement Instrumentation on In Situ Density

Fabricio Leiva-Villacorta[1] and David H. Timm[2], P.E., M.ASCE

[1]Graduate Research Assistant, Department of Civil Engineering, Auburn University, 238 Harbert Engineering Center, Auburn, AL 36849; PH (334) 844-4320; email: leivafa@auburn.edu
[2]Gottlieb Associate Professor of Civil Engineering, Auburn University, 238 Harbert Engineering Center, Auburn, AL 36849; PH (334) 844-6282; e-mail: timmdav@auburn.edu

## ABSTRACT

The National Center for Asphalt Technology (NCAT) Test Track has generated valuable information related to the installation of gauges, methodology for data collection and data analysis. However, information regarding the effect of pavement instrumentation on in situ properties is nonexistent. The objective of this study was to evaluate the effect of pavement strain gauges on the in situ density. In 2009, during construction of two test sections it was necessary to remove and replace the bottom lift since quality control targets were not met. Prior to removal, cores were taken in the gauge array and outside the array so that a two-sample T-test could be conducted for each section, respectively. The results showed that there were no significant differences between air void content and consequently no negative effect caused by the presence of pavement instrumentation. Additional density testing was conducted in and around the gauge array in these two test sections after completion of the surface layer. The data indicated no difference in measured density resulting from the presence of embedded instrumentation.

## INTRODUCTION

Pavement instrumentation is an important tool to monitor in situ pavement material performance and critical structural responses under environmental and loading conditions. The National Center for Asphalt Technology (NCAT) Test Track is a facility where instrumentation is used to study many issues pertaining to mechanistic-empirical (M-E) design. The NCAT Test Track is a 1.7 mile Pavement Test Facility and consists of 46 test sections with various asphalt mixtures. Ten million equivalent single axle loads are applied over a two-year period with field performance documented weekly. It has been in operation for ten years and the construction of the fourth experiment took place in 2009. Inclusion of instrumented sections started in 2003 (second cycle) where only eight sections were utilized for a structural experiment. Currently, that number has been doubled to sixteen sections.

Information obtained from the last two Test Track research cycles that is related to the installation of gauges, methodology for data collection and data analysis has been well documented (*Timm et al., 2004; Timm, 2008; Timm, 2009*). This information

includes full details of not only gauge installation but also analyses of variability on gauges responses (*Willis and Timm, 2009*). Gauges placed in different areas and tested under the same loading conditions can return different measurements based on material property (i.e. density) variability (*Willis & Timm, 2009*). However, information regarding the effect of pavement instrumentation on in situ properties and how it could contribute to material variability has not been studied.

A typical structural section at the Test Track contained 12 strain gauges and two pressure cells (Figure 1). One earth pressure cell is placed at the top of the base material layer and the other on top of the subgrade material layer. The strain gauge array is centered along the outside wheel path of the pavement structure. The array consists of two rows of three longitudinal gauges and two rows of three transverse gauges. Each asphalt gauge has an offset of 2 ft from each other (*Timm, 2009*). All these devices are connected to a data acquisition unit. During gauge installation, trenches for the conduits containing the cables and cavities for pressure cells were excavated. Prior to the placement of the next layer these elements were hand-backfilled and hand-compacted. Finally, all the asphalt strain gauges were covered with the respective mix sieved through the No. 4 sieve prior to compaction by rollers.

Installation of pavement instrumentation requires disruption and modification of material properties of the existing base layer. This process also incorporates complexity during construction which may involve time delays and additional difficulty in quality control (*Llenin et al., 2006*). Inclusion of objects with lighter density (asphalt strain gauges made out of nylon coated with a Teflon polymer) that do not belong to a typical pavement structure, and the use of finer material to cover them may provide a localized zone of low density and higher binder content. One study suggested that settlement over the gauge array due to poor compaction has led to premature failure in terms of fatigue cracking (Hugo et al., 1997). Another study suggested that variation in construction techniques seemed to have a significant effect on rutting performance (*Mulvaney, 2004*). These two studies were also performed at Accelerated Pavement Testing facilities with similar purposes to the NCAT Test Track and both studies pointed out the effect of construction on material properties and consequently on pavement performance. However, these issues have not been attributed, directly, to the presence of pavement instrumentation.

**Test Sections**

In 2009, during construction of two new test sections that included embedded instrumentation, it was necessary to remove and replace the bottom lift of asphalt since quality control targets were not initially met. These two sections are part of a pooled experiment where the asphalt thickness was designed to be seven inches over the same aggregate base and subgrade to enable a detailed analysis using the new Mechanistic-Empirical Pavement Design Guide (MEPDG). Therefore, tighter specifications and a detailed quality control process had to be enforced to meet the objective of this group of sections. The bottom lift of these two sections were approximately three inches of high recycled content (RAP = 50%) asphalt mix, built over consistent subgrade and granular base. The only difference between the mixtures

is that one was produced with hot mix asphalt technology (section N10) and the other one with a warm mix asphalt technology (section N11). When the bottom lift of material was first produced and placed, it was discovered that the asphalt contents were too high and the bottom lift had to be removed and replaced. This problem was attributed to varying moisture contents in the plant RAP stockpile that were not identified during original production. To maximize the knowledge gained from repaving, it was decided to core the pavement in and around the gauge array to evaluate the effects of pavement instrumentation on density.

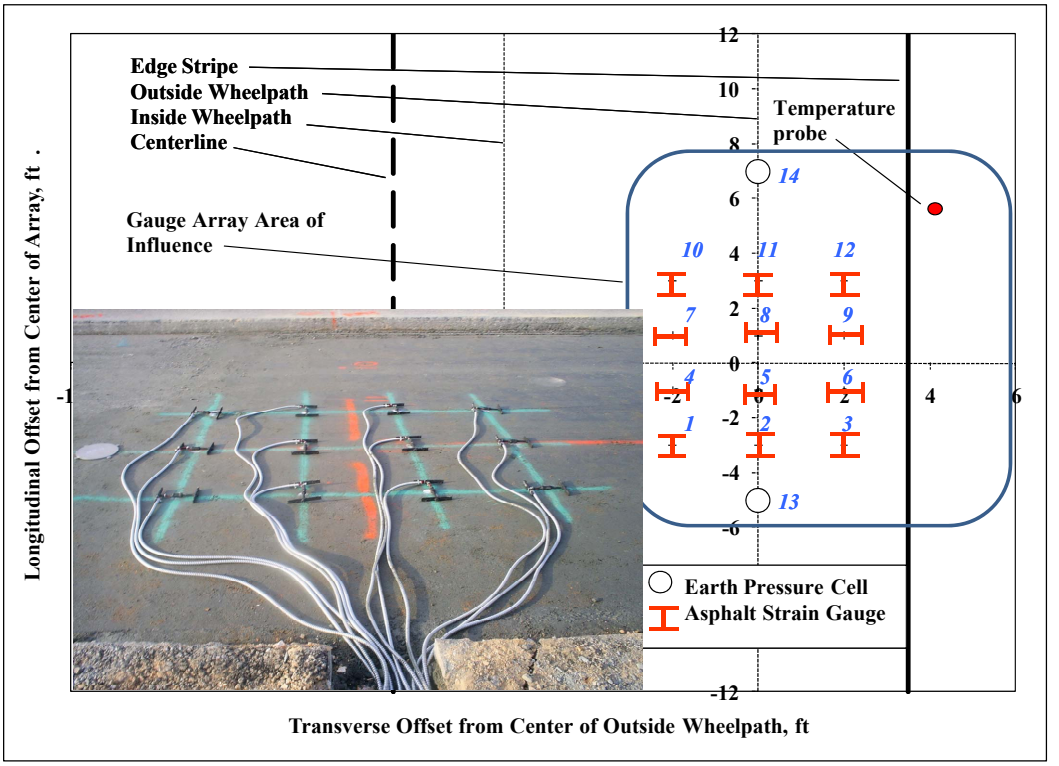



Figure 1: Sensor arraignment.

## OBJECTIVES

The main objective of this study was to evaluate the effect of pavement strain gauges on the in situ density of the pavement structure. The second objective was to perform an analysis of variability of the in-situ relative density.

## SCOPE

To accomplish the objectives, extraction of cores in two test sections that included embedded instrumentation was performed. Cores were taken from inside the gauge array and outside the array. Air void contents were compared using a two-sample T-test. At the end of construction, relative density was measured using the Pavement Quality Indicator (PQI) above the asphalt strain gauges embedded at the bottom of

the pavement. Density testing was also conducted in areas between gauges and outside the gauge array. An analysis of variance was used to evaluate material variability in terms of density.

## METHODOLOGY

### *Part 1: analysis of cores*

To develop the first part of this project an extraction of cores from the first lift of sections N10 and N11 was performed. Figure 2 shows the locations of the cores in relation to the gauge array for both sections. Six cores were taken outside the gauge array and six were taken within the area of influence of the gauge array. Air void contents were calculated for each core to be use as parameter of comparison. Bulk specific gravity was obtained using AASHTO T 166 (Bulk Specific Gravity of Compacted Bituminous Mixtures Using Saturated Surface-Dry Specimens) and the corresponding theoretical maximum specific gravity (rice density) was obtained using AASHTO T 209 . It should be emphasized that the coring was conducted after placement of the first lift of HMA of approximately three inches.

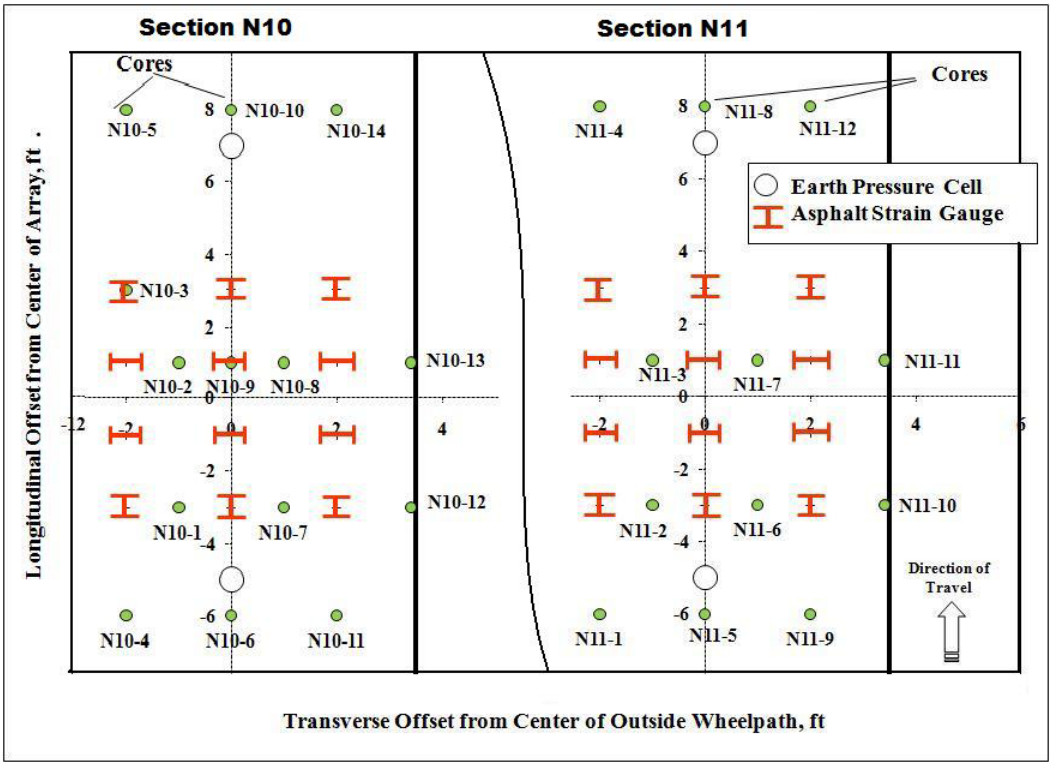



Figure 2: Core locations and air voids results for sections N10 and N11.

### *Part 2: analysis of relative density*

The second part of this investigation involved the use of a Pavement Quality Indicator (PQI) to acquire relative density in and around the gauge array in N10 and N11. Figure 3 shows the different locations where the relative density was obtained. Density testing was conducted on top of asphalt gauges, between gauges and around the gauge array.

As can be seen in Figure 3, five groups were formed to perform the analysis of variance: "Between" (between gauges), "Gauges" (on top of asphalt strain gauges), "Center Line", "Middle" and "Shoulder". These five groups were selected to analyze the change in density inside and outside the gauge array (all of them within the area of influence of the array). Along each two-hundred fot section there are randomly randomly selected stations within 50-foot sub-sections. These random stations are used for quality control and acceptance locations during construction. This factor was incorporated into the analysis to compare the variability of the density of the entire section against the testing area.

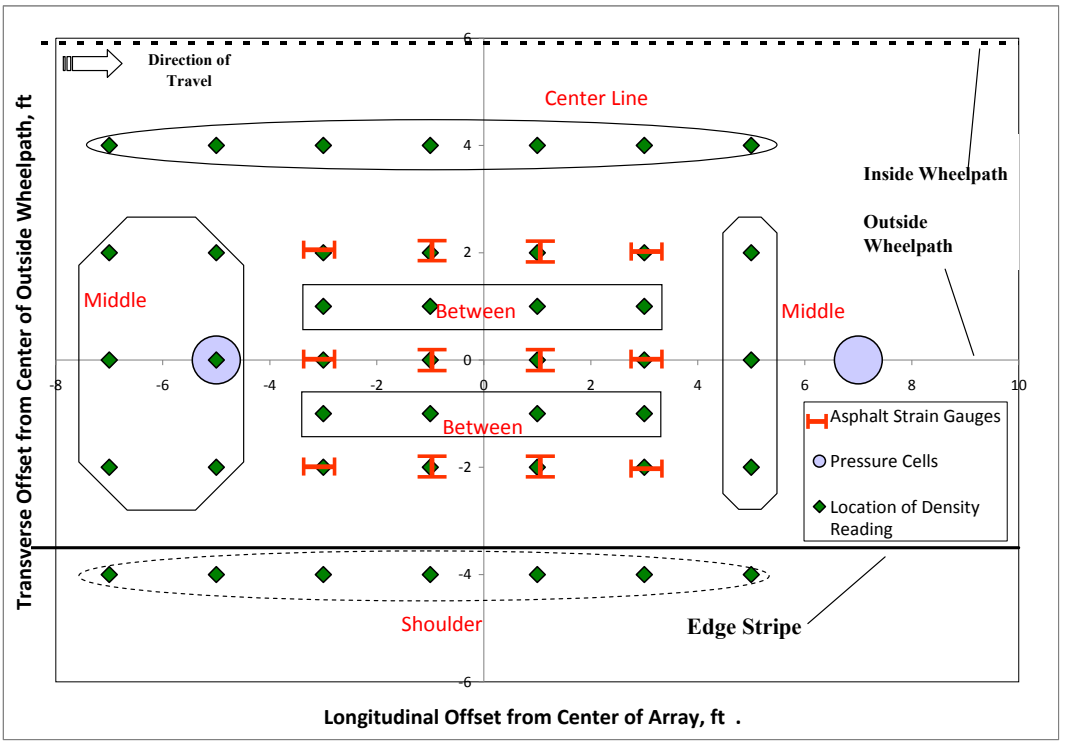



Figure 3: Density testing locations.

## RESULTS AND DISCUSSION

### *Part 1: analysis of cores*

Table 1 shows the value of air voids obtained for each extracted specimen. Two treatment groups were formed to use a two-sample t-test: OUT – outside gauge array, IN – inside gauge array. On average, air voids content for cores located outside (OUT) was 6.1% for section N10 and 6.7% for section N11, respectively. The average air voids content for cores located within the area of influence of the array (IN) was 6.8% for N10 and 6.4% for N11, respectively.

In the case of section N10, two cores were extracted from two different gauge locations. The rice density used to compute the air voids content for the cores IN and OUT was also applied to these two cores. The calculated air voids were 10.2% and 7.7%. These results were higher compared with the average IN-results of 6.8%. It was

determined that the lighter material placed on top of each gauge (passing the No. 4 sieve) and the lighter piece of gauge attached to the core were responsible for the higher air voids contents. Therefore, it would be expected to obtain lower results if rice values containing similar properties of the cores were used. Unfortunately there was not enough material to obtain the rice value for each specimen. When examining field specimens it was determined that there was visibly good compaction around the gauge itself as shown in Figure 4.

Table 1: Calculated air voids for each extracted specimen

| Section N10 | | | Section N11 | | |
|---|---|---|---|---|---|
| **Core ID** | **Location** | **Air Voids,%** | **Core ID** | **Location** | **Air Voids,%** |
| N10-1 | IN | 6.5 | N11-2 | IN | 6.3 |
| N10-2 | IN | 6.8 | N11-3 | IN | 6.1 |
| N10-7 | IN | 6.9 | N11-6 | IN | 6.4 |
| N10-8 | IN | 6.2 | N11-7 | IN | 6.5 |
| N10-12 | IN | 8.0 | N11-10 | IN | 6.4 |
| N10-13 | IN | 6.6 | N11-11 | IN | 6.7 |
| Average | | 6.8 | Average | | 6.4 |
| N10-4 | OUT | 6.3 | N11-1 | OUT | 6.5 |
| N10-5 | OUT | 5.8 | N11-4 | OUT | 6.2 |
| N10-6 | OUT | 5.7 | N11-5 | OUT | 7.4 |
| N10-10 | OUT | 5.4 | N11-8 | OUT | 6.8 |
| N10-11 | OUT | 6.8 | N11-9 | OUT | 6.3 |
| N10-14 | OUT | 6.5 | N11-12 | OUT | 6.9 |
| Average | | 6.1 | Average | | 6.7 |


Figure 4: Specimens containing parts of pavement instrumentation.

Figure 5 shows a comparison of mean air voids contents and the results of a two sample t-test for sections N10 and N11. For section N10 the calculated p-value for an average difference of 0.7% air voids was 0.05 and it was considered marginal. Further

results were considered before drawing a general conclusion. When the same statistical test was performed on section N11 there was no evidence to conclude that the factors IN and OUT were different. In other words, the mean air voids contents were statistically indistinguishable. Overall, the results indicated that the presence of pavement instrumentation was not responsible for differences in the parameter air voids.

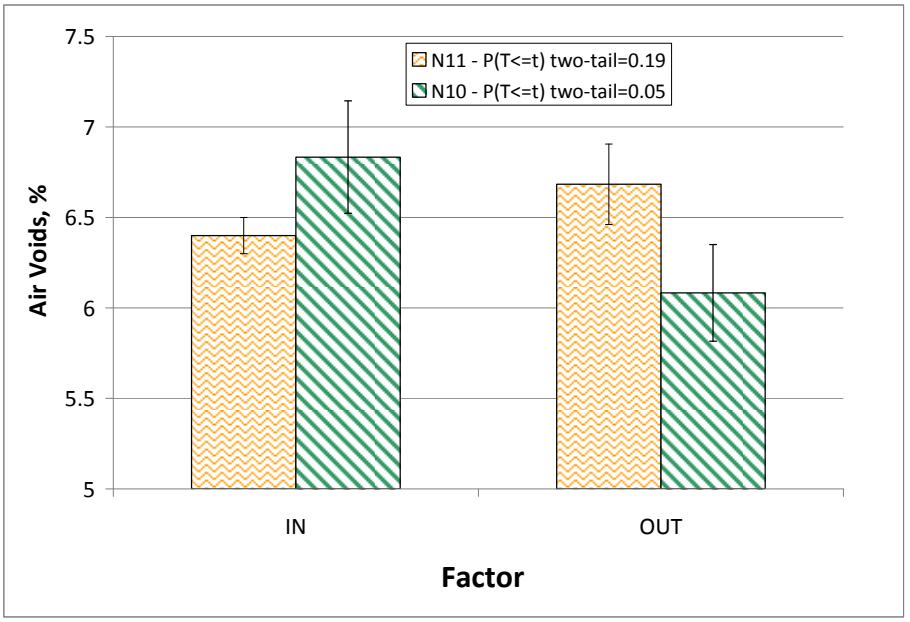



Figure 5: Two sample t-Test for sections N10 and N11.

***Part 2: analysis of relative density***

Figure 6 shows the density distribution of Section N10. First, a bar chart contains the average density of the five locations including the average density obtained at four random marks (RM) located along the entire section. Second, a contour plot shows the change in density along the studied area. Overall, it can be seen in Figure 6 that small differences on the mean density among groups with similar standard deviations (varying from 0.2 to 0.4) were characteristic of these sections. In addition, the contour plot shows a uniform density along the tested area with a density range equivalent to 2 pcf.

Figure 7 shows the density distribution of Section N11. Similar results were found for this section. Small differences in mean density were obtained among groups and uniform density was observed along the studied area. Also small standard deviations were obtained (varying from 0.2 to 0.5) and led to small range in density (2 pcf).

An analysis of variance shown in Table 2 indicated that there are not significant differences between factors (Between, Gauges, Center Line, Middle, Shoulder and Random Mark) for these two sections at a significance level of 5%. This means that the density measured along these three sections was not affected by the presence of pavement instrumentation. In addition, no significant difference among groups also

means that the mat is uniform along the entire section since the Random Mark group covers other representative areas of the section outside the tested area.

The analysis of variance also indicates that both sections had similar variability explained by the similar calculated MSError which is an estimate of the pooled variance (*Montgomery, 2008*). This also shows that the use of different technologies (HMA vs WMA) did not produce a significant difference in the variability of the in-situ density.

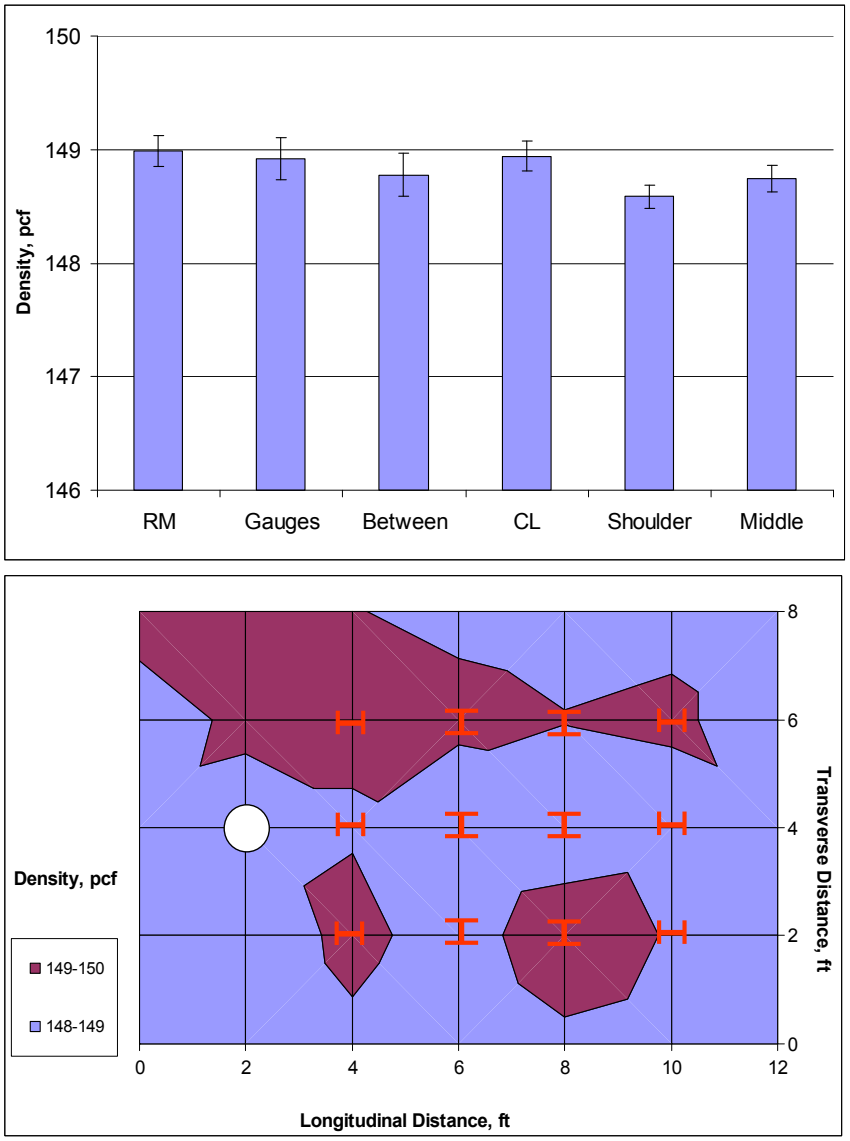



Figure 6: Density distribution for Section N10.

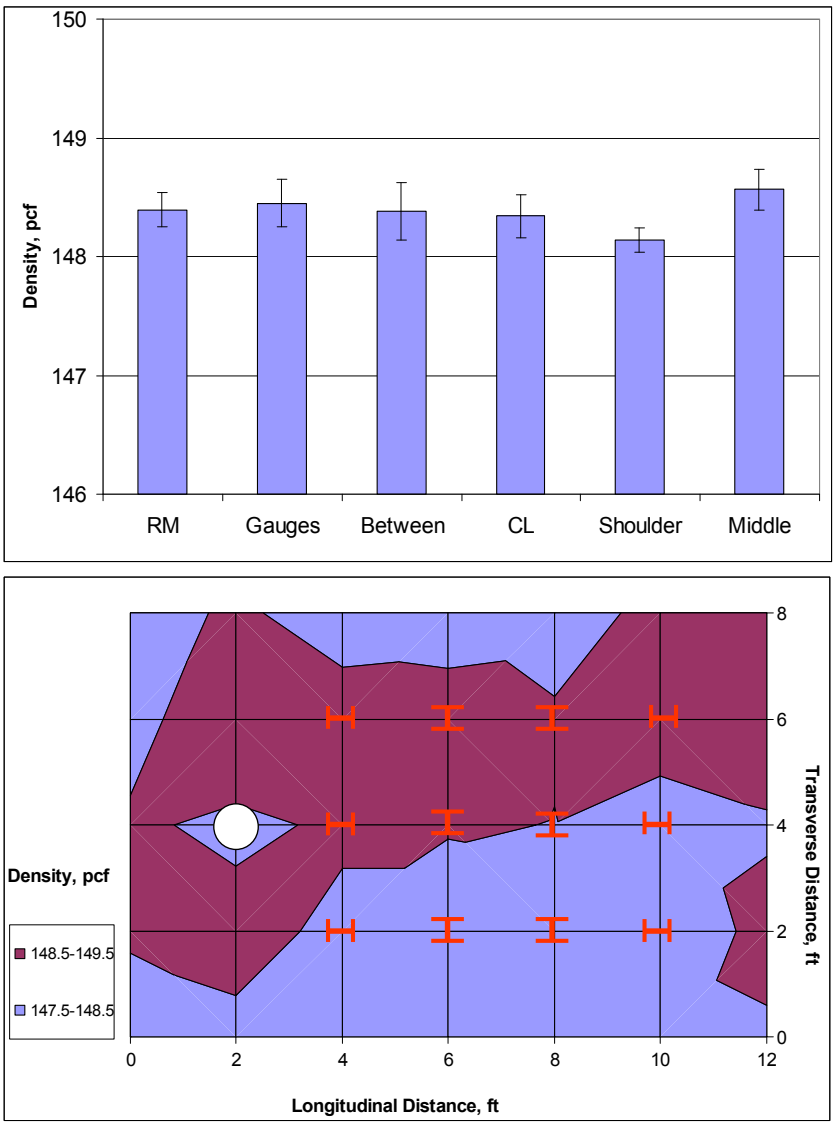



Figure 7: Density distribution for Section N11.

Table 2: ANOVA Table for Relative Density Analysis

| Section | Source | DF | SS | MS | F | P-value |
|---|---|---|---|---|---|---|
| N10 | Location | 5 | 1.00084 | 0.20017 | 2.24 | 0.064 |
| | Error | 52 | 4.64071 | 0.08924 | | |
| | Total | 57 | 5.64155 | | | |
| N11 | Location | 5 | 0.8624 | 0.1725 | 1.36 | 0.256 |
| | Error | 52 | 6.6068 | 0.1271 | | |
| | Total | 57 | 7.4691 | | | |

## CONCLUSIONS

The following conclusions can be drawn from this study:

- Statistical analysis performed on field samples taken from two sections provided evidence to conclude that the pavement instrumentation did not affect the overall air voids content of the area of influence of the gauge array.
- Analysis of variance performed on relative density provided enough information to disregard a negative effect on the mat relative density (in and outside the gauge array) due to the presence of asphalt gauges.
- Overall, the uniformity of the mat was not compromised by the presence of instrumentation devices.

## REFERENCES

Hugo, F., T. Scullion, N. Lee, K. Fults, and T. Visser. "A Rational Evaluation of Pavement Performance Using the Texas Mobile Load Simulator." *Proceedings of the 8th International Conference on Asphalt Pavements*, Seattle, Washington, 1997, Volume II, pp. 1125-1243.

Llenin, J.A, T.K. Pellinen, and D.M. Abraham. "Construction Management of a Small-Scale Accelerated Pavement Testing Facility." *Journal of Performance of Constructed Facilities*, Volume 20:3, August 2006, pp. 229-236.

Montgomery D. C., "*Design and Analysis of Experiments"*. 7th ed., John Wiley & Sons, INC. July, 2008.

Mulvaney, Ronald, Minnesota Department of Transportation; "*MnROAD Mainline Rutting Forensic Investigation"*. Second International Conference on Accelerated Pavement Testing. Minneapolis, Minnesota. September 2004.

Timm, D. H. "*Design, Construction and Instrumentation of the 2006 Test Track Structural Study"*. Draft report. National Center for Asphalt Technology, Auburn University, Ala., 2008.

Timm, D. H., A. L. Priest, and T. V. McEwen. "*Design and Instrumentation of the Structural Pavement Experiment at the NCAT Test Track"*. Report No. 04-01. National Center for Asphalt Technology, Auburn University, Ala., 2004.

Timm, D., "Design, *Construction and Instrumentation of the 2006 Test Track Structural Study"*. National Center for Asphalt Technology: NCAT Report 09-01, February 2009.

Willis, R. and Timm D., **"***Repeatability of Asphalt Strain Measurements Under Falling Weight Deflectometer Loading"*. Transportation Research Record: Journal of the Transportation Research Board, No. 2094, Washington, D.C., 2009.

# Flexible Pavement Quality Assurance Using Ground Penetrating Radar

Z. Leng[1] and I. Al-Qadi[2]
[1]Graduate Research Assistant, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 205 N. Mathews Ave., MC-250, Urbana, IL 61801; PH: (217)8930705; FAX: (217)8930601; email: zleng2@illinois.edu
[2]Founder Professor of Engineering, Director of Illinois Center for Transportation, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 205 N. Mathews Ave., Urbana, IL 61801; PH: (217)2650427; FAX: (217)8930601; email: alqadi@illinois.edu

## ABSTRACT

To produce a high-quality and long-lasting pavement, quality assurance (QA) is critical. For flexible pavements, layer thickness and in-situ density are the two most important characteristics to monitor. This paper presents the successful application of ground penetrating radar (GPR) to nondestructively and continuously measure asphaltic layer thickness and density. Ground penetrating radar surveys were conducted at a five-lane full-scale test site. Each lane is composed of four sections with the same asphaltic mixture but different densities. The GPR signal reflection was used to predict the layer thickness profile. Specific gravity models developed in an earlier study were utilized to predict the asphalt mixture density profile. The GPR-predicted thickness and density were compared to the core-measured values. Results showed that GPR achieved an average relative error of 4.9% in layer thickness measurement without using calibration cores and an average relative error of 2.6% in asphalt mixture bulk specific gravity using one calibration core.

## INTRODUCTION

To produce a high-quality and long-lasting pavement, quality assurance (QA) is critical. For flexible pavements, layer thickness and asphaltic mixture density are the two most important characteristics to monitor because they are highly related to the pavement performance. In construction practice, statistically based QA specifications are usually followed. The basic objective of these specifications is to specify and measure quality characteristics (such as thickness and in-situ density) that are related to pavement performance, and then to pay the contractor for the quality provided (Hand and Epps 2006). One of the main purposes of using statistically based specifications is to minimize the errors from random sampling because most of the available quality measurement methods, such as laboratory testing on field cores and in-situ density measurement by nuclear gauges, can only provide information at discrete test locations. Depending on the spacing of the sampling locations, an uncertainty or risk exists in the measurement and hence the decision because the population of the tested random samples is only a small fraction of the evaluated materials. Therefore, a reliable and rapid test

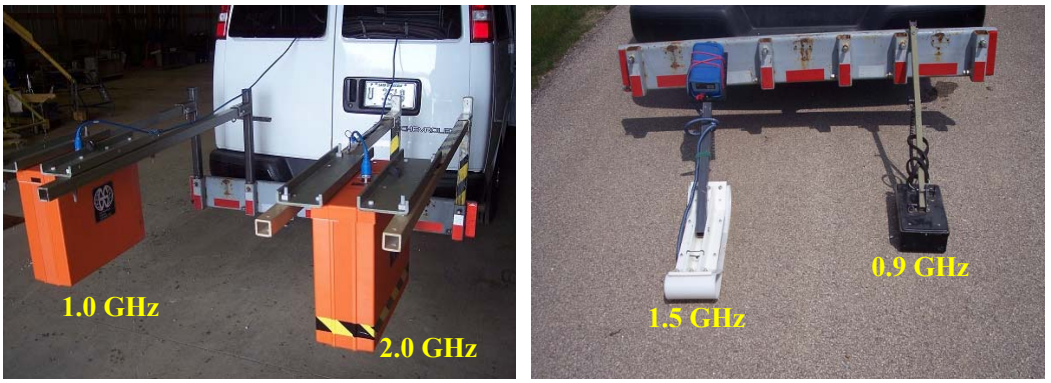
method that covers a relatively large area is desired to enhance the data confidence level.

The objective of this study is to investigate the performance of ground penetrating radar (GPR) as a QA tool for both asphaltic mixture layer thickness and in-situ density measurements, considering that GPR survey is nondestructive, rapid, and continuous. To achieve the objective of this study, a five-lane full-scale test site was designed and constructed having various layer thickness and density profiles. Pavement layer thickness and density were predicted using GPR measurements and validated by ground-truth measurements of field cores.

## GROUND PENETRATING RADAR SYSTEM

The application of GPR is based on transmitting electromagnetic (EM) signals toward the ground and receiving reflected signals from locations having dielectric contrast. From the received signals and the two-way travel times, GPR measurements can be used to obtain information about the pavement structure, such as the material layer thickness and subsurface defects.

As shown in Figure 1, the GPR antennae can be grouped into air-coupled and ground-coupled (Leng et al. 2009). Air-coupled antennae are usually used for pavement evaluation at highway speed and may provide high resolution if high frequency is used. However, to detect defects in deep subgrade, ground-coupled antennae may be used at relatively low frequency. Ground-coupled antennas provide greater penetrating depth than air-coupled antennae at the same frequency.




(a) Air-coupled antennae (b) Ground-coupled antennae

**Figure 1. Types of GPR antennae.**

For the same type of GPR system, the frequency of the antenna plays an important role in the penetrating depth of the GPR signal and the data resolution: the higher the antenna frequency, the lower the penetrating depth, and the higher the resolution. For pavement survey, the most commonly used central frequencies of the air-coupled antennae are 1.0 GHz and 2.0 GHz, and the central frequencies of the ground-coupled antennae are usually within the range of 80 MHz to 1.5 GHz (Saarenketo and Scullion 2000).

## THEORETICAL BACKGROUND ON ASPHLATIC MIXTURE LAYER THICKNESS AND DENSITY ESTIMATION USING GPR

For GPR applications, the most important material property is the dielectric constant. As shown in Eq. (1), a material's dielectric constant, $\varepsilon_r$, determines the GPR signal transmission speed, $v$, within this material.

$$v = \frac{c}{\sqrt{\varepsilon_r}} \tag{1}$$

where, $c$ is the speed of light in free space of 11.8 in/ns (30 cm/ns). Table 1 lists the typical dielectric constant values of some common materials which are related to pavement engineering (Al-Qadi 2009, Daniels 2004, and Jol 2009).

**Table 1. Typical Dielectric Constant Values.**

| Material | Dielectric Constant |
|---|---|
| Air | 1 |
| Water | 81 |
| Concrete | 3-18 |
| HMA | 3-10 |
| Asphalt | 3 |
| Limestone | 5-9 |
| Granite | 4-6 |
| Dry Sand | 3-5 |
| Saturated Sand | 20-30 |
| Silts | 5-30 |
| Clays | 5-40 |

For a layered structure, such as a flexible pavement, the dielectric constant of an asphaltic surface layer, $\varepsilon_{AM}$, can be estimated from the amplitudes of the reflected pulses (Figure 2) as follows (Al-Qadi et al. 2003, and Lahouar et al. 2002):

$$\varepsilon_{HMA} = \left( \frac{1 + A_o / A_p}{1 - A_o / A_p} \right)^2 \tag{2}$$

where, $A_p$ is the amplitude of the incident GPR wave obtained by collecting data over a copper plate placed on the surface of the pavement; and $A_o$ is the amplitude of the surface reflection. Once the dielectric constant of layer is known, the layer thickness can be calculated using the following equation:

$$d_{HMA} = \frac{c t_{HMA}}{2\sqrt{\varepsilon_{HMA}}} \tag{3}$$

where, $d_{HMA}$ is the asphaltic layer thickness and $t_{HMA}$ is the two-way travel time of the GPR signal within the asphaltic layer. Eq.s (2) and (3) have been successfully used in many studies to obtain high-accuracy pavement thickness profiles

(Al-Qadi et al. 2003, Al-Qadi and Lahouar 2004, Al-Qadi and Lahouar 2005, Saarenketo and Scullion 2000, and Maser 2006).

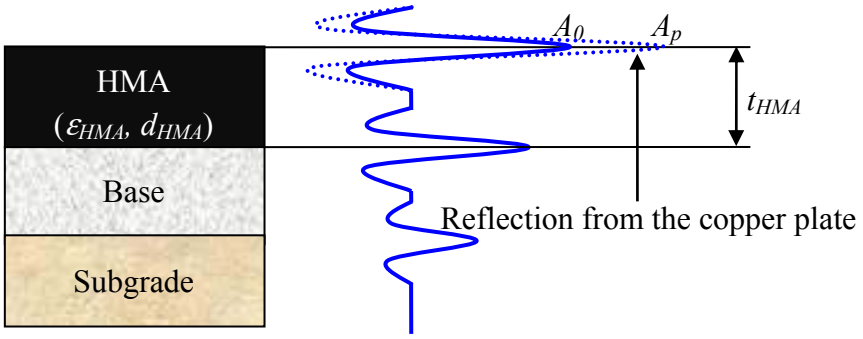



**Figure 2. Typical GPR signal for a flexible pavement system.**

To explore the feasibility of using GPR to measure the asphaltic mixture density or specific gravity, empirical exponential model between asphaltic materials air void content and its dielectric constant, based on statistical regression, was proposed (Saarenketo and Scullion 2000). Utilizing EM mixing theory, Al-Qadi et al. (2010) further developed several models to predict the bulk specific gravity of asphalt mixture. These models are based on the fact that the dielectric constant of a mixture is dependent on the dielectric and volumetric properties of its components. For asphaltic material, which is a mixture of air, asphalt, and aggregate, the compaction process reduces the air void content, hence, increases its bulk dielectric constant. Al-Qadi et al. (2010) in their study considered the complex refractive index model, Rayleigh model, and Bottcher model. The study found that the Rayleigh model, Eq. (4), performed the best based on laboratory test results. Hence, the Rayleigh model was selected in this study to predict the in-situ bulk specific gravity of asphaltic mixture.

$$G_{mb} = \frac{\dfrac{\varepsilon_{AM} - \varepsilon_b}{\varepsilon_{AM} + 2\varepsilon_b} - \dfrac{1-\varepsilon_b}{1+2\varepsilon_b}}{\left(\dfrac{\varepsilon_s - \varepsilon_b}{\varepsilon_s + 2\varepsilon_b}\right)\left(\dfrac{1-P_b}{G_{sb}}\right) - \left(\dfrac{1-\varepsilon_b}{1+2\varepsilon_b}\right)\left(\dfrac{1}{G_{mm}}\right)} \tag{4}$$

where, $G_{mb}$ and $G_{mm}$ are the bulk and maximum specific gravity of asphaltic mixture, $G_{sb}$ and $G_b$ are the bulk specific gravity of aggregate and binder, $P_b$ is the binder content, $\varepsilon_b$ is the dielectric constant of binder, and $\varepsilon_s$ is the dielectric constant of aggregate.

## TEST SITE DESIGN AND CONSTRUCTION

To evaluate the performance of GPR as a QA tool for measuring the asphaltic layer thickness and in-situ density, a five-lane full-scale test site was designed and constructed. The following variables were considered in the test site design: mix type, aggregate type, asphalt type, asphalt content, and air void content. Three basic mix designs, commonly used in construction practice, were selected: limestone surface mix, granite surface mix, and limestone binder mix;

PG64-22 binder was used. To evaluate the effect of asphalt type, a fourth mix was added using PG70-22 with the gravel surface mix. To evaluate the effect of asphalt content on GPR data, a fifth mix was used: the asphalt binder content of the gravel surface mix was increased by 1%. In total, five mixes, as presented in Table 2, were constructed. Each mix was compacted at four density level targets: 4%, 6%, 9%, and 12% air void content.

**Table 2. Mixes Used in the Testing Site.**

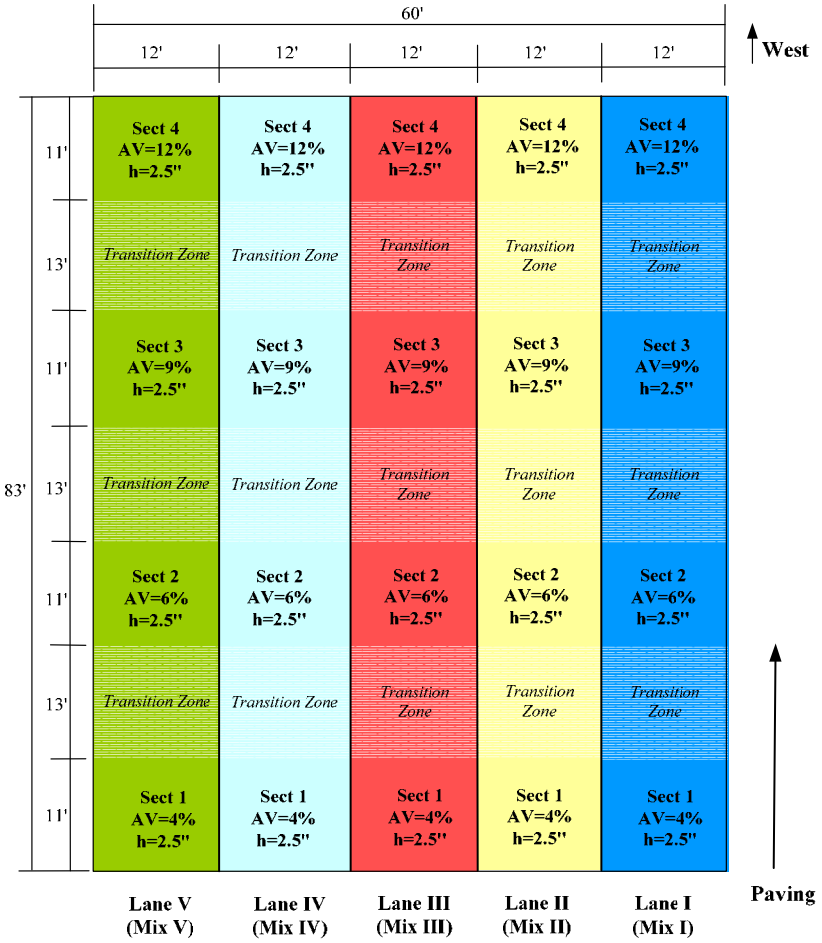
| Mix # | Mix Type | Asphalt Type | Asphalt Content |
|---|---|---|---|
| Mix I | Limestone Surface Mix | PG64-22 | Optimum |
| Mix II | Granite Surface Mix | PG70-22 | Optimum |
| Mix III | Granite Surface Mix | PG64-22 | Optimum |
| Mix IV | Granite Surface Mix | PG64-22 | Optimum+1% |
| Mix V | Limestone Binder Mix | PG64-22 | Optimum |

The test pavement sections were built on a large flexible pavement parking lot. As shown in Figure 3, five lanes, each of which has four sections at various densities, were constructed. Each section is 12 ft (3.6 m) wide and 11 ft (3.3 m) long. A 13 ft (3.9 m) long transition area separated each two adjacent sections to allow the compactor to maneuver to achieve uniform compaction in the test sections. To receive a clear GPR signal reflection at the bottom of the asphaltic layer, a 1/8 in (3.2 mm) thick steel plate was placed on the existing pavement surface prior to paving each section (Figure 4).

## DATA COLLECTION AND ANALYASIS

Ground penetrating radar surveys were conducted over each test lane using a 2.0 GHz air-coupled antenna. The GPR surveys were conducted at a speed of approximately 20 mph (32 km/h), and the data were collected at a rate of 1 scan per in (1 scan per 2.54 cm). The GPR survey line location is labeled in Figure 4, and a typical raw GPR image of the test lane is shown in Figure 5. In Figure 5, the plateau-shapes are the strong reflections of the GPR signals from the steel plates underneath the asphaltic surface, which show the accurate locations of the bottom of the surface layer.

After the raw GPR data was collected, the dielectric constant profile of each test lane was obtained using Eq. (2) and is shown in Figure 6. The dielectric constant of each lane exhibits a decreasing trend along the survey direction, from the high-density section to the low-density section. In general, the lanes composed of limestone mixes (Lane I and Lane V) have greater dielectric constant than the lanes having granite mixes.




*AV and h represent the target air void and layer thickness, respectively

**Figure 3. Asphaltic material overlay construction layout.**

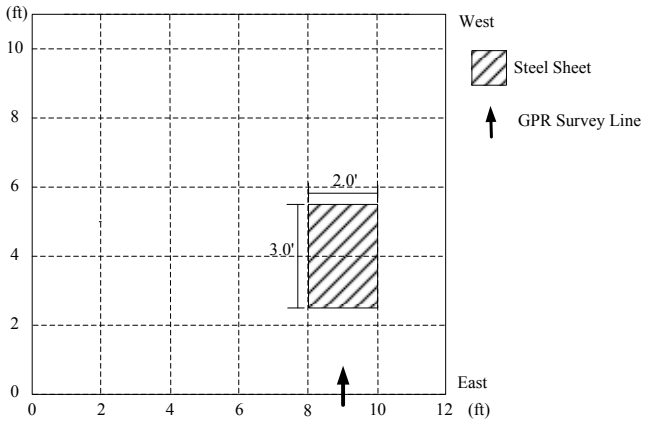



**Figure 4. Locations of steel sheet in each section.**

With known dielectric constant profiles, the thickness profile of each test lane was predicted using Eq. (3), as shown Figure 7. To validate the accuracy of the predicted values, one core was extracted from each test section above the steel plate. The locations of the cores are shown as dashed lines in Figure 7. The predicted thicknesses at the coring locations were compared to the measured

thickness of the cores. As shown in Table 3, the relative prediction error of thickness is within the range of 0 to 12.4% and the average is 4.9%. It should be noted that the overlay was placed on flexible pavement, which makes it difficult to detect the bottom of the surface layer. A more accurate reading can be expected if the underlying layer is composed of a non-asphaltic material.

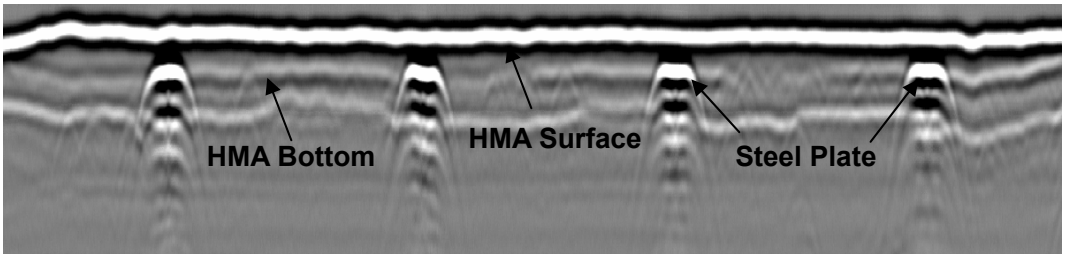



**Figure 5. Typical GPR image for the test lane.**

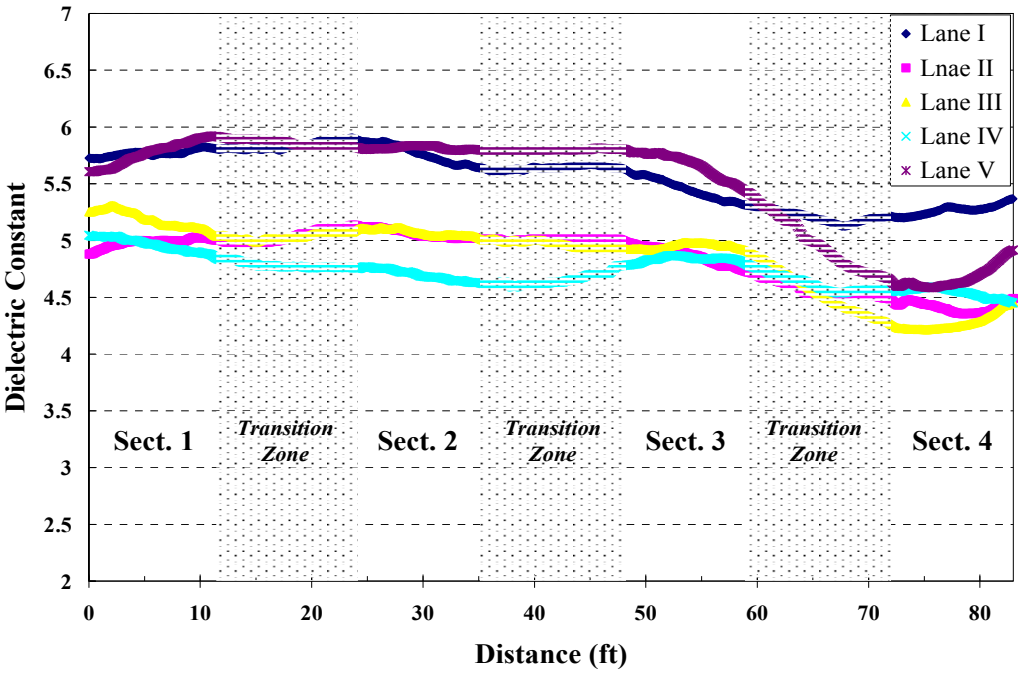



**Figure 6. Asphaltic mixture dielectric constant profile of each test lane.**

In addition to the thickness profiles, the asphaltic materials density profiles were predicted based on the measured dielectric constants using Eq. (4) and is shown in Figure 8. In the right-hand side of Eq. (4), the values of five variables can either be obtained from the mix design ($P_b$, $G_{sb}$, $G_{mm}$) or considered as a constant ($G_b$, $\varepsilon_b$), as shown in Table 4. However, the value of $\varepsilon_s$ depends on the aggregate type (Table 1). Therefore, the density and dielectric constant of the core obtained from Section 1 of each lane were utilized to back-calculate the value of $\varepsilon_s$. Table 4 shows the back-calculated values of $\varepsilon_s$ of each mix. The accuracy of the predicted $G_{mb}$ (Figure 8) was validated by comparing the measured $G_{mb}$ of the cores from Sections 2, 3, and 4 to their GPR-predicted values. As shown in Table

3, the average prediction error of the asphaltic mixture bulk specific gravity is 2.6%. The prediction errors for Section 4 materials are generally larger than those for Sections 2 and 3. This is due to the test method used to measure $G_{mb}$ in the lab, which is the saturated surface dry method (ASSHTO 2001). This method usually works well for specimens with lower air void contents (less than 7%), but it may result in high error for specimens with large air void contents.

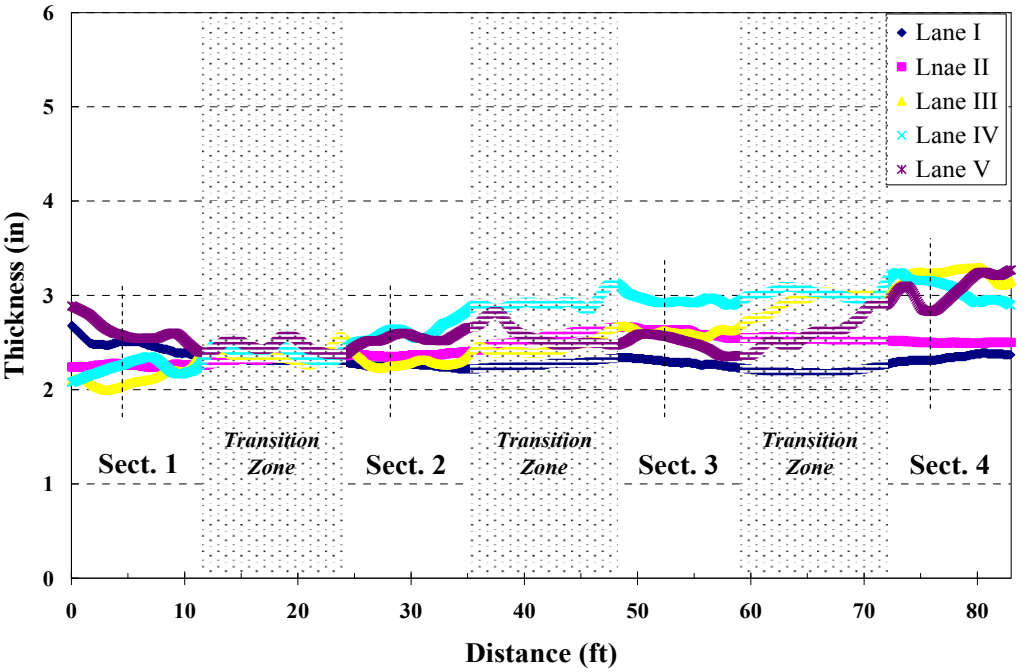



**Figure 7. Asphaltic surface thickness profile of each test lane.**

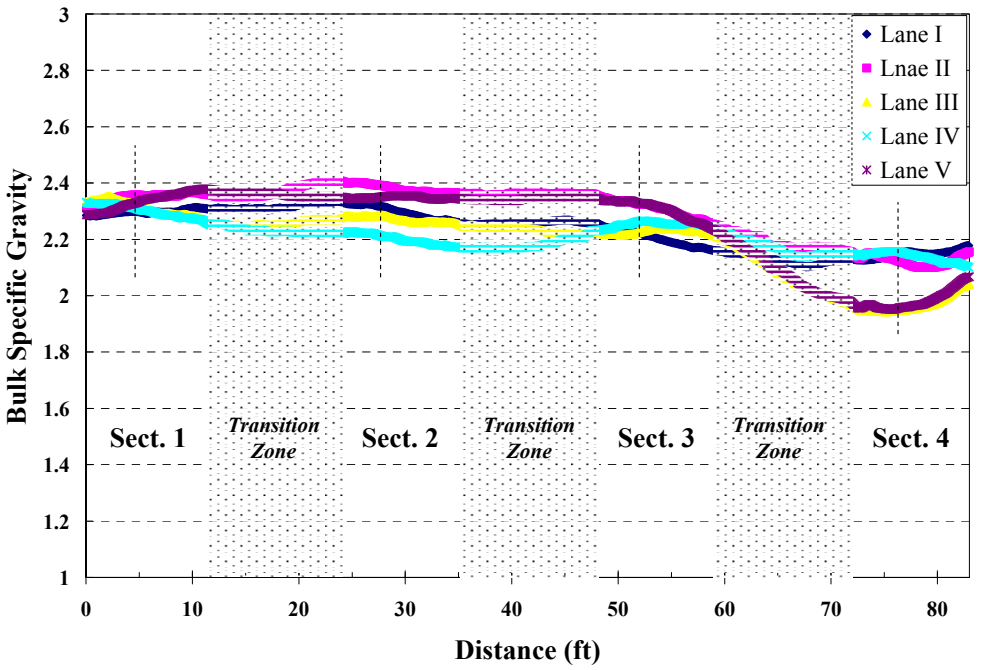



**Figure 8. Asphaltic bulk specific gravity profile of each test lane.**

**Table 3. Relative Errors in Asphaltic Layer Thickness and $G_{mb}$ Prediction.**

| Core # | Core Thickness (in) | GPR Thickness (in) | Relative Error (%) | GPR $G_{mb}$ | Core $G_{mb}$ | Relative Error (%) |
|---|---|---|---|---|---|---|
| I-1 | 2.32 | 2.49 | 7.2 | 2.298 | 2.298 | - |
| I-2 | 2.21 | 2.25 | 1.7 | 2.315 | 2.297 | 0.8 |
| I-3 | 2.32 | 2.30 | 0.9 | 2.214 | 2.211 | 0.1 |
| I-4 | 2.15 | 2.35 | 9.5 | 2.152 | 2.136 | 0.7 |
| II-1 | 2.30 | 2.27 | 1.2 | 2.355 | 2.355 | - |
| II-2 | 2.27 | 2.35 | 3.4 | 2.389 | 2.357 | 1.3 |
| II-3 | 2.62 | 2.63 | 0.5 | 2.325 | 2.297 | 1.2 |
| II-4 | 2.42 | 2.49 | 2.7 | 2.121 | 2.217 | 4.4 |
| III-1 | 2.23 | 2.06 | 7.7 | 2.333 | 2.333 | - |
| III-2 | 2.43 | 2.27 | 6.6 | 2.285 | 2.265 | 0.9 |
| III-3 | 2.73 | 2.74 | 0.6 | 2.229 | 2.232 | 0.1 |
| III-4 | 3.08 | 3.24 | 5.1 | 1.948 | 2.166 | 10.1 |
| IV-1 | 2.26 | 2.36 | 4.4 | 2.315 | 2.315 | - |
| IV-2 | 2.61 | 2.74 | 5.2 | 2.209 | 2.252 | 1.9 |
| IV-3 | 2.86 | 2.99 | 4.5 | 2.265 | 2.234 | 1.4 |
| IV-4 | 2.77 | 3.11 | 12.4 | 2.157 | 2.184 | 1.2 |
| V-1 | 2.49 | 2.61 | 4.7 | 2.323 | 2.323 | - |
| V-2 | 2.55 | 2.64 | 3.5 | 2.352 | 2.304 | 2.1 |
| V-3 | 2.48 | 2.60 | 4.8 | 2.328 | 2.277 | 2.2 |
| V-4 | 2.63 | 2.91 | 10.7 | 1.959 | 2.205 | 11.2 |
| | Ave. Prediction Error in Thickness (%) | | 4.9 | Ave. Prediction Error in $G_{mb}$ (%) | | 2.6 |

**Table 4. Determination of the Aggregate Dielectric Constants.**

| Core # | $\varepsilon_{HMA}$ | $G_{mb}$ | $P_b$ | $G_b$ | $G_{sb}$ | $G_{mm}$ | $\varepsilon_b$ | $\varepsilon_s$ |
|---|---|---|---|---|---|---|---|---|
| I-1 | 5.7716 | 2.298 | 6.0 | 1.015 | 2.603 | 2.481 | 3 | 7.0 |
| II-1 | 4.9941 | 2.355 | 5.4 | 1.015 | 2.641 | 2.528 | 3 | 5.8 |
| III-1 | 5.2503 | 2.333 | 5.4 | 1.015 | 2.641 | 2.501 | 3 | 6.2 |
| IV-1 | 5.0019 | 2.315 | 6.4 | 1.015 | 2.641 | 2.468 | 3 | 5.9 |
| V-1 | 5.7267 | 2.323 | 5.1 | 1.015 | 2.607 | 2.505 | 3 | 6.8 |

## SUMMARY

This study presents the successful application of GPR as a QA tool for flexible pavement construction. Based on the GPR data collected using a 2.0 GHz air-coupled antenna, the asphaltic layer thickness and density profiles were predicted for a five-lane test site. The study resulted in the following conclusions:

- The asphaltic mixture dielectric constant is dependent on the aggregate type used in the mix. The larger the dielectric constant of the aggregate, the higher the dielectric constant of mix.

- The dielectric constant of the asphaltic material is affected by its density. For the same mix, the higher the density, the greater the measured dielectric constant.
- Ground penetrating radar predicted asphaltic layer thickness with an error of 4.9% without using calibration cores.
- The specific gravity model developed in an early study provides an effective way to predict asphaltic materials specific gravity profiles using GPR-measured dielectric constants. An average prediction error of 2.6% was obtained using the Rayleigh model when using a calibration core.

## ACKNOWLEDGEMENT

The authors would like to acknowledge the input of Al Larkin of the Federal Aviation Administration (FAA). This publication is based on the results of an FAA project, Non-Destructive Testing and Evaluation (NDTE) Technologies for Airport Pavement Acceptance and Quality Assurance Activities. The FAA project is conducted in cooperation with the Center of Excellence for Airport Technology (CEAT) and the University of Illinois at Urbana-Champaign. The contents of this study reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of CEAT or FAA. This paper does not constitute a standard, specification, or regulation.

## REFERENCES

Al-Qadi, I. L. (2009). *Class Notes for Pavement Maintenance and Rehabilitation,* University of Illinois at Urbana-Champaign.

Al-Qadi, I. L., Lahouar, S., and Loulizi, A. (2003). "Successful application of ground penetrating radar for quality assurance-quality control of new pavements." *Transportation Research Record: Journal of the Transportation Research Board* 1861, Transportation Research Board of the National Academies, Washington, D.C., 86-97.

Al-Qadi, I. L., and Lahouar, S. (2005). "Measuring layer thicknesses with GPR-theory to practice." *Construction and Building Materials* 19(10), 763-772.

Al-Qadi, I. L., and Lahouar, S. (2004). "Use of GPR for thickness measurement and quality control of flexible pavements." *The Journal of APPT*, 501–528.

Al-Qadi, I. L., Leng, Z., Lahouar, S., and Baek, J. (2010). "In-place hot-mix asphalt density estimation using ground-penetrating radar." *87th Annual Meeting of the Transportation Research Board (CD-ROM)*, Transportation Research Board of the National Academies, Washington, D.C.

American Association of State Highway and Transportation Officials (2001). *Standard Specifications for Transportation Materials and Methods of Sampling and Testing.* (22nd Ed.), AASHTO, Washington, D.C.

Daniels, D. J. (2004). *Ground Penetrating Radar,* The Institution of Electrical Engineers, London, United Kingdom.

Hand, A. J. T., and Epps, J. A. (2006). "Fundamentals of percent within limits and quality control-quality assurance compaction specifications." *Factors Affecting Compaction of Asphalt Pavements*, Transportation Research Board of the National Academies, Washington, D.C., 140-162.

Jol, H. M. (2009). *Ground Penetrating Radar: Theory and Applications*, Elsevier Science Ltd., Oxford, United Kingdom.

Lahouar, S., Al-Qadi, I. L., Louliz, A., Clark, T. M., and Lee, D. T. (2002). "Approach to determining in situ dielectric constant of pavements: development and implementation at interstate 81 in Virginia." *Transportation Research Record: Journal of the Transportation Research Board* 1806, Transportation Research Board of the National Academies, Washington, D.C., 81-87.

Leng, Z., Al-Qadi, I. L. Lahouar, S., and Baek, J. (2009). “Selection of antenna type and frequency for pavement survey using ground penetrating radar.” *86th Annual Meeting of the Transportation Research Board (CD-ROM)*, Transportation Research Board of the National Academies, Washington, D.C.

Maser, K. R., T. J. Holland, Roberts, R., and Popovics, J. (2006). "NDE methods for quality assurance of new pavement thickness." *International Journal of Pavement Engineering* 7(1), 1-10.

Saarenketo, T., and Scullion, T. (2000). "Road evaluation with ground penetrating radar." *Journal of Applied Geophysics* 43(2-4), 119-138.

# Evaluation of the Rolling Wheel Deflectometer as a Structural Pavement Assessment Tool in Louisiana

**Mostafa A. Elseifi [1], Ahmed Abdel-Khalek[2], Kevin Gaspard[3], Zhongjie "Doc" Zhang[4], and Said Ismail[5]**

**Abstract:** The Rolling Wheel Deflectometer (RWD), which measures pavement surface deflection at traffic speed, offers the potential to characterize the structural conditions of the road network without major delays and in a cost-effective way. The objective of this study was to evaluate the repeatability of RWD deflections and to determine the effects of test speed and pavement conditions on the measured deflections. To achieve this objective, 16 different test sites representing a wide array of pavement conditions were tested. Results of the field testing program determined that the repeatability of RWD measurements was acceptable with an average coefficient of variation of 14% at a testing speed of 32 km/h. The influence of testing speed on the measured deflections was minimal. Since the test speed is restricted by the posted speed limit, testing can be conducted at different speeds while allowing for direct comparison of the measured deflections. RWD deflection measurements appear to properly reflect pavement conditions and structural integrity of the road network by providing for a greater average deflection and scattering for sites in poor conditions.

[1] Assistant Professor, Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA 70803, email: elseifi@lsu.edu
[2] Graduate Research Assistant, Louisiana State University
[3] Senior Pavement Research Engineer, Louisiana Transportation Research Center, LSU, 4101 Gourrier Ave., Baton Rouge, LA 70808
[4] Pavement Geotechnical Research Administrator, Louisiana Transportation Research Center, LSU, 4101 Gourrier Ave., Baton Rouge, LA 70808
[5] Management Systems Engineer, Louisiana Department of Transportation and Development, Office of Planning and Programming

## Introduction

The current Pavement Management System (PMS) for the Louisiana Department of Transportation and Development (LADOTD) is based on pavement condition measurements that are collected once every two years using the Automatic Road Analyzer (ARAN®) system that provides a continuous assessment of the road network (Khattak et al. 2008). Conditions of the pavement are assessed using cracking, rutting, roughness, patching, and faulting measurements, which are collected for every state highway in Louisiana. Collected data are reported every 0.16 of a km and are analyzed to calculate a Condition Pavement Index (CPI) on a scale from zero to 100. The CPI varies from 95 to 100, 85 to 94, 65 to 84, 50 to 64, and 49 or less for very good, good, fair, poor, and very poor roads, respectively.

The current PMS system may be substantially improved if the structural conditions of in-service pavements are considered in selecting suitable maintenance and rehabilitation methods. Structural assessment of in-service pavements would help avoid applying routine preventive treatment techniques on structurally-deficient pavements. Surface deflection measured using the Falling Weight Deflectomer (FWD) is a popular method that allows to determine the structural capacity of in-service pavements at the project level. However, due to the associated cost and the slow testing process, the use of deflection testing in network pavement management activities has been limited. In addition, lane closures due to FWD testing may compromise the safety of the traveling public and highway workers, which may not justify the worth of the collected data.

The Rolling Wheel Deflectometer (RWD) offers great potential benefits for continuous deflection measurements at the network level in order to assess the structural conditions of in-service pavements and to identify uniform and homogeneous pavement sections. This innovative system, which measures deflections at traffic speeds, offers the potential to characterize the structural integrity of the road network without major delays and in a cost-effective way. In spite of these promising benefits, the repeatability of the measurements needs to be evaluated as well as the relationship between deflection measurements and pavement conditions of the roadway.

This study describes a detailed field evaluation of the RWD system in Louisiana through which 16 different test sites representing a wide array of pavement conditions were tested. To quantify the repeatability, deflection measurements were collected in triplicates, which allowed estimating the variability in the measurements and its relationship to pavement conditions. Comparison was also established between RWD measurements and pavement conditions as described by the Pavement Condition Index (PCI). The collected data were incorporated in a deflection-based Geographical Information System (GIS) map that was developed to illustrate the use of this technology at the network level.

## Background

The rolling wheel deflectometer was developed by Applied Research Associates (ARA), Inc. to measure pavement surface deflections at traffic speeds and to characterize the load carrying capacity of in-service pavements. The first Qwest /

Dynatest RWD prototype was introduced in the late 1990s and was designed to perform measurements on airfield pavements at a maximum speed of 9 km/h (Briggs et al. 2000). The latest version of the RWD was introduced in 2003 and can collect deflections at traffic speeds. It consists of a 16.2-m long semitrailer applying a standard 18 kN load on the pavement structure over the rear single axle using fixed steel plates, see Figure 1. As RWD travels on top of the pavement, triangulation lasers mounted on a 64.8 cm aluminum beam and placed at 20.3-cm intervals are used to measure surface deflections. The beam is mounted on the right side of the semitrailer to follow the right wheel path on the right lane, which is usually the weakest location on the pavement structure. Three spot lasers are placed in front of the loaded wheel to define the unloaded surface and one spot laser is placed right on top of the loaded dual-tire assembly to measure the deflected surface. The laser sensors are set to collect a reading at a fixed interval of 15 mm at all truck speeds. Prior to the field testing program described in this study, a more accurate and stable deflection measurement system customized for pavement applications was installed. The upgraded system has a 100-mm measurement deflection range. This was the first testing program conducted with the new and improved laser deflection system.


**Figure 1.** General Overview of the Rolling Wheel Deflection System

Since its introduction in 1996, the RWD has gradually transitioned from conceptual testing to real-world applications by incorporating its measurements into PMS activities at the network level (Groog 2004). Accordingly, a number of studies were conducted to evaluate RWD technology in the assessment of in-service pavements at the network level. Gedafa and co-workers evaluated the use of RWD for network deflection measurements in Kansas and compared these data to FWD deflection measurements (Gedafa et al. 2008). The authors highlighted that remaining life models adopted in Kansas make use of one input from pavement surface deflection

requiring time-consuming measurements with FWD. The use of the RWD would allow for deflection testing at regular speeds without traffic protection or lane closure. Results of this study showed that there is no significant difference between center deflections measured using the RWD and the FWD. Since the structural conditions of the pavement network did not change significantly over a 4-year period, the authors recommended collecting RWD deflection data at 4-year intervals.

A study was conducted by Vavrik et al. (2008) to develop a methodology that incorporates RWD measurements into PMS activities. In this study, the authors developed a PMS for Champaign County (Illinois) that incorporated deflection measurements obtained with the RWD. A treatment matrix, used as a decision-making tool, determines the recommended treatment method based on the Pavement Condition Index, traffic volume, and RWD deflection data. Based on this study, the authors concluded that RWD is an important component of the proposed PMS and in any pavement preservation programs in order to avoid applying preventive maintenance treatments on pavements that are not structurally sound.

A study was conducted in Virginia to evaluate the effectiveness of RWD as a pavement structural assessment tool (Diefenderfer 2010). Deflection testing was conducted at three sites in order to evaluate the influence of pavement types, temperature, and surface texture on the repeatability of the measurements and their correlation to FWD. Results of this study indicated that while the range of RWD and FWD deflection measurements was similar, the results of both testing methods did not correlate well. In addition, RWD deflection measurements were not repeatable on all test sections. However, it is noted that FWD and RWD testing measurements were not conducted concurrently and were phased out by a five-month testing gap.

**Field Testing Program**

The complete field testing program performed by LADOTD consisted of two phases. In the first phase, the complete asphalt road network (about 2250 km) in District 5 was tested using the RWD deflection system based on ARA standard testing protocol. In the second phase, 16 road-sections (each 2.4 km), referred to as research sites, were selected and were used for a detailed evaluation of RWD technology. The test plan consisted of conducting RWD and FWD measurements on the selected flexible and surface treatments pavement test sites. However, only the RWD measurements for the research sites are presented in this paper. The field testing program for RWD and FWD testing was conducted successfully with no major problems during the course of the experiment.

Selected test sites were representative of the pavement network in Louisiana in terms of pavement classification and design. All sections were asphalt-surfaced since the use of RWD on concrete pavement surfaces has not been validated. However, two composite sections with a known concrete layer underneath the asphalt surface were included to evaluate RWD's use for such pavement types. To assess the effects of truck speed on the measured deflection, RWD testing was conducted on the test sections at different speeds (i.e., 32, 48, 64, 80, and 97 km/h). However, test speed was restricted by the posted speed limits on a number of sites. Pavement temperature was recorded in conjunction with each test. To assist in the analysis, pavement

design of the selected sites was obtained using cores and construction documents. In addition, the test plan included supportive measurements such as roughness, pavement temperature, and distress survey for the selected sites.

Table 1 presents a description of the research sites. It is noted that only Site 7 was selected on the interstate highway system and allowed for testing at 96 km/h. However, testing at 80 km/h was conducted on eight of the 16 sites. Road segments were selected to represent different pavement conditions as described by the CPI and with varying Hot-Mix Asphalt (HMA) thickness and base type. As shown in Table 1, traffic volume widely varied in the selected sections from an Annual Average Daily Traffic (AADT) of 244 to 29,350; these traffic volumes range from low to heavy. Testing was conducted in December 2009. The pavement surface temperature ranged from -1.5 to 21°C with an average temperature of 9°C during the testing process.

**Table 1. Description of the Research Sites**

| Site ID | Pavement Type | Subgrade/Base Type | AADT | CPI (2009) | Condition-CPI | IRI (2009) |
|---|---|---|---|---|---|---|
| **1** | HMA | Granular | 1,185 | 81 | good | 145 |
| **2** | HMA | Granular | 2,585 | 92 | good | 89 |
| **3** | HMA | Stabilized | 6,398 | 99 | very good | 55 |
| **4** | HMA | Stabilized | 244 | N/A[1] | very good | N/A |
| **5** | HMA | Stabilized | 3,490 | 98 | very good | 64 |
| **6** | HMA | Stabilized | 5,623 | 99 | very good | 61 |
| **7** | Composite | Granular | 29,357 | 97 | very good | 71 |
| **8** | HMA | Granular | 6,409 | 70 | fair | 149 |
| **9** | HMA | Stabilized | 3,859 | 99 | very good | 57 |
| **10** | HMA | Granular | 650 | 64 | poor | 179 |
| **11** | HMA | Stabilized | 800 | 57 | poor | 309 |
| **12** | HMA | Stabilized | 1,898 | 77 | fair | 192 |
| **13** | HMA | Stabilized | 7,017 | 87 | good | 99 |
| **14** | Composite | Granular | 2,963 | 63 | poor | 202 |
| **15** | S. Treated | Stabilized | 424 | 82 | good | 202 |
| **16** | S. Treated | Stabilized | 244 | 60 | poor | 257 |

[1]: not available - newly-constructed in 2009.

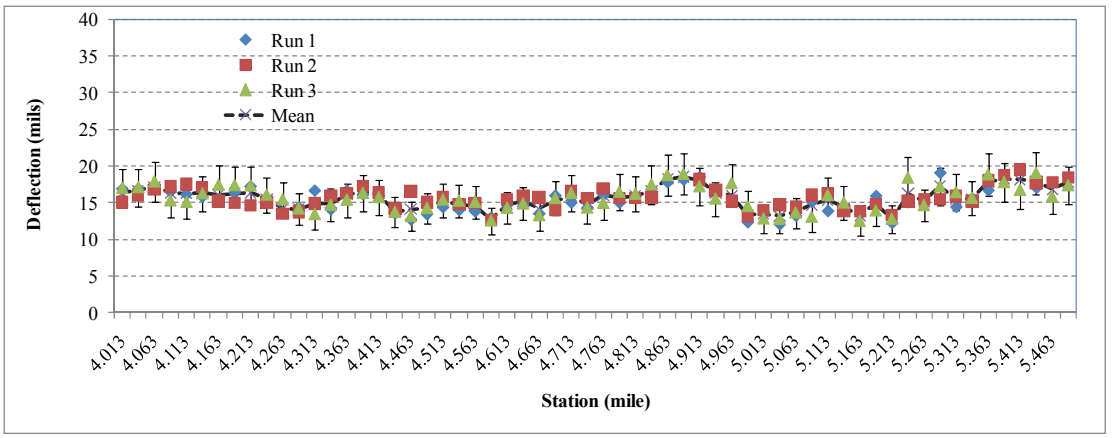
## Data Processing

During RWD testing, laser deflection readings are measured at 15-mm intervals. Irrelevant data such as measurements collected on top of bridge, sharp horizontal and vertical curves, and at traffic signals were removed. Erroneous data may also be obtained if the pavement surface is wet or in areas with severe cracking at the

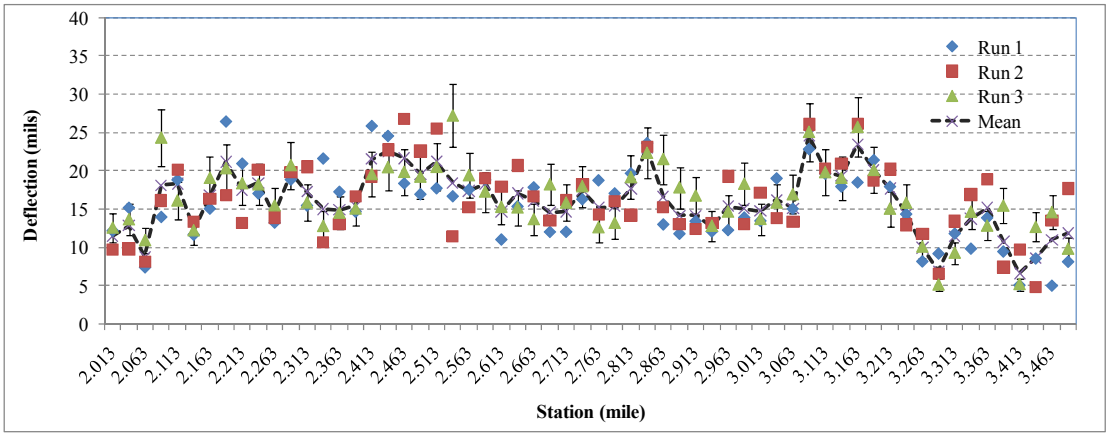
pavement surface. Valid deflection measurements were then averaged every 0.16 km (average of 10,728 individual readings) to provide a representative deflection measurement for each test interval (Vavrik et al. 2008). The averaging process reduces the error in the individual measurements caused by bouncing and random vibrations of the truck to within ± 1% of the global mean (Steele and Vavrik 2010). Surface deflections were corrected for variation in pavement temperature by shifting the measurements to a standard temperature of 20°C using the AASHTO 1993 method.

## Results and Analysis

Figure 2 (a and b) illustrates the measured pavement surface deflections for Sites 4 and 8. In these figures, the individual runs are presented along with the average of the measurements. As shown in these figures, measurements were more scattered in Site 8, which was reported to be in fair condition, than in Site 4, which was newly reconstructed in 2009. The uniformity of the measurements through the length of the test section is also evident in Figure 2(a).




(a)




(b)

**Figure 2.** Variation of Pavement Deflection for (a) Site 4 and (b) Site 8

The repeatability of RWD measurements is presented in Figure 3 as described by the coefficient of variation (COV) for the 16 research sites at a testing speed of 32 km/h. As shown in this figure, the repeatability of the measurements was acceptable with a COV ranging from 6 to 20% with an average of 14%. By comparing the conditions of the test sites presented in Table 1 to the deflection variability reported in Figure 3, one may observe that the majority of the sites in good conditions had better uniformity and less variability than sites in poor conditions. However, Site 9 had high deflection variability while being in very good conditions.

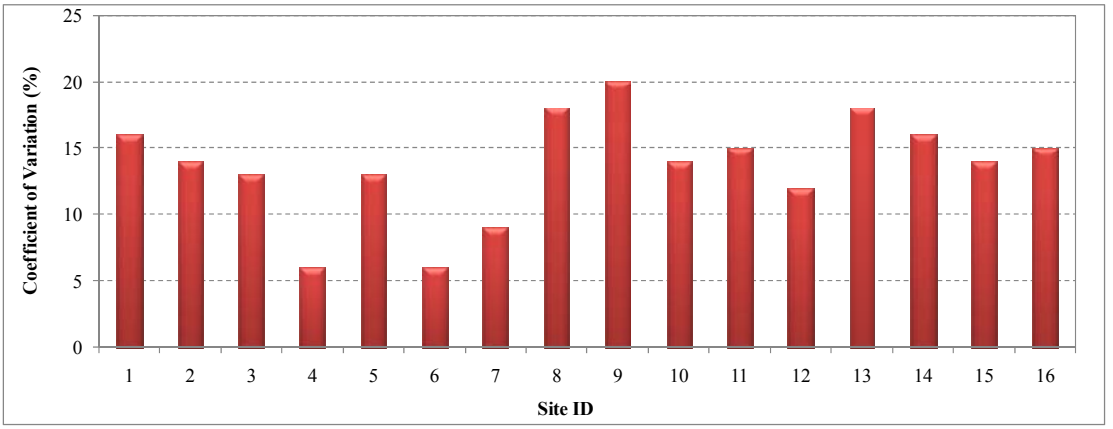



**Figure 3.** Repeatability of RWD Measurements for the Research Sites

### Effect of Speed on RWD Deflection Measurements

To assess the effects of truck speed on the measured deflection, RWD testing was conducted on the research sites at different speeds (i.e., 32, 48, 64, 80, and 96 km/h). Figure 4 illustrates the variation of the measured deflections with truck speed. As shown in this figure, the influence of the testing speed on the measured deflection was minimal and was not statistically significant. Since the test speed is restricted by the posted speed limit and the safe operation of a tractor and a semitrailer, this would allow comparing pavement surface deflections measured at different speeds and road conditions.

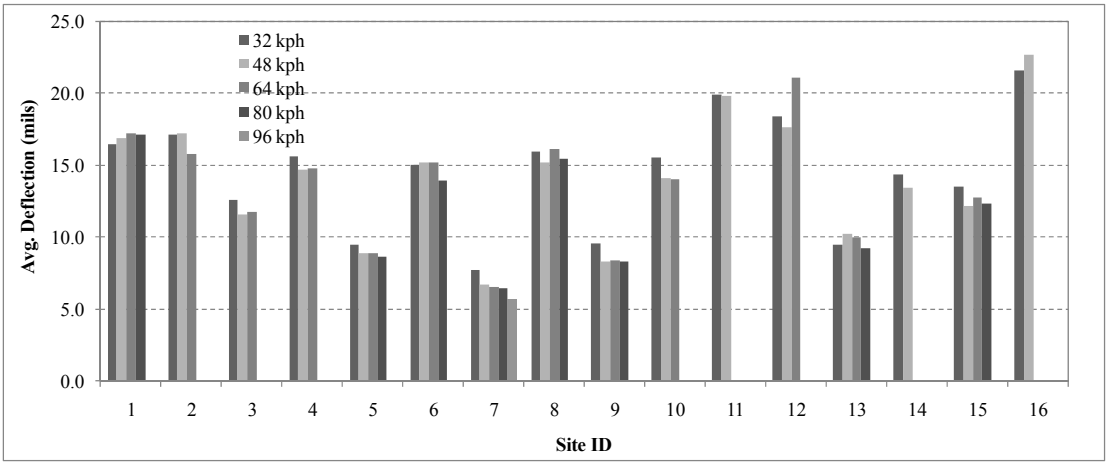



**Figure 4.** Average RWD Deflection Measured at Different Speeds

## Incorporation of RWD Measurements into a Deflection-Based GIS Map

RWD measurements were analyzed to develop a deflection-based GIS map displaying the variation of the average RWD deflection for the different test sections. Deflection-based GIS maps may be used to identify homogeneous sections, areas of distresses on selected sites, and to display the response of RWD to different pavement conditions. Figure 5 presents a GIS-based map showing the average RWD deflection for each site and its relationship to pavement conditions as defined by the CPI. As shown in this figure, the average RWD deflection appears to properly reflect the structural integrity and pavement conditions on each site by showing a greater average deflection for sites in poor conditions. However, the measured deflection is also a function of the pavement structural capacity. Therefore, RWD temperature-corrected deflections should be interpreted in conjunction with the pavement structural capacity and integrity.

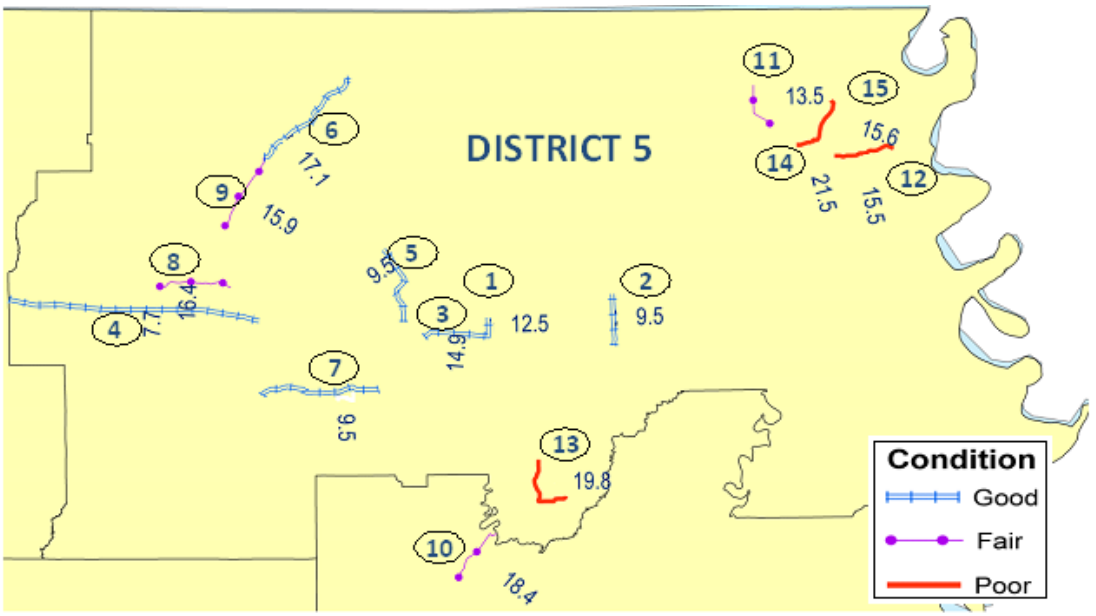



**Figure 5.** Deflection-Based GIS Map for the Research Sites

## Summary and Conclusions

The objective of this study was to evaluate the repeatability of RWD deflections and to determine the effects of test speed and pavement conditions on the measured deflections. The collected data were also incorporated into a deflection-based GIS map that was developed to demonstrate the use of this technology at the network level. The field testing program for RWD was conducted successfully with no major problems during the course of the experiment. Success of testing operations requires avoiding testing during rain or on wet pavements and to average individual readings to eliminate random errors due to truck bouncing and vibrations. Based on the results of the field testing program, the following conclusions may be drawn:

- Repeatability of RWD measurements was acceptable with an average coefficient of variation of 14%.
- The influence of the testing speed on the measured deflections was minimal. Since the test speed is restricted by the posted speed limit, testing can be conducted at different speeds while allowing for direct comparison of the measured deflections.
- RWD deflection measurements appear to properly reflect pavement conditions and structural integrity of the road network by providing for a greater average deflection and scattering for sites in poor conditions.

This study represents a first step towards evaluating and supporting the implementation of RWD technology in structural assessment of road conditions at the network level. Based on the results presented in this study, further research is underway to establish the relationship between RWD and FWD and to develop a methodology to utilize RWD measurements in assessing the structural integrity of

pavements at the network level. Research is also underway to quantify the cost-effectiveness of this technology as compared to standard FWD deflection testing.

## Acknowledgments

The financial support provided by the Louisiana Transportation Research Center (LTRC) is greatly appreciated. The contents of this paper do not necessarily reflect the official views or policies of the Louisiana Department of Transportation and Development or the Louisiana Transportation Research Center. The authors would like to acknowledge the assistance of R. Bandaru of Louisiana State University and D. Steele of ARA, Inc.

## References

Briggs, R.C., Johnson, R.F., Stubstad, R.N., and Pierce, L. (2000). "A comparison of the rolling weight deflectometer with the falling weight deflectometer." Nondestructive testing of pavements and backcalculation of moduli, third volume, ASTM STP 1375, S.D. Tayabji and E.O. Lukanen, Eds., American Society for Testing and Materials, West Conshohocken, PA, 2000, 444-456.

Diefenderfer, B.K. (2010). Investigation of the rolling wheel deflectometer as a network-level pavement structural evaluation tool. Virginia Transportation Research Council, Final Report VTRC 10-R5.

Gedafa, D.S., Hossain, M., Miller, R., and Steele, D. (2008). "Network level pavement structural evaluation using rolling wheel deflectometer." Paper No. 08-2648 Presented at the 87th Transportation Research Board Annual Meeting, Washington, D.C.

Gedafa, D.S. Hossain, M., Miller, R., and Van, T. (2010). "Estimation of remaining service life of flexible pavement from surface deflections." *Journal of Transportation Engineering*, ASCE, Vol. 136, No. 4, 342-352.

Groog, M.G., and Hall, J.W. (2004). "Measuring Pavement Deflection at 55 mph." *Public Roads*, Vol. 67, No. 4.

Khattak, M.J., Baladi, G.Y., Zhang, Z., and Ismail, S. (2008). "A review of the pavement management system of state of Louisiana – Phase I." *Transportation Research Record: Journal of the Transportation Research Board*, 2084, Washington, D.C.

Vavrik, W.R., Blue, J., and Steele, D.A. (2008). "Rolling wheel deflectometer-based pavement management system success: Champaign County, IL." Paper No. 08-2728 Presented at the 87th Transportation Research Board Annual Meeting, Washington, D.C.

# Evaluation of Recycled Asphalt Shingles in Hot Mix Asphalt

Shu Wei Goh[1] and Zhanping You[2]

[1]Ph.D. Student and Research Assistant, Department of Civil and Environmental Engineering, Michigan Technological University, Houghton, Michigan 49931, USA. Tel: (906)487-2528, Fax: (906)487-1620, Email: sgoh@mtu.edu
[2]Associate Professor, Department of Civil and Environmental Engineering, Michigan Technological University, Houghton, Michigan 49931, USA. Tel: (906)487-1059, Fax: (906)487-1620, Email: zyou@mtu.edu

## ABSTRACT

The objective of this study was to evaluate the performance of asphalt mixtures containing different dosages of recycled asphalt shingles (RAS) using the dynamic modulus (|E*|) test with a universal testing machine and rutting test with the asphalt pavement analyzer (APA). In this study, 5% and 10% RAS were used in the Superpave mixtures. A mixture without RAS was used as the control mixture. All the mixture samples were compacted at 86 gyrations under different temperatures. The 10% RAS had the highest air void level and the control mixture had the lowest air void level. The dynamic modulus and rutting tests were conducted to evaluate the performance of the mixtures with RAS. The results show that the |E*| for RAS mixtures were higher as compared to the control mixture when tested at high temperature (i.e. 39.2°C). Additionally, the added RAS decreased the rut depth significantly after 8000 cycles using the APA.

Keywords: Recycled Asphalt Shingles, Dynamic Modulus, Asphalt Mixes, Asphalt Pavement Analyzer, Flexible Pavements

## BACKGROUND

Every year, approximately eleven million tons of asphalt roofing shingles are disposed of in landfills. These shingles come from two sources, scraps and rejects from the manufacturing of new shingles and the shingles removed during reconstruction of existing roofs, also known as tear-offs (California Integrated Waste Management Board 2008). One alternative to landfilling these shingles is to recycle them into hot mix asphalt (HMA) pavements.

There is strong evidence that asphalt roofing shingles can be recycled into HMA to reduce costs without causing detrimental effects on performance (1989; 1996; 1997). The composition of asphalt shingles is similar to that of HMA in that both contain asphalt binder, aggregates and mineral fines. In addition, shingles contain fibers that can serve as reinforcement of the HMA (Abdulshafi et al. 1997). The binder

used to make shingles is stiffer than the binder typically used in HMA; and the aggregates used to surface the shingles are of a higher quality than those used in HMA. Approximately 10 million tons of tear-off shingle waste and 1 million tons of manufacturer scrap are produced each year in the US. These shingles must be processed to reduce their size and also to remove contaminants from the tear-off shingles. All of the waste shingles could be disposed of if two percent recycled shingles were used in all HMA produced in the US, which is less than current applicable specifications allow. The cost savings of using recycled shingles in HMA has been estimated to range from $0.50 to $2.80 per ton depending on virgin material costs and whether tear-off or manufacturer scrap shingles are used (Foo et al. 1999). At the same time, there are concerns about the presence of asbestos and polycyclic aromatic hydrocarbons in recycling shingles and studies have found that these concerns are minimal. The addition of shingles to HMA has been found to improve the rut resistance of the mixture and make the HMA easier to compact during construction (Hanson et al. 1997). Although large amounts of shingles have been found to increase the moisture susceptibility of the mix, small amounts have been found to have an insignificant impact on the mix. There have been several field trials of HMA pavements containing shingles performed by different states. The field trials have generally proved that the addition of shingles to HMA has been successful as the performance of these sections has been similar to that of the control sections (Amirkhanian and Vaughan 2001; Janisch and Turgeon 1996).

## OBJECTIVE

The objective of this study is to evaluate the performance of asphalt mixtures containing different dosages of RAS using the dynamic modulus (|E*|) testing and rutting test from asphalt pavement analyzer (APA).

## RECYCLE ASPHALT SHINGLES (RAS)

In this study, the recycled asphalt shingles (RAS) were obtained from Grand Rapid Asphalt, a paving firm in the State of Michigan. The RAS received were collected and processed at Crutchall Resource Recycling, a resource recycler located in Grand Rapids, Lansing, Kalamazoo, Warren, and Flint. Tear off shingles - shingles that has been removed from existing building - were used in this study. One of the differences between manufacturer scrap shingles and tear off shingles is the fact that tear offs usually contain other construction debris while manufacturer scrap shingles do not. This debris can be plastic, fiberglass or asbestos, chunks of wood, nails, or other debris that might be found on the roof. Thus, shingles that are tear offs need to be thoroughly processed to remove these debris.

In order to thoroughly investigate the composition of RAS received, the asphalt binder in the shingles was extracted and recovered using the procedure described in the ASTM D2172 standard specification, which is the procedure typically used in recovering binders from asphalt mixtures. Figure 1(a) and (b) show the shingles before and after it was extracted.

(a) (b)

**Figure 1. Shingles (a) before and (b) after the process of extraction and recovery**

From Figure 1, it is observed that the remaining shingles contain a fair amount of debris, mainly chunks of wood, fiberglass and pieces of plastic. The gradation testing of the shingles in this study was performed by Grand Rapids Asphalt. A total of four replicates gradation tests, each test contain approximately 1000g RAS sample, were conducted and these results are shown in Table 1.

**Table 1. Gradation Test for Recycled Asphalt Shingles**

| Sieves size | % Passed | | | | AVERAGE |
|---|---|---|---|---|---|
| | Sample 1 | Sample 2 | Sample 3 | Sample 4 | |
| 3/4 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 1/2 | 98.5 | 99.3 | 98.4 | 99.1 | 98.8 |
| 3/8 | 93.9 | 96.9 | 95.0 | 98.6 | 96.1 |
| 4 | 81.4 | 89.9 | 88.2 | 95.5 | 88.8 |
| 8 | 77.6 | 87.0 | 83.4 | 93.1 | 85.3 |
| 16 | 65.5 | 71.7 | 70.7 | 74.7 | 70.7 |
| 30 | 47.8 | 50.2 | 48.0 | 51.1 | 49.3 |
| 50 | 37.8 | 41.7 | 39.9 | 43.3 | 40.7 |
| 100 | 31.7 | 35.3 | 34.0 | 35.4 | 34.1 |
| 200 | 25.4 | 29.1 | 27.8 | 27.7 | 27.5 |
| Asphalt Content | 30.26% | 36.25% | 39.21% | 44.23% | 37.49% |

Based on the test results, it was observed that the sample sizes varied; however, a consistent result is observed among the four replicates. It is observed that a lot of dust (aggregate passing 200 sieve size) was in the RAS collected. For the asphalt binder content, it was found that RAS contains high binder content, which is around 37.49% based on the sample weight. These asphalt binders were highly aged and they were stiffer compared to the neat binder.

## SAMPLE PREPARATION

In this paper, asphalt mixtures with 5% and 10% RAS (by weight) were studied. It was assumed that the additional 5% and 10% RAS would not affect the

binder grade significantly and hence it is notable that the binder grade of RAS mixture was not evaluated in this study. The control and RAS mixtures were batched, mixed and compacted in the lab. The control mixture was mixed at 165°C and compacted at 153°C; and all the RAS mixtures were mixed at 180°C and compacted at 165°C. Same gyration number, 86 gyrations, was applied for all the mixtures (control, 5% RAS and 10% RAS mixture) during the compaction process using the Superpave™ gyratory compactor. A total of three replicates were applied to each dynamic modulus and APA rutting tests in order to meet the Superpave™ criteria.

The gradations for both control and RAS mixtures are shown in Table 1. A total of six mixtures (three mixtures for dynamic modulus test and another three for APA rutting test) were produced for each control, 5% RAS and 10% RAS mixtures in this evaluation.

Based on Table 2, it is observed that when RAS was added, it increases the air void level. Even when higher compaction temperature was used for RAS mixture, the air void level for 5% and 10% RAS mixture increased by 0.61% and 3.20%, respectively. The reason for the increased air void is mainly because RAS is stiffer than the original asphalt and need additional compaction or a higher compaction temperature to achieve the desired air void level.

**Table 2. Gradation of Control RAS mixture**

| | Percent Passing | | |
|---|---|---|---|
| **Sieve Size** | **Control Mixture** | **5% RAS Mixture** | **10% RAS Mixture** |
| 19mm | 100.0% | 100.0% | 100.0% |
| 12.5mm | 100.0% | 99.9% | 99.9% |
| 9.5mm | 99.1% | 99.1% | 99.1% |
| 4.75mm | 75.0% | 75.0% | 75.0% |
| 2.36mm | 55.9% | 55.9% | 55.9% |
| 1.18mm | 41.3% | 41.3% | 41.3% |
| 0.80mm | 27.5% | 27.5% | 27.5% |
| 0.30mm | 14.5% | 14.5% | 14.5% |
| 0.1mm | 7.5% | 7.5% | 7.5% |
| 0.075mm | 5.5% | 5.5% | 5.5% |
| **Asphalt Content:** | 5.57% | 5.57% (3.97% of new binder ) | 5.57% (2.09% of new binder ) |
| **Average Air Void Level:** | 6.10% | 6.71% | 9.30% |

## DYNAMIC MODULUS TESTING

The dynamic modulus (|E*|) is the ratio of stress to strain under haversine (or sinusoidal) loading conditions and is used as one of the material characterization inputs in the MEPDG to model pavement performances. The |E*| test in this study was performed based on the AASHTO TP62-03 Standard (AASHTO:TP62 2003). Control, 5% and 10% RAS mixtures were evaluated and a total of three replicate samples were

used for each type of mixture in this study. Three different temperatures were used (i.e. -5°C, 21.3°C and 39.2°C) and the frequencies ranged from 0.1 Hz to 25 Hz in this test. The results of the |E*| test are shown from Figure 2 to Figure 4.

Based on Figure 2, it is observed that |E*| for control mixtures is significantly higher than both 5% and 10% RAS mixture at low testing temperature (i.e. -5°C). However, one interesting finding is that when the test temperature increased, the difference in |E*| between the control and RAS mixture decreased. Furthermore, it is observed that the |E*| of RAS mixture surpassed the control mixture when tested at 39.2°C. A bigger |E*| at a high temperature (i.e. 39.2°C) indicated that the asphalt mixture has higher rutting resistant; and a lower |E*| at low temperature (-5°C) indicated that the asphalt mixture has lower fatigue cracking potential (Witczak 2007; Witczak 2005). In this study, the RAS mixtures favored both fatigue cracking and rutting resistant of asphalt mixture compared to the control mixtures. It was suspected that the higher fatigue cracking resistant nature of RAS mixture was due to the fiberglass in RAS mixture (Abtahi et al.); and the higher rutting resistant was due to the aging of the asphalt shingles.

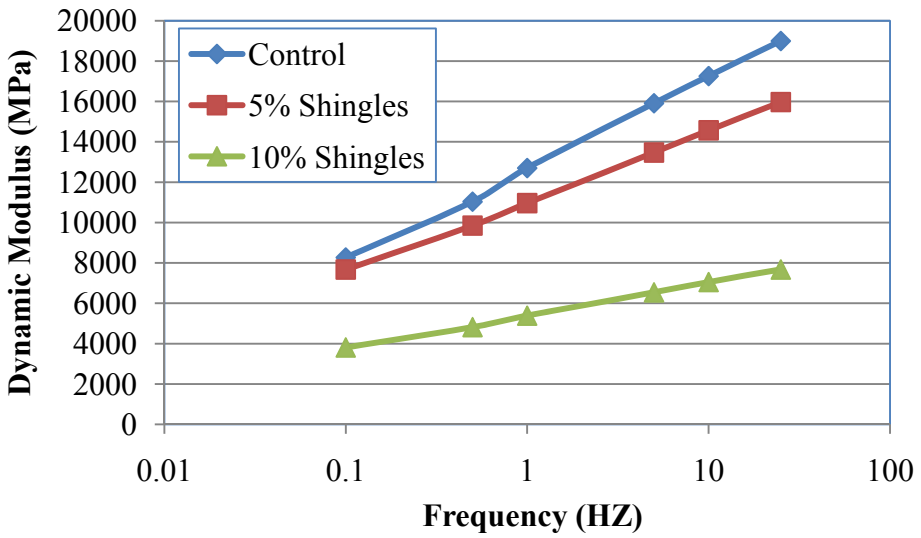



**Figure 2. Comparison of Dynamic Modulus Test Results for Control, 5% RAS and 10% RAS Mixtures at -5°C, Where the Average Air Void Content of the Control Samples is 6.10%, 5% RAS is 6.71%, 10% RAS is 9.30%)**

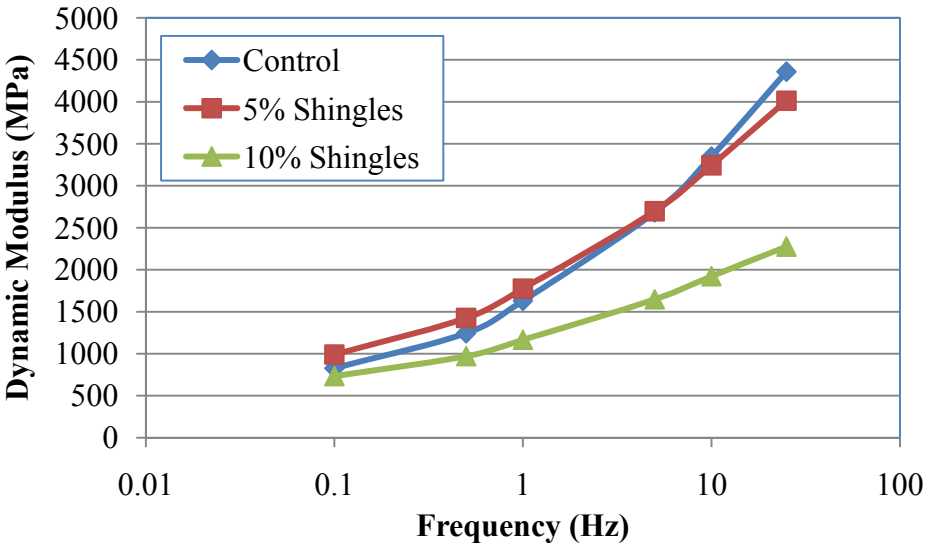



**Figure 3. Comparison of Dynamic Modulus Test Results for Control, 5% RAS and 10% RAS Mixtures at 21.3°C, Where the Average Air Void Content of the Control Samples is 6.10%, 5% RAS is 6.71%, 10% RAS is 9.30%)**

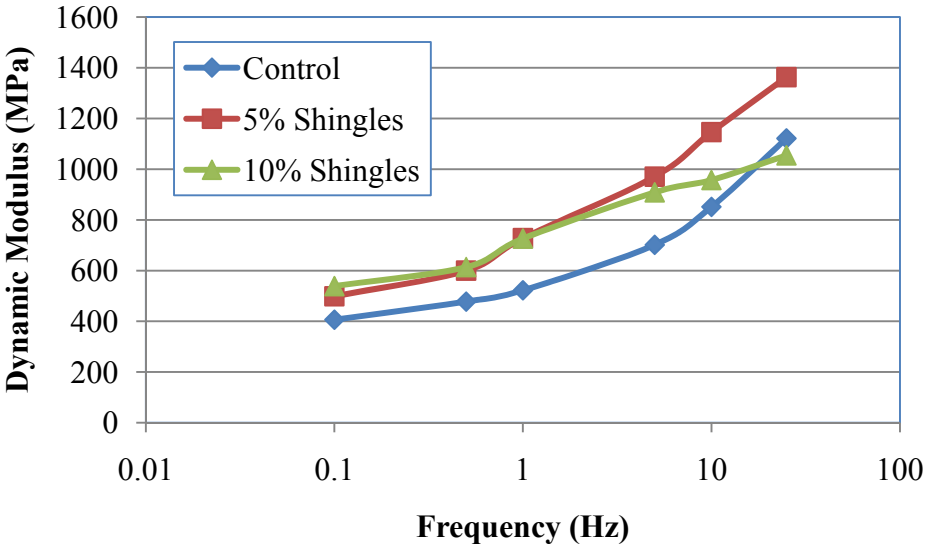



**Figure 4. Comparison of Dynamic Modulus Test Results for Control, 5% RAS and 10% RAS Mixtures at 39.2°C, Where the Average Air Void Content of the Control Samples is 6.10%, 5% RAS is 6.71%, 10% RAS is 9.30%)**

## APA RUTTING TEST

The rutting tests were conducted using the Asphalt Pavement Analyzer (APA) device based on the AASHTO TP63-03 standard specification (AASHTO:TP63 2003) at 58°C. The purpose of this test was to determine the rut resistance for the RAS

mixture and compare the results with the control mixture. In this test, the 5 and 10% RAS mixtures were tested and compared with the control mixtures. The results of the APA test are presented in Figure 5.

Based on the test conducted, it was found that the RAS mixture has a significantly lower rutting depth compared to the control mixture, especially for 10% RAS mixture. The final depth after 8000 cycles for the control and 10% RAS mixture are 3.48mm and 0.79mm, respectively. It was found that for the 5 and 10% RAS, the asphalt rutting after 8000 test cycles had decreased by 72% and 94%, respectively. This indicated that the use of RAS could significantly improve the rutting resistant of the asphalt pavement. This is most likely due to the highly aged nature of the RAS binder. The results were in line with the findings from |E*| testing which revealed a higher observed |E*| for the RAS mixture tested at a higher temperature of 39.2°C as shown in Figure 4.

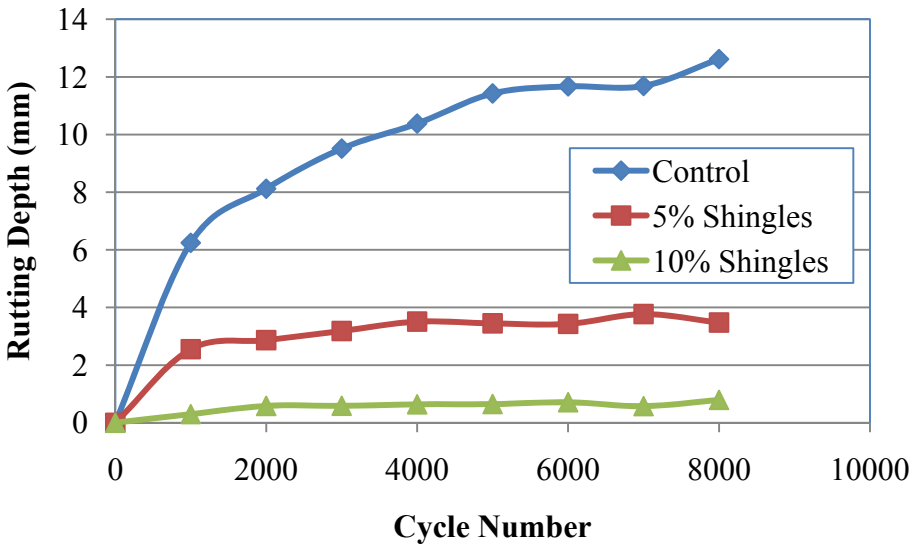



**Figure 5. Comparison of APA Rutting Test Results for Control, 5% RAS and 10% RAS Mixtures at 58°C**

## SUMMARY

This paper discusses the composition of recycled asphalt shingles (RAS) and the effect of RAS on HMA mixture. The composition of RAS was investigated by extracting and recovering the asphalt binder from the RAS. It was found that RAS contains a fair amount of debris, mainly chunks of wood and pieces of plastic. In addition, it was found that RAS contains as high as 37.49% of asphalt binder based on its weight.

The performance of 5 and 10% RAS mixture were evaluated using the dynamic modulus (|E*|) test and rutting test from asphalt pavement analyzer (APA). Based on |E*| testing, it was found that the RAS mixtures improved the rutting resistance of asphalt mixture compared to the control mixtures. Additionally, it was

found that the higher rutting resistant nature of the RAS mixture was due to the highly aged asphalt in the RAS.

For the APA testing, similar findings from |E*| was observed. It was found that when 5 and 10% RAS was used, the asphalt rutting decreased by 72% and 94%, respectively, after 8000 test cycles. This indicated that the use of RAS could significantly improve the rutting resistant of the asphalt pavement.

**ACKNOWLEDGEMENT**

The authors wish to express their gratitude to James Vivian, Moua Lee, Brad Diener, and David Porter at Michigan Technological University for their help in running some preliminary tests.

**REFERENCES**

AASHTO:TP62. (2003). "Standard Method of Test for Determining Dynamic Modulus of Hot-Mix Asphalt Concrete Mixtures." American Association of State Highway and Transportation Officials.

AASHTO:TP63. (2003). "Determining Rutting Susceptibility of Asphalt Paving Mixtures Using the Asphalt Pavement Analyzer (APA)." American Association of State Highway and Transportation Officials.

Abdulshafi, O., Kedzierski, B., Fitch, M., and Muhktar, H. (1997). "Evaluation of the Benefits of adding Waste Fiberglass Roofing Shingles to Hot-Mix-Asphalt." CTL Engineering, Incorporated, Ohio Department of Transportation, Federal Highway Administration.

Abtahi, Sayyed Mahdi, Sheikhzadeh, Mohammad, and Hejazi, Sayyed Mahdi. "Fiber-reinforced asphalt-concrete - A review." *Construction and Building Materials*, 24(6), 871-877.

Amirkhanian, S. N., and Vaughan, K. M. (2001). *Field Evaluation of Use of Waste Shingles in Asphalt Mixtures*, Clemson University, South Carolina Department of Transportation, Federal Highway Administration.

California Integrated Waste Management Board (2008). "Asphalt Roofing Shingles Recycling." Reference [cited 2008 April]; Available from: http://www.ciwmb.ca.gov/condemo/Shingles/.

Foo, K. Y., Hanson, D. I., and Lynn, T. A. (1999). "Evaluation of roofing shingles in hot mix asphalt." *Journal of Materials in Civil Engineering*, 11(1), 15-20.

Hanson, Doughlas I., Foo, Kee Y., and Lynn, Todd A. (1997). "Evalution of Roofing Shingles in HMA." National Cetner for Asphatl Technology.

Janisch, D. W., and Turgeon, C. M. (1996). *Minnesota's Experience with scrap Shingles in Bituminuous Pavements*, Minnesota Department of Transportation.

Witczak, M.W. (2007). "Specification Criteria for Simple Performance Tests for Rutting." *NCHRP Report 580*, Arizona State University, Tempe, Arizona.

Witczak, M. W. (2005). *Simple Performance Tests: Summary of Recommended Methods and Database*, Transportation Research Board.

# Structural Coefficients and Life-Cycle Cost

Kendra P. Davis[1] and David H. Timm[2]

[1]Technical Director, Georgia Asphalt Pavement Association, 2160 Satellite Boulevard Suite 100, Duluth, GA 30097; PH (770) 710-2158; e-mail: kdavis@gahotmix.com
[2]Gottlieb Associate Professor of Civil Engineering, Auburn University, 238 Harbert Engineering Center, Auburn, AL 36849; PH (334) 844-6282; e-mail: timmdav@auburn.edu

## ABSTRACT

As transportation agencies search for relief from highly strained budgets, a recent study from National Center for Asphalt Technology (NCAT) shows that hot mix asphalt (HMA) pavements can be designed approximately 18.5% thinner using the 1993 AASHTO Design Guide for flexible pavements. This savings is the result of recalibrating the asphalt layer coefficient. The previously recommended layer coefficient, used by the Alabama Department of Transportation (ALDOT), of 0.44 comes from the AASHO Road Test, which took place near Ottawa, Illinois from 1958 to 1960. The recalibrated average layer coefficient of 0.54 was determined by utilizing performance and traffic data from test sections at the NCAT Test Track, and reflects advancements made in HMA materials and construction over the last 50 years. While the direct savings of 18.5% HMA thickness is obvious, the savings in initial construction and life cycle costs are not as straightforward. However, there will always be a savings in the initial cost and user delay costs because thinner pavements are cheaper to build and take less time to construct. The economic benefits of changing to the new layer coefficient will vary depending upon what layer coefficient an agency is currently using. A survey of state transportation agencies found that there is little consistency in the layer coefficients used across the nation. Many states use different coefficients for varying types of mix or use a different coefficient beyond a certain depth of HMA, while others use the same value regardless of mix type or depth. This inconsistency between the states does not seem to follow a particular trend (e.g., higher values in the southern portion of the country). The wide array of layer coefficients used across the nation suggests that this value needed to be revisited, and has lead many states toward implementing mechanistic-empirical design systems. This paper describes the recalibration efforts at the NCAT Test Track followed by a discussion of life-cycle costs related to using a new structural coefficient. Data from the survey of state agencies are also presented relative to the new NCAT findings and life-cycle cost analysis.

## INTRODUCTION

While there is a noticeable shift in interest toward mechanistic-empirical (M-E) pavement design across the nation, the actual implementation of the upcoming DARWin ME 2.0 software will likely not take place for several more years for many transportation agencies. This issue, combined with the current economic downturn,

has created a need to revisit the current pavement design methodology to ensure that hot mix asphalt (HMA) pavements are being constructed at optimal thickness. The current design system used by many state agencies is the 1993 American Association of State Highway and Transportation Officials (AASHTO) Design Guide, which was developed based on the results of the AASHO Road Test. This road test took place from 1958 to 1960 near Ottawa, Illinois, and was conducted to relate axle loadings to required pavement thickness. There are many limitations associated with using the road test results for pavement design, including the application of only 1.1 million equivalent single axle loads (ESALs), the use of only one soil type for subgrade, and using a maximum HMA thickness of 6 inches. Additionally, stone matrix asphalt (SMA), polymer-modified mixes, and Superpave mixtures did not exist during the test; therefore, current design procedures do not reflect the advancements that have been made in paving materials and design over the last 50 years.

Using the AASHO Road Test findings, equations were developed and have been modified over time to determine the necessary HMA thickness for a given set of parameters (traffic, subgrade resilient modulus, etc.). One parameter needed to calculate HMA thickness is the layer coefficient, which is a term used to quantify the relative structural contribution of a pavement layer. For HMA, AASHTO recommends using a layer coefficient of 0.44 (AASHTO 1993). This value was taken as a weighted average of values obtained from the AASHO Road Test, which ranged from 0.33 to 0.83 (HRB 1962). This range of values is surprisingly broad, and the AASHTO Design Guide recommends that each state determine coefficients individually if possible. However, the layer coefficient is a complex parameter that is dependent upon many factors, including material stiffness, compressive strength, tensile strength, moisture conditions, and the layer's relative position within the pavement cross section (AASHTO 1993; Pologruto 2001). A literature review revealed that a synthesis of nationwide HMA layer coefficients did not exist.

A sensitivity analysis of the inputs to the AASHTO flexible pavement design equation showed that the HMA layer coefficient is the most influential input on the resulting HMA thickness, more so than the traffic level, subgrade resilient modulus, change in serviceability, reliability or variability (Davis and Timm 2009). Therefore, recalibrating this value would provide the greatest potential benefit, and an optimal pavement cross section. While many researchers have recalibrated this value in the past (Van Wyk et al. 1983; Corree and White 1989; Hossain et al. 1997; Romanoschi and Metcalf 1999; Pologruto 2001; Jess and Timm 2005; von Quintus 2007), there have been very few changes implemented as a result of their studies. For the investigations that included dense-graded HMA mixtures, the resulting layer coefficients ranged from 0.44 to 0.64, indicating a possible need to increase the AASHTO-recommended HMA layer coefficient. The authors believe these new values have not been implemented because they were calibrated primarily using deflection measurements rather than pavement performance data.

While increasing the HMA layer coefficient would have obvious direct benefits to the initial construction costs of a pavement, the change in life-cycle costs is not as readily apparent. Life cycle cost analysis (LCCA) is a valuable tool used to compare the costs of pavement alternatives beyond the initial construction costs. The costs are compared over an analysis period, which is a timeframe that includes at least

one rehabilitation cycle of each alternative. All future costs of rehabilitation are discounted back to the present to compare the net present value amounts for each alternative, and sometimes social "costs" are included, such as user delay costs due to the construction work zone. If these costs are considered, increasing the layer coefficient would have even more measureable benefits because thinner pavements take less time to construct, and therefore would cause a further decrease in the life cycle costs of a HMA pavement alternative.

## OBJECTIVE

The objectives of this work were to briefly describe the process used to recalibrate the HMA layer coefficient, demonstrate the effect the new layer coefficient has on the life cycle cost of HMA pavements, and present data from a nationwide survey on layer coefficients.

## SCOPE OF WORK

A nationwide survey was conducted to determine the layer coefficients used for new HMA construction in each state to better quantify the potential savings available to the other states. The recalibration of the HMA layer coefficient was conducted using traffic and performance data from the 2003 and 2006 NCAT Test Track research cycles. Only the structural sections from each research cycle were used in the study, for a total of 14 sections representing cross sections, materials, and mixes from various states. The test sections were exposed to an accelerated loading rate of 10 million ESALs per three-year cycle, and detailed traffic loadings and pavement performance data from each cycle provided an optimal scenario for recalibration by performing a least squares regression between measured and predicted pavement performance. The newly recommended layer coefficient provides an approximate 18.5% savings in HMA thickness (relative to the AASHTO recommended coefficient), and the effect this savings has on life cycle costs of new pavement and overlays was quantified using the LCCAExpress software.

## NATIONWIDE LAYER COEFFICIENTS

A literature review on layer coefficients revealed that a recent survey of states' values had not been conducted. Therefore, a survey was sent out to state transportation agencies and state asphalt pavement associations to determine the values being used for HMA layer coefficients, if applicable. The survey was only for dense-graded, newly constructed HMA mixtures. If no response was received, an online search of state specifications was used to obtain the pavement design method and/or layer coefficient value. The results from the survey are shown in Table 1.

**Table 1.** Results from nationwide HMA layer coefficient survey

| State | HMA Layer Coefficient |
|---|---|
| Alabama | 0.30 base mixes, 0.54 surface & intermediate mixes (revised from 0.44 in 2009) |
| Alaska | Not applicable, mechanistic design |
| Arizona | 0.44 |
| Arkansas | 0.44 surface & intermediate mixes, 0.36 base mixes |

| State | HMA Layer Coefficient |
|---|---|
| California | Not applicable, gravel factor design |
| Colorado | 0.40 - 0.44 |
| Connecticut | 0.44 surface and intermediate, 0.34 base |
| Delaware | 0.44 surface and intermediate, 0.32 base |
| Florida | 0.44 |
| Georgia | 0.44 above 4.5”, 0.30 below 4.5” |
| Hawaii | Not applicable, gravel factor design |
| Idaho | Not applicable, gravel factor design |
| Illinois | Not applicable, mechanistic design |
| Indiana | Not applicable, mechanistic design |
| Iowa | 0.46 (crushed content > 60%), 0.44 surface, 0.40 intermediate and base |
| Kansas | 0.42 surface, 0.34 base |
| Kentucky | 0.44 surface, 0.42 intermediate, 0.40 base |
| Louisiana | 0.44 surface and intermediate, 0.33 base |
| Maine | 0.44 down to 4”, 0.34 below 4” |
| Maryland | 0.44 |
| Massachusetts | 0.44 surface, 0.34 base |
| Michigan | 0.42 surface and intermediate, 0.36 base |
| Minnesota | Not applicable, granular equivalency design |
| Mississippi | 0.44 |
| Missouri | Not applicable, mechanistic-empirical design |
| Montana | 0.41 |
| Nebraska | 0.44 |
| Nevada | 0.35 |
| New Hampshire | 0.38 surface and intermediate, 0.34 base |
| New Jersey | 0.44 |
| New Mexico | 0.44 |
| New York | 0.42 |
| North Carolina | 0.44 surface, 0.40 intermediate and base |
| North Dakota | 0.34 – 0.40 |
| Ohio | 0.43 surface and intermediate, 0.36 base |
| Oklahoma | 0.44 |
| Oregon | 0.42 |
| Pennsylvania | 0.44 surface, 0.40 base |
| Rhode Island | 0.44 |
| South Carolina | 0.44 surface, 0.34 intermediate and base and > 4” deep |
| South Dakota | 0.40 |
| Tennessee | 0.40 |
| Texas | Not applicable, mechanistic-empirical design |

| State | HMA Layer Coefficient |
|---|---|
| Utah | 0.40 |
| Vermont | 0.32 – 0.36 |
| Virginia | 0.44 surface, 0.40 intermediate and base |
| West Virginia | 0.44 |
| Washington | 0.44 |
| Wisconsin | 0.44 |
| Wyoming | 0.44 100 gyration mixes, 0.40 50 gyration mixes |

As seen in Table 1, there is a wide range of HMA layer coefficients (0.30 to 0.54) used for new HMA pavement design across the nation. Several states have switched to mechanistic or mechanistic-empirical design methodologies, and no longer use the AASHTO method. California, Hawaii, Idaho, and Minnesota use gravel factors or granular equivalency factors for pavement design, which equate the strength of other pavement layers to the strength of a granular base. Georgia, Maine, and South Carolina use varying coefficients based on depth in the pavement. Forty percent of the states use varying coefficients based upon the mix type, and the remainder of the states (38%) simply use one value regardless of depth or mix type. Twelve of those states use the AASHTO-recommended layer coefficient of 0.44. It is interesting to note that prior to 2009, no state used a value exceeding 0.44 (except for Iowa with high crushed aggregate content) despite the many advances in asphalt mixture technology mentioned previously. The potential consequence of this was less-than-optimal pavement sections as recognized by ALDOT and served as the motivation for recalibration.

## RECALIBRATION OF THE LAYER COEFFICIENT

A recalibrated layer coefficient was found for each of the 14 structural sections used in this investigation. The test sections from the 2003 and 2006 NCAT Test Track research cycles are shown in Figs. 1 and 2, respectively. Detailed traffic (axle passes, axle weights) and performance (IRI) data were used in the 1993 AASHTO Design Guide flexible pavement design equation and AASHTO traffic equations to calculate predicted and measured ESALs, respectively. The IRI data were collected for each wheelpath converted to serviceability (PSI) values using the Al-Omari/Darter equation for use in the AASHTO equations (Al-Omari and Darter 1994). To be conservative, the wheelpath with the lowest serviceability values was used for recalibration. Sections that did not exhibit significant decreases in serviceability were assigned artificial terminal serviceability levels to be conservative. The predicted and measured ESALs were compared, and a least squares regression was performed to minimize the differences between these values to determine the optimal layer coefficient for each section.

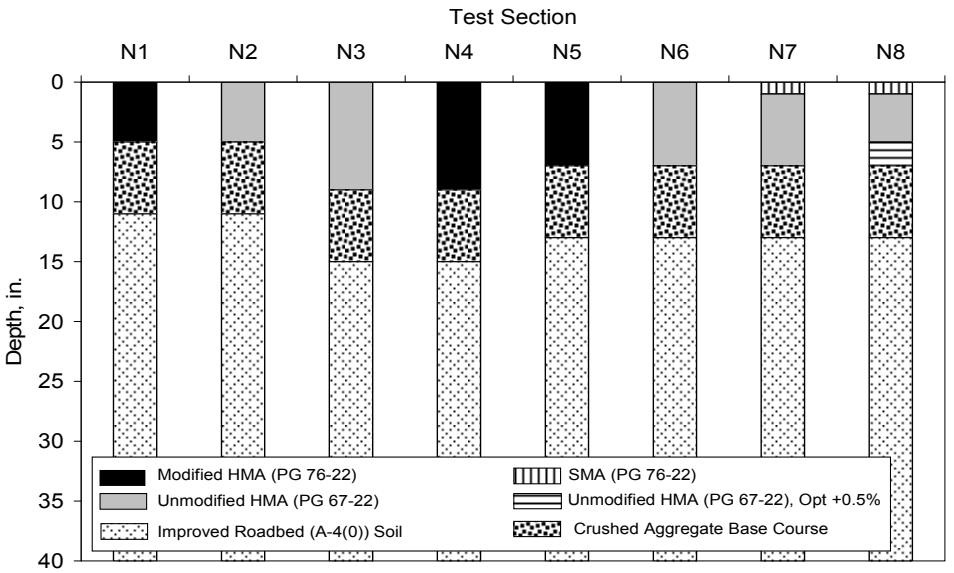



**Figure 1.** 2003 NCAT Test Track structural sections (Priest and Timm 2006)

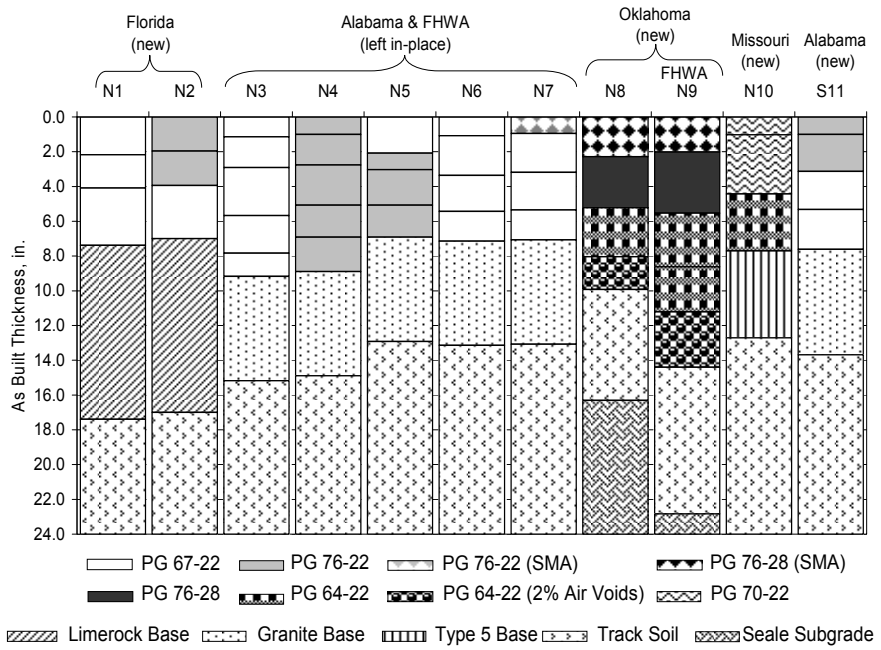



**Figure 2.** 2006 NCAT Test Track structural sections (Timm 2009)

Fig. 3 shows the resulting HMA layer coefficients by test section, as well as the overall average of 0.54. Only two sections had layer coefficients less than 0.44, and both had probable debonding and/or slippage failures (Willis and Timm 2006; Davis and Timm 2009). A full description of the layer coefficient recalibration study has been documented elsewhere (Davis and Timm 2009; Davis and Timm in press).

Using a layer coefficient of 0.54 instead of 0.44 results in an approximate 18.5% HMA thickness savings. As shown in Table 1, ALDOT implemented the new layer coefficient for all new pavements and overlay designs in 2009. It is important to note that the recalibration was done in the Southeastern United States on a variety

of asphalt mixtures. Care should be taken when transferring the results to other parts of the country with different mixtures. The results may not be directly applicable.

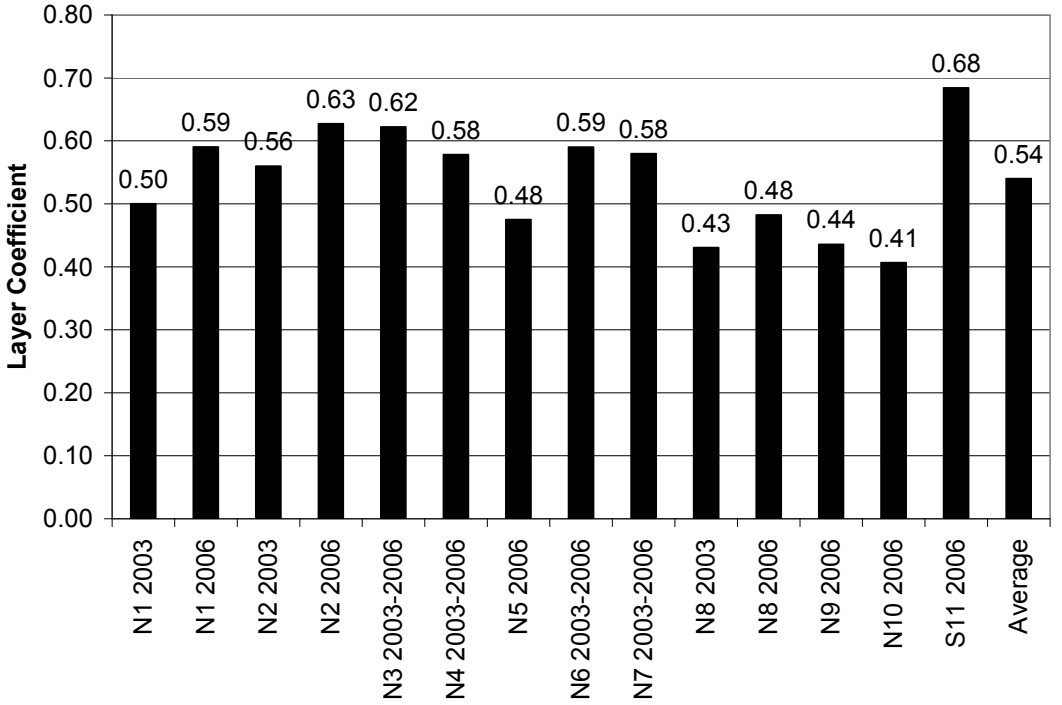



**Figure 3.** Computed layer coefficients for NCAT Test Track structural sections (Davis and Timm 2009)

## IMPACT ON LIFE CYCLE COSTS

Although the immediate 18.5% savings in thickness is an obvious direct benefit to overall HMA thickness, the life cycle cost savings are not as easy to quantify. LCCAExpress was used to compare the costs of two new pavement sections; one designed with the AASHTO recommended 0.44 layer coefficient, and the other with the new 0.54 layer coefficient. LCCAExpress follows Federal Highway Administration (FHWA) guidelines for calculating life cycle costs, and requires inputs such as pavement layer thicknesses, material costs, discount rate, traffic information, and a schedule of rehabilitation activities (Walls and Smith 1998). To perform a comparison of the life cycle costs using each layer coefficient, a hypothetical pavement design was created as shown in Table 2.

**Table 2.** Hypothetical scenario for new full-depth pavement construction

| **Input** | **Scenario 1: $a_1$ = 0.44** | **Scenario 2: $a_1$ = 0.54** |
|---|---|---|
| Project Length | 2 miles | |
| Number of Lanes | 2 in each direction | |
| Life Span | 40 years | |
| Discount Rate | 4% | |
| Current Traffic Volume | 20,000 vehicles/day | |
| Percent Trucks | 12% | |
| Traffic Growth Rate | 1% | |
| Rehabilitation Activities @ Cost | Patching: 5% @ \$10/yd$^2$<br>Milling: 2" @ \$5/yd$^2$<br>Inlay: 2" @ \$65/ton (1.5"), \$54/ton (0.5") | |
| HMA Thickness @ Cost | Wearing: 1.5" @ \$65/ton<br>Binder: 2" @ \$54/ton<br>Base: 7.5" @ \$52/ton | Wearing: 1.5" @ \$65/ton<br>Binder: 2" @ \$54/ton<br>Base: 5.5" @ \$52/ton |

The design shown in Table 2 is typical for what might be found on a heavily trafficked U.S. highway or state route. Scenario 1 is 11" thick while Scenario 2 is 9" thick. The two-inch reduction was applied to the base mixture resulting in an overall pavement thickness reduction of approximately 18.5%. The prices provided are roughly based on Georgia bid prices, and do not reflect current market conditions. For simplicity, a production rate of 200 tons/hour was assumed for the calculation of initial construction and rehabilitation durations. The rehabilitation activities, milling and inlay of 2 inches, were scheduled to occur at 15, 27, and 39 years. User delay costs were included in the analysis. These costs were calculated based upon the amount of time lost to stopping, time lost sitting in traffic, as well as the cost of idling. These costs, which come from the FHWA Bulletin (Walls and Smith 1998), were escalated to account for inflation. The costs vary based upon vehicle type and inflation rate, and are shown in Table 3.

**Table 3.** Delay cost rates used in LCCAExpress

| **Category** | **Passenger Cars** | **Single Unit Trucks** | **Combination Trucks** |
|---|---|---|---|
| Added Cost (\$/1000 stops) | 118.71 | 228.82 | 1,026.49 |
| Added Time (hrs/1000 stops) | 5.84 | 8.07 | 20.72 |
| Value of Time (\$/veh-hr) | 16.47 | 26.37 | 31.73 |

Fig. 4 shows the comparison of initial construction and total life-cycle costs for each scenario. Both the initial and life-cycle costs are less expensive for the section designed using 0.54, as expected. There is a 14% savings for initial construction costs, and a 10% cost savings over the 40-year life of the pavement. Additionally, there was a 14% savings in user delay costs (a decrease from \$2,134 to \$1,835) since the thinner pavement would take less time to build, and therefore mean less time spent sitting in traffic while work zones are in effect.

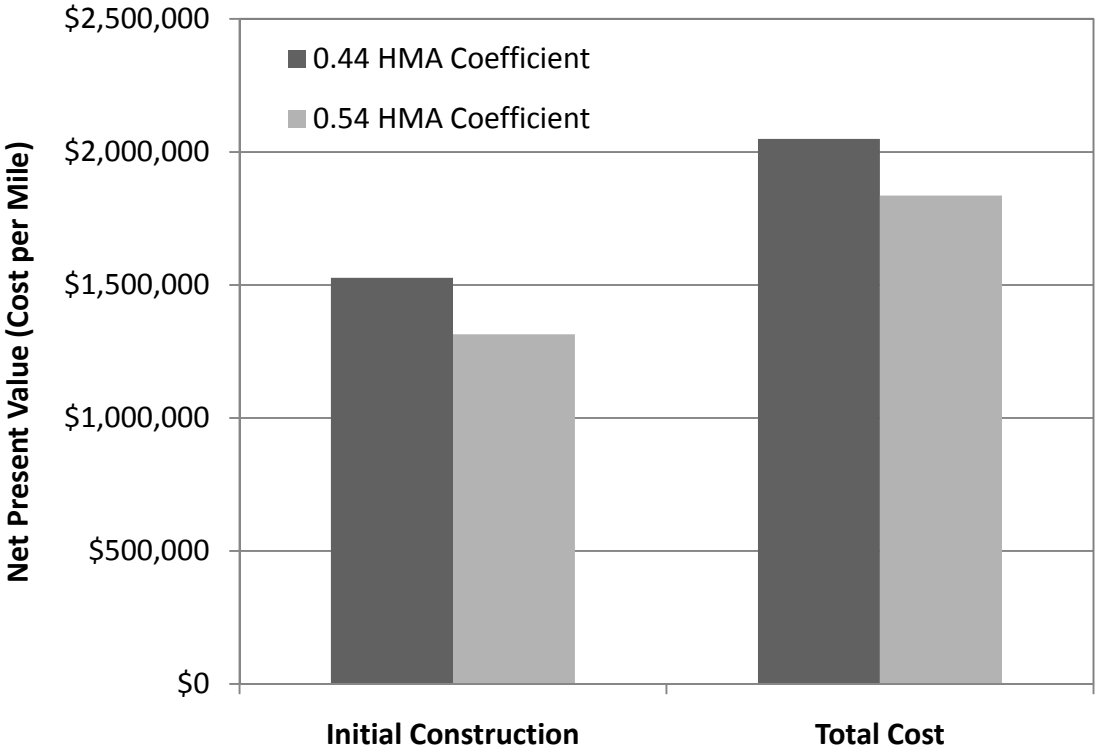



**Figure 4.** LCCA results for new pavement construction (Davis 2010)

## CONCLUSIONS

Based upon data obtained from a nationwide survey on HMA layer coefficients, there was very little consistency in the values used and they seemed to be capped at the AASHTO recommended value of 0.44. A sensitivity analysis showed that the HMA layer coefficient is the most influential parameter on resulting HMA thickness using the AASHTO Design Guide flexible pavement design equation, and therefore recalibrating this value would provide the greatest potential economical benefit. Fourteen structural sections from the 2003 and 2006 NCAT Test Track cycles were used to recalibrate the layer coefficient, and the resulting average layer coefficient was 0.54. Using this value instead of the AASHTO recommended coefficient of 0.44 results in an approximate 18.5% HMA thickness savings. This change has been implemented by ALDOT for all new HMA pavement and overlay designs. This thickness savings is obviously beneficial, but its effect on life-cycle costs is more complex. An example LCCA calculation showed that for a new pavement section, using the 0.54 coefficient saved approximately 10% over the entire life cycle, and the initial construction costs were 14% lower. User delay costs were also 14% less because the thinner pavement would take less time to build. Care should be taken when considering the use of this coefficient in other states because it is a product of the environment and materials used in the study.

## REFERENCES

AASHTO (1993). *AASHTO Guide for Design of Pavement Structures*. American

Association of State Highway and Transportation Officials, Washington, D.C.
Al-Omari, B. and Darter, M.I. (1994). "Relationships between international roughness index and present serviceability rating." *Transportation Research Record*, 1435, 130-136.
Corree, B.J. and White, T.D. (1989). "The synthesis of mixture strength parameters applied to the determination of AASHTO layer coefficient distributions." *Asphalt Paving Technology*, 58, 109-141.
Davis, K.D. (2010). "How the new structural coefficient impacts life cycle costs." *Hot Mix Asphalt Technology*, Vol. 15 No. 4, 34-35.
Davis, K.D. and Timm, D.H. (in press). "Recalibration of the asphalt layer coefficient." *ASCE Journal of Transportation Engineering*.
Davis, K.D. and Timm, D.H. (2009). "Recalibration of the asphalt layer coefficient." *NCAT Rep. No. 09-03*, Auburn, Ala.
Hossain, M., Habib, A., and LaTorella, T.M. (1997). "Structural layer coefficients of crumb-rubber modified asphalt concrete mixtures." *Transportation Research Record*, 1583, 62-70.
HRB (1962). *The AASHO Road Test.* Special Reports 61A, 61C, 61E. Highway Research Board.
Jess, J.C. and Timm, D.H. (2005). "Structural coefficients for new asphalt mixtures." *ALDOT Project No. 930-559*, National Center for Asphalt Technology, Auburn University, Auburn, Ala.
Pologruto, M. (2001). "Procedure for use of falling weight deflectometer to determine AASHTO layer coefficients." *Transportation Research Record*, 1764, 11-19.
Priest, A.L. and Timm, D.H. (2006). "Material Properties of the 2003 NCAT Test Track Structural Study." *NCAT Rep. No. 06-01*, Auburn, Ala.
Romanoschi, S. and Metcalf, J.B. (1999). "Simple approach to estimation of pavement structural capacity." *Transportation Research Record*, 1652, 198-205.
Timm, D.H. (2009). "Design, construction and instrumentation of the 2006 test track structural study." *NCAT Rep. No. 09-01*, Auburn, Ala.
Van Wyk, A., Yoder, E.J. and Wood, L.E. (1983). "Determination of structural equivalency factors of recycled layers by using field data." *Transportation Research Record*, 898, 122-132.
Von Quintus, H.L. (2007). "Evaluation of procedure to assign structural layer coefficients for use in flexible pavement design." *Kansas Department of Transportation Rep. No. KS-07-9*, Topeka, Kansas.
Walls, J. and Smith, M.R. (1998). "Life cycle cost analysis in pavement design – interim technical bulletin." *FHWA Report No. FHWA-SA-98-079*, Federal Highway Administration, Washington, D.C.
Willis, J.R. and Timm, D.H. (2006). "Forensic investigation of a rich bottom pavement." *National Center for Asphalt Technology (NCAT) Rep. No. 06-04,* Auburn, Ala.

# Impact of Pavement Surface Monitoring Frequency on Pavement Management Decision Making

Syed Waqar Haider[1], Gilbert Y. Baladi[2], and Karim Chatti[2]

[1]Assistant Professor, Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824; PH (517) 353-9782; FAX (517) 432-1827; email: syedwaqa@egr.msu.edu
[2]Professor, Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824.

## ABSTRACT

Highway agencies collect pavement condition data over time to manage their network in a cost-effective manner. However, there exist differences among highway agencies for data collection monitoring frequency. While most agencies collect sensor-based data (roughness, rut depth, and faulting) on a yearly basis, image-based data (cracking and other pavement defects) are collected every other year or on a three-year basis. The main reason for the differences is the relative difficulties (cost, resources, and technology) in collecting and processing imaging data. While monitoring frequency has an impact on performance prediction at project-level, it can also affect pavement management system (PMS) decisions at the network-level. Such decisions include identification of project boundaries, optimum timing of pavement fixes, and selection of the most cost-effective pavement preservation strategy. In this paper, PMS data from the state of Washington and MnRoad for flexible and rigid pavements were analyzed to determine the impact of monitoring frequency on the PMS decisions. Results of the analyses show that monitoring frequency may affect the identification of project boundaries, optimum fix type, and timing. These effects have implications on the cost effectiveness of the pavement management strategy.

## INTRODUCTION

Monitoring pavement surface condition over time is an essential part of pavement management and is required for the development and/or calibration of pavement performance models (Haider et al. 2010; McGhee 2004). The time series condition data could be used to accomplish the goals and objectives of State Highway Agencies (SHA). In particular, the data could be used to determine the needs for pavement preservation and rehabilitation. In general, the SHA's goals and objectives dictate the type and details of the pavement distress data to be collected, and the method and frequency of data collection. The frequency of pavement distress monitoring can be divided into three broad categories.

1. Short frequency (1 to 2 years); the data are typically used to develop or calibrate pavement performance models.
2. Long frequency (3 to 4 years); the data are typically used for assessing the general condition of the pavement network.
3. Localized monitoring; the data are typically used to address an unexpected pavement performance.

In practice, differences exist among agencies between the monitoring frequency used for pavement surface distress (imaging) and that used for sensor-measured features (roughness, rut depth, and joint faulting). These differences are mainly related to the relative difficulties (cost, resources, and technology) in collecting and processing imaging data. Many agencies collect sensor data more frequently than images. Most of the highway agencies monitor pavement condition at 1-, 2-, or 3-year frequencies (McGhee 2004). Thus, the frequency of distress data collection could further add to the uncertainty in performance prediction (Haider et al. 2010). The uncertainty associated to the method of condition data collection (manual, semi-automated, and automated) has been well documented in the literature (Albitres et al. 2007; Benson et al. 1988; FORMAT 2003; FORMAT 2005; McQueen and Timm 2005; Rada et al. 2004; Wang and Gong 2002). However, there is a need to investigate the effects of condition monitoring frequency on pavement management decisions related to highway maintenance, rehabilitation, and reconstruction. The monitoring data are used to establish maintenance and rehabilitation priorities, to select maintenance and rehabilitation strategies and to project pavement performance (Hicks and Mahoney 1981). Determining the effects of data collection frequency will assist highway agencies to bridge the information gap for pavement management purposes and their support for the evaluation and calibration of the performance models. Also, several pavement surface condition indicators are collected at the same frequency although the optimum frequency could be different from one indicator to another depending upon the progression with time. Consequently, a smallest monitoring interval should be adopted to minimize its impacts among various condition indicators.

In general, the project scoping processes for the selection of the boundaries of pavement projects and the type of pavement fix are based on one or combination of the following three methods:

1. The predicted pavement performance or the predicted pavement Remaining Service Life (RSL) using the observed or measured time dependent distress data. The advantage of RSL method is that the pavement rate of deterioration is included in the process. However, if the pavement performance and/or RSL predictions are uncertain or accurate time series data are unavailable, the PMS based decisions could be inaccurate and less effective.
2. The latest available distress data or the latest calculated distress index (DI) using the pavement distresses is the most commonly used method. This practice has some inherent shortcomings: (i) the pavement rates of deterioration are excluded from the process, and (ii) the empirical nature of determining DI. These limitations could lead to inaccurate decisions regarding: (a) the identification of the appropriate boundaries of the pavement projects, (b) the determination of the optimum timing for pavement preservation and rehabilitation, and (c) the selection of the most-cost effective pavement fixes.
3. The observations of the pavement surface conditions made by designated staff members who frequently drive over the network and observe the surface conditions. Although this method has been used by most SHAs for a long time, the method is ineffective, the observations of the pavement surface

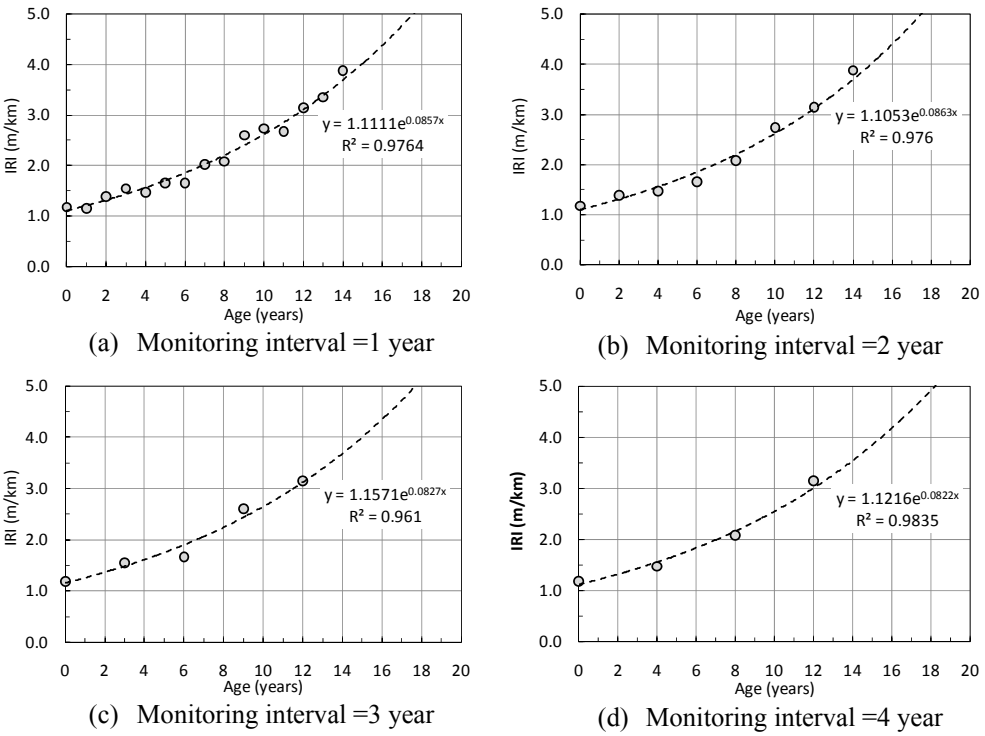
conditions are typically not documented, and the pavement rates of deterioration are excluded.

Typically, most Pavement Management System (PMS) decision making is based on pavement performance prediction by using the observed or measured pavement performance. However, if the performance prediction (or RSL) is uncertain or over or under-predicted, the PMS will be much less effective. The main objective of the paper is to assess the impact of condition monitoring interval on the PMS decision by utilizing the observed pavement performance in the field. The observed pavement performance also reflects the state's practices for pavement maintenance and rehabilitation. In this paper, two types of analysis were performed: (a) effect of monitoring interval on performance prediction, (b) impact of monitoring interval on PMS decisions. It should be noted that the analyses have been performed from two data sources (the MnRoad and the state of Washington) where adequate time series performance data for different performance measures (cracking, rutting and IRI) were available. The results of MnRoad and Washington data analyses are presented below.

## EFFECT OF MONITORING INTERVAL ON PERFORMANCE PREDICTION

The MnRoad performance data were used to illustrate the effects of monitoring interval on the performance prediction. The availability of performance data for 16 years (time series structural and functional distresses from 1994 to 2008) for various pavement sections designated as cells was the main motivation for utilizing the MnRoad performance data and investigating the impact of monitoring frequency on performance prediction. It should be noted that IRI and rutting distresses are monitored using sensor-based technology while cracking severity and extent are measured using imaged-based sensors. A particular distress type growth can be captured using a specific model form depending on the progression over time. For example, historically IRI over time can be modeled using exponential form while power and logistic models are used to represent rutting and cracking growth, respectively. Figure 1 shows an example of IRI prediction models with varying monitoring interval for a flexible pavement section (cell 3). The following observations can be made from these results:

- An exponential form of model fits best to IRI time series data.
- The variability of the data with time is very low; this can be seen from an accurate fit (high $R^2$) of the model through observed IRI data.
- The monitoring intervals of 1 to 4 years were obtained by eliminating the observed data from every other year for 2 year interval, every two years for 3-year interval and so on.
- The impact of monitoring interval is minimal i.e., the predicted life varies from 17 to 18 years to reach a threshold of 5 m/km (in this case). The impact of monitoring frequency is minimal because of very low variability of the IRI data over time. Therefore, the accuracy of monitored data will have an important role in investigating the impact of monitoring interval on performance prediction.




(a) Monitoring interval =1 year (b) Monitoring interval =2 year

(c) Monitoring interval =3 year (d) Monitoring interval =4 year

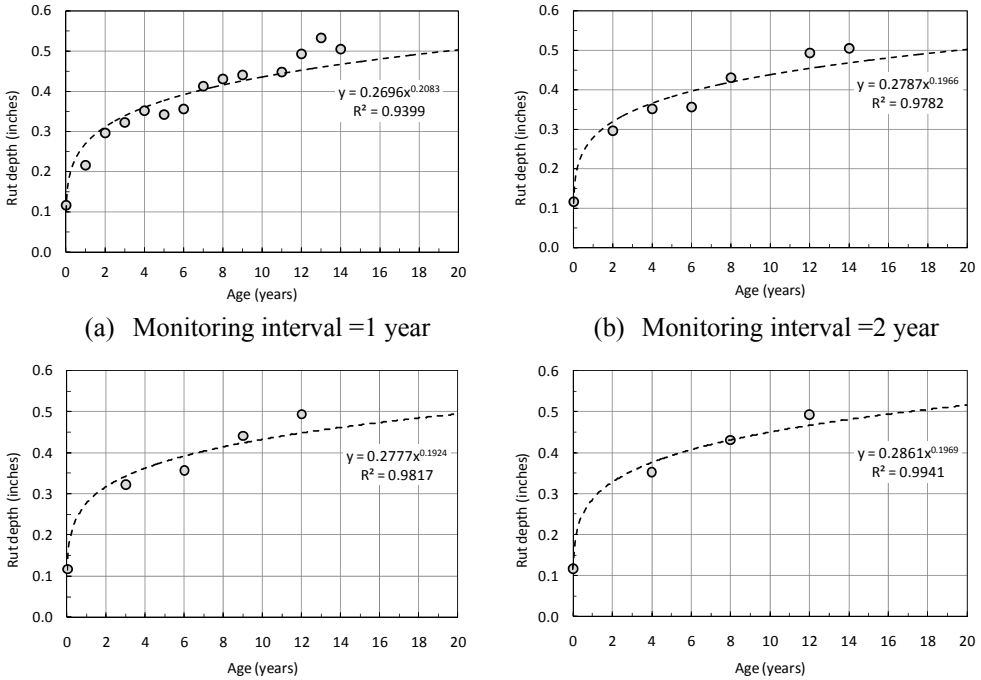
**Figure 1. Effect of monitoring interval on IRI prediction (cell 3 – flexible)**

To demonstrate the impact of accuracy or variability on performance prediction with different monitoring intervals, consider the case of rutting performance in another flexible pavement (cell 4) as shown in Figure 2. The figure shows an example of rutting prediction models with varying monitoring interval. The following observations can be made from these results:

- A power model form fit seems adequate for rut depth time series data.
- The variability of the data with time is high; this can be seen from a less accurate fit (lower $R^2$ as compared to IRI shown above) of the model.
- Again, the monitoring intervals of 1 to 4 years were obtained by eliminating the observed data from every other year for 2 year interval, every two years for 3-year interval and so on.
- The impact of monitoring interval seems significant i.e., the predicted life varies from 17 to more than 20 years to reach a threshold of 0.5 inch rut depth. The impact of monitoring frequency on performance prediction is significant because of more variability of the rut depth data over time. Therefore, the accuracy of monitored data has an important role in investigating the impact of monitoring interval on performance prediction.

At the project level, theoretically it is possible to monitor (rutting) a new pavement more frequently in the initial years (i.e., first 6 to 8 years) and subsequently the monitoring interval can be increased without sacrificing the accuracy in rutting prediction. However, in case of cracking higher monitoring intervals could be possible in the initial years of a project when not cracking is expected while higher monitoring intervals are warranted after the initiation of cracks. In summary, the monitoring intervals can be adjusted based on the expected distress growth over time

(i.e., the model form). On the other hand, from the network level data collection perspective, given the high variability of pavement surface conditions and ages along the entire length of network, it is not practical to use variable monitoring frequency for both sensor- and image-based data.




(a) Monitoring interval =1 year (b) Monitoring interval =2 year

(c) Monitoring interval =3 year (d) Monitoring interval =4 year

**Figure 2. Effect of monitoring interval on rut depth prediction (cell 4 – flexible)**

## IMPACT OF MONITORING INTERVAL ON PMS DECISIONS

The pavement condition monitoring frequency may impact the PMS decisions in three different ways:

- The identification of the boundaries of pavement projects.
- The determination of the optimum timing of pavement fixes.
- The selection of the most cost-effective pavement preservation strategy.

The boundary of the project on a same road may vary depending on the monitored condition. For example, the pavement condition may vary between different sections on the same project (uniform condition) if condition data are collected every year (1 year interval) versus every other year (2 years interval). Consequently, the timing of a particular fix will be affected by the non-uniformity of road sections. For example, some of the pavement sections having lower extent of distress will be triggered for a fixed even if those don't need any attention at that time. In addition, these discrepancies in decision making will lead to inconsistent fix type. Thus all these impacts of monitoring frequencies on PMS decision aspects will significantly affect the optimization of strategy at the network level. In order to investigate the effects of monitoring interval on the above mentioned aspects related to PMS decision, three types of analyses were conducted by utilizing observed performance data on a flexible pavement project (11 miles) in the state of Washington: (1) Effect of monitoring interval on uniform section boundaries based

on yearly assessment; (2) Consequence of monitoring interval on uniform sections based on time series data; and (3) Impact of monitoring interval on predicted distress growth.

The Washington PMS data were considered in the analyses because of the availability of ten years pavement condition data at one year interval. The availability of ten year data has facilitated the analyses for simulating 1, 2, and 3 year data collection intervals, respectively. The analyses were based on the data schemes listed in Table 1. The data schemes in the table indicate that for the one year pavement distress data collection frequency, the data are available every year. Whereas for the two year frequency, the observed and measured pavement distress data are available in odd years only (1999, 2001, 2003, etc.). The data for the even years were assumed to be the same as the data collected for the last odd year. For the three year frequency, the pavement distress data are collected once every three years. The distress data for the two years where data are not collected were assumed to be the same as the data from the last data collection activity. Finally, it was also assumed that decisions made on the basis of one-year distress data collection frequency are the most accurate and representative of actual conditions, and hence, those were used in the analysis as datum. Due to space limitations, the results of analyses are presented for a flexible pavement section from the Washington PMS data below; however, similar results were found in case of rigid pavements.

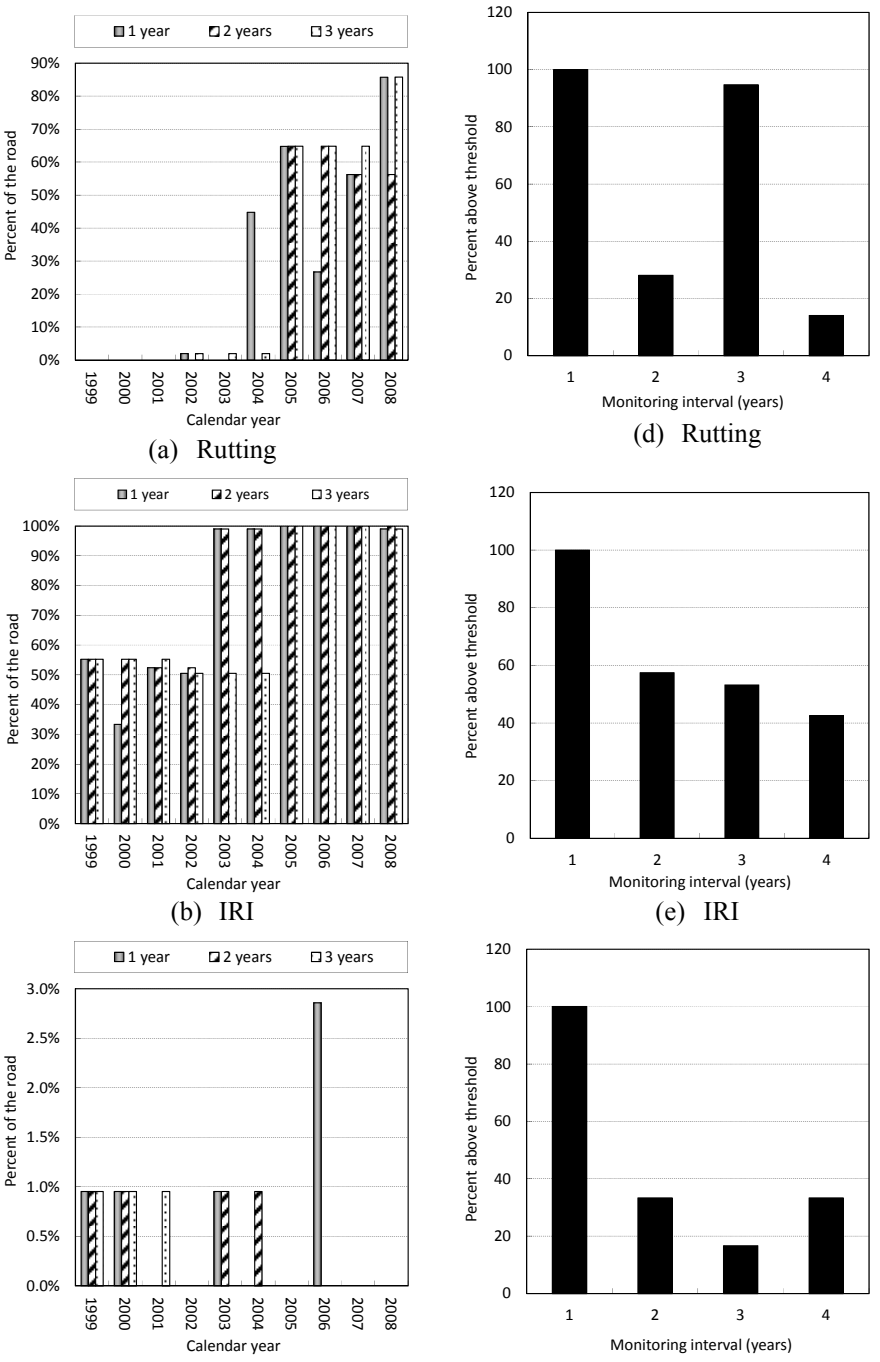
**Table 1. Distress data schemes for 1, 2 and 3 years pavement distress data collection frequency**

| Distress data survey year | Data collection frequency | | |
|---|---|---|---|
| | Every year | Every other year | Every third year |
| 1999 | ● | ● | ● |
| 2000 | ● | 1 | 2 |
| 2001 | ● | ● | 3 |
| 2002 | ● | 1 | ● |
| 2003 | ● | ● | 2 |
| 2004 | ● | 1 | 3 |
| 2005 | ● | ● | ● |
| 2006 | ● | 1 | 2 |
| 2007 | ● | ● | 3 |
| 2008 | ● | 1 | ● |

• Distress data are collected during the indicated year, 1. The pavement distress data for the indicated even years were assumed to be the same as the pavement distress data collected during the previous odd year. 2. The distress data for these years are assumed to be the same as the distress data collected during the previous year. 3. The distress data for these years were assumed to be the same as the distress data collected two years earlier.

**Impact of monitoring interval on uniform section boundaries.** Figure 3 shows the effects of distress data monitoring interval on the percent of the pavement sections reaching the specific threshold values of more than 0.25 inch, 60 inch/mile, and 50 feet for rut depth, IRI and longitudinal cracking in 0.1-mile section, respectively. It can be seen from the figure, that the percentages of the road reaching the threshold values are a function of the data collection frequency. Hence, the accuracy of the PMS based decisions depends on the year the decisions are being made relative to the year during which the distress data were observed or measured. For example, decisions made in the year 2006 (see Figures 3a to 3c) indicate that:

- Based on every year data collection frequency, about 25 percent of the pavement along the 10.37 miles pavement segment of SRID 005 has accumulated more than 0.25 inch rut depth; about 100 percent showed IRI of more than 60 inch/mile and about 2.75 percent showed 50 feet having low severity longitudinal cracking.
- Based on two and three years distress data collection frequencies, the above three percentages are 65, 100, and 0 percent, respectively.




(a) Rutting (d) Rutting

(b) IRI (e) IRI

(c) Longitudinal cracking (f) Longitudinal cracking

**Figure 3. Impact of monitoring interval on PMS decision (SRID 005 – Flexible)**

The above observations imply that the accuracy of the lengths of pavement projects and the boundaries of the pavement projects are affected by the data collection frequency. In addition, the data in Figures 3a to 3c also indicate that:

- If all decisions are to be made during odd years only, the decisions will be the same regardless if the pavement distress data were collected on a yearly or every other year basis. Stated differently, relative to the one year data collection frequency, the two-year frequency causes errors every other year where the data are not collected. These errors could be minimized if the pavement projects are to be selected in odd years for the next two-year period (2-year pavement preservation program).
- The three-year data collection frequency causes errors in the distribution and locations of the pavement sections reaching the threshold values. These errors are relatively higher than those caused by the two-year distress data collection frequency. Once again, the errors could be decreased if all decisions were made during the same year where the distress data are collected. The decisions must address three years pavement preservation program.

**Consequence of monitoring intervals on uniform section based on time series data.** Another way to investigate the impact of monitoring frequency on the PMS decision is to capture or flag pavement sections for a fix based on a particular distress threshold in the time series data at different monitoring intervals. The observed or measured time series data, for a particular pavement performance measure (e.g., rutting, IRI, or cracking), will vary in magnitude for 1, 2, and 3 year monitoring intervals (see Figures 3d to 3f). For example, if a pavement section develops a high magnitude (above the threshold) rut depth during any year, the section will be tagged for fixing in the 1 year data collection frequency. However, the section will not be captured to be fixed in the 2 and/or 3 year data collection frequency if the rut depth exceeded the threshold during the years where the data are not collected. To illustrate, the time series (ten years, 1999 to 2008) rut depth, IRI and longitudinal cracking data of the flexible pavement section SRID 005 indicate that some areas of the roads have been fixed during the 10-year period. The data for the three types of distress were analyzed to assess the impact of data collection frequency to capture 0.1-mile sections of the road when the threshold values, as mentioned before, were reached.

The analysis consists of searching the 10-year time series data to identify each 0.1-mile pavement section where the distress data exceeded each of the three threshold values stated above. In the analysis, the number of 0.1-mile sections that were captured in the 1-year data collection frequency was used as a datum and it was expressed as 100 percent capture. The number of 0.1-mile pavement sections that were captured in each of the 2 and 3 year frequency, was expressed as percent of the number of 0.1-mile sections captured using the 1-year frequency data. Figures 3d to 3f present the results for the three distress types. The data in the figure indicate that monitoring interval has a significant effect on the number of 0.1-mile pavement sections that were captured to be fixed. In all the three figures the data illustrate that lower data collection frequency (longer monitoring intervals) decrease the chance of capturing the pavement sections to be fixed. The impacts of the monitoring intervals seem to be more significant for rutting and cracking data for this particular pavement

section. Similar results were obtained for the rigid pavement section included in the database obtained from Washington DOT.

To further assess the impact of monitoring interval on the selection of pavement project boundaries, distributions of several distress types were generated for three monitoring intervals. Figures 4a to 4c show the distributions of various magnitudes of distresses along the same flexible pavement section SRID 005. It can be seen from the figure that for various distress magnitudes, the monitoring interval affects the distribution and hence, the selection of the boundaries of pavement projects.

**Impact of monitoring interval on distress growth.** Finally, the time series performance data were used to determine the yearly growth in the various distress types for three monitoring intervals. In reality, the yearly distress growth is a function of various factors including the (a) variability in the measured data over time, (b) inherent random variability with time, and (c) type and extent of the pavement fixes applied on the pavement surface over time

The yearly growth of each distress type for 1, 2, and 3 year monitoring intervals and the frequency distribution of the growth rates are illustrated in Figures 4d to 4f. The growth frequency distribution curves in the figure indicate that the 1 year monitoring interval has a wider distribution than the 2 and the 3 year intervals. This implies that low monitoring interval captures more distress growth rate information. It should be noted that the negative growth rates are the direct results of pavement fixes.

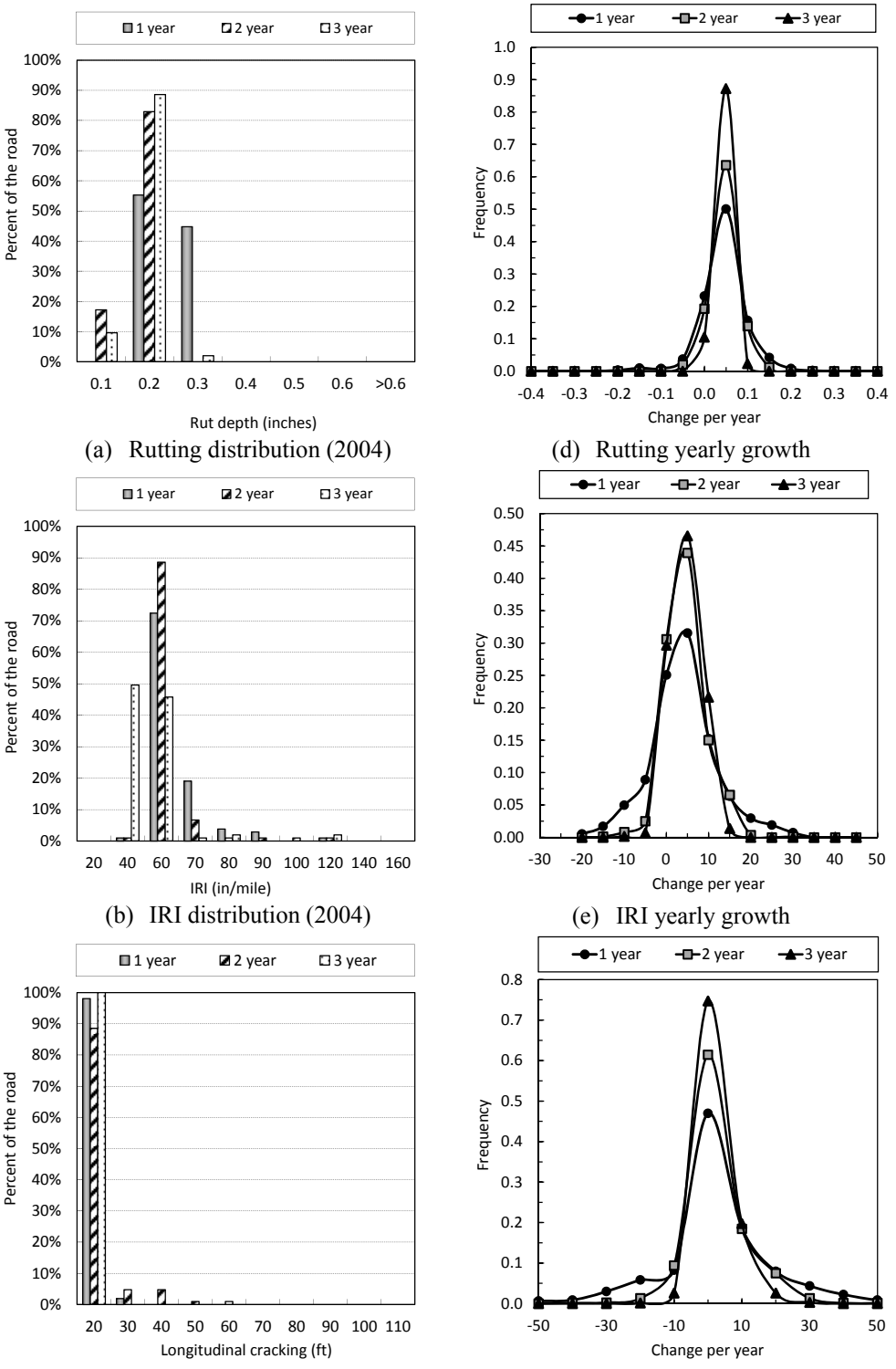
## CONCLUSIONS

Results of the analyses for investigating the effects of monitoring interval on performance prediction and on the PMS-based decisions can be summarized as follows:

1. Monitoring interval can have a significant impact on performance prediction, especially when the variability in the time series data is high. This discrepancy in the performance prediction can affect the PMS decision for future maintenance and rehabilitation actions at both project and network analyses.
2. Monitoring interval has a significant effect on the identification of the boundaries of pavement projects that are in need of repair. Such effects vary from one year to another and are a function of the year during which an agency makes decisions relative to the year the distress data are collected.
3. Longer monitoring interval decreases the chance of capturing or flagging pavement sections exceeding the specified distress thresholds.
4. Shorter monitoring intervals capture more distress growth information. The implication is that monitoring intervals affect the accuracy of the pavement performance prediction models developed using the time series distress data and hence, the accuracy of the PMS-based decisions.

## ACKNOWLEDGEMENT

The authors wish to thank the Federal Highway Administration (FHWA) for the financial support of the study. Washington DOT and Mr. Benjamin Worel from Minnesota DOT are acknowledged for providing the PMS data.




(a) Rutting distribution (2004)

(d) Rutting yearly growth

(b) IRI distribution (2004)

(e) IRI yearly growth

(c) Longitudinal cracking (2004)

(f) Longitudinal cracking yearly growth

**Figure 4. Impact of monitoring interval on distress distributions and yearly growth rate (SRID 005 – Flexible)**

## REFERENCES

Albitres, C. M. C., Smith, R. E., and Pendleton, O. J. (2007). "Comparison of Automated Pavement Distress Data Collection Procedures for Local Agencies in the San Francisco Bay Area." *Transportation Research Record*, 86th Annual Meeting CD ROM.

Benson, K., Elkins, G., Uddin, W., and Hudson, W. (1988). *Comparison of Methods and Equipment to Conduct Pavement Distress Surveys*, Transportation Research Board.

FORMAT. (2003). "Optimized Pavement Condition Data Collection Procedures." Project funded by the European Community under 'Competitive and sustainable Growth Programme (1998-2002).

FORMAT. (2005). "Final Technical Report: Fully Optimized Road Maintenance." Project funded by the European Community under 'Competitive and sustainable Growth Programme (1998-2002).

Haider, S. W., Baladi, G. Y., Chatti, K., and Dean, C. M. (2010). " Effect of Pavement Condition Data Collection Frequency on Performance Prediction." *Transportation Research Record*, 89th Annual Meeting CD ROM.

Hicks, R. G., and Mahoney, J. P. (1981). "Collection and Use of Pavement Condition Data." *NCHRP Synthesis of Highway Practice 76*.

McGhee, K. H. (2004). "Automated Pavement Distress Collection Techniques." *NCHRP Synthesis 334*.

McQueen, J. M., and Timm, D. H. (2005). "Statistical analysis of automated versus manual pavement condition surveys." *Transportation Research Record*(1940), 55-62.

Rada, G. R., Simpson, A. L., and Hunt, J. E. (2004). "Collecting and interpreting long-term pavement performance photographic distress data: Quality control-quality assurance processes." *Transportation Research Record*(1889), 97-105.

Wang, K., and Gong, W. "Automated Pavement Distress Survey: A Review and A New Direction."

# Thermal Cracking Prediction Model and Software for Asphalt Pavements

Eshan V. Dave[1], Sofie Leon[2], Kyoungsoo Park[2]
[1]Department of Civil Engineering, University of Minnesota - Duluth, 1405 University Drive, Duluth, MN, USA, evdave@umn.edu
[2]Department of Civil and Environmental Engineering,
University of Illinois at Urbana-Champaign, 205 N. Mathews Ave., Urbana, IL, USA
leon7@illinois.edu, kpark16@illinois.edu

## Abstract

Thermally induced cracking in asphalt pavements remains to be one of the prominent distress mechanisms in regions with cooler climates. At present, the AASHTO Mechanistic-Empirical Pavement Design Guide (MEPDG) is the most widely deployed pavement analysis and design procedure. For thermal cracking predictions, MEPDG utilizes a simplified one-dimensional stress evaluation model with a simple Paris-law (i.e. linear elastic fracture mechanics) based crack propagation procedure. The user-friendly graphical interface for MEPDG makes it an attractive design procedure of choice, however, the over simplicity of the model and lack of a physics-based representation to accurately capture the nonlinear fracture behavior of rate-dependent asphalt concrete reduce(s) the reliability of predictions. This study presents an interactive thermal cracking prediction model that utilizes a nonlinear finite element based thermal cracking analysis engine which can be easily employed using a user-friendly graphical interface. The analysis engine is comprised of (1) the cohesive zone fracture model for accurate simulation of crack initiation and propagation due to thermal loading and (2) the viscoelastic material model for time and temperature dependent bulk material behavior. The graphical user interface (GUI) is designed to be highly interactive and user-friendly in nature, and features screen layouts similar to those used in the AASHTO MEPDG, thus minimizing transition time for the user. This paper describes the individual components of the low temperature cracking prediction software (called LTC Model) including details on the graphical user interface, viscoelastic finite element analysis, cohesive zone fracture model, and integration of various software components for thermal cracking predictions.

## Introduction and Background

Asphalt pavements are usually constructed in a continuous manner without presence of periodic joints resulting in smooth driving characteristics. During the periods of severe low temperature climatic events, caused by low absolute temperatures or high cooling rates, thermal stresses build up in the continuous asphalt concrete layer. Damage accumulates as thermally induced stresses approach the material strength, and eventually cracks are formed. This type of cracking is commonly referred to as thermal cracking. Thermal cracking in asphalt pavements is often formed in periodic manner and is a prevalent form of pavement distress and damage mechanism in areas with cold climates. Figure 1 shows a typical thermal crack in asphalt pavement. The cracks in the figure have been sealed and hence more clearly visible in the picture.

Figure 1. Thermal cracking in asphalt concrete pavements.

In order to tackle thermal cracking distress from a design perspective, the most widely accepted pavement design guide in United States, the AASHTO Mechanistic Empirical Pavement Design Guide (MEPDG), utilizes a one-dimensional viscoelastic analysis program with a Paris law cracking criteria based on linear elastic fracture mechanics (LEFM). The analysis program in the MEPDG was developed by Roque et al. (1996). A number of studies in recent years have demonstrated that fracture in asphalt concrete is a highly non-linear phenomenon, typically characterized as quasi-brittle behavior. This has been demonstrated through modeling (Song et al. 2006), and through laboratory experiments (Wagoner et al. 2005; Li et al. 2006) amongst others.

An accurate model is necessary to design asphalt concrete pavements that are resistant to thermal cracking. The model must represent the time and temperature dependent viscoelastic material behavior and capture the nonlinear fracture behavior of quasi-brittle materials. Cohesive zone fracture models allow for accurate and efficient representation of the quasi-brittle fracture in asphalt concrete (Song et al. 2006), while a viscoelastic finite element analysis procedure, such as recursive-incremental scheme (Yi and Hilton 1993; Zocher et al. 1997; Dave et al. 2010) captures the rate- and temperature- dependent material behavior. Thus, a cohesive zone fracture model with a viscoelastic finite element analysis engine is a suitable analysis procedure for thermal cracking simulation in asphalt pavements. This type of procedure has been successfully utilized to model thermal cracking in various pavement test sections (Marasteanu et al. 2007; Dave et al. 2008). Previous studies have utilized commercial finite element software with cohesive zone fracture models and user-defined viscoelastic material models. However, use of commercial software is a major hindrance in wide-spread deployment of such analysis procedures to public and private agencies. In the current study, a stand-alone analysis and design software to predict thermal cracking performance of asphalt concrete pavements is presented. It provides an intuitive and user-friendly graphical user interface (GUI) as a means to perform rigorous viscoelastic finite element analysis with cohesive zone modeling. Description of various components of the analysis model is presented in the next section.

## Low Temperature Cracking Prediction Software

This section describes various components of the thermal cracking prediction software. The GUI and input file generator are described first, followed by the viscoelastic finite element analysis engine and the cohesive zone fracture model.

Graphical User Interface (GUI)

The standalone low-temperature cracking prediction software unifies several analysis modules into a user-friendly GUI. A graphical representation of the interaction between the GUI and the analysis modules is shown in Figure 2. The GUI collects and compiles the input conditions provided by the user and calls the Input File Generator to generate all necessary files for the finite element analysis engine. As described in Figure 2, the three main outcomes of the Input File Generator are the geometric data file, material data file and temperature boundary condition file. Finally, the GUI executes the finite element simulation to evaluate the potential of thermal cracking. The details of the codes listed in Figure 2 are discussed in the remainder of this section.

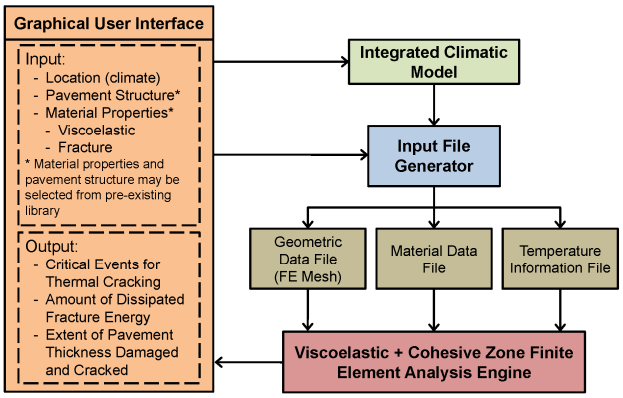
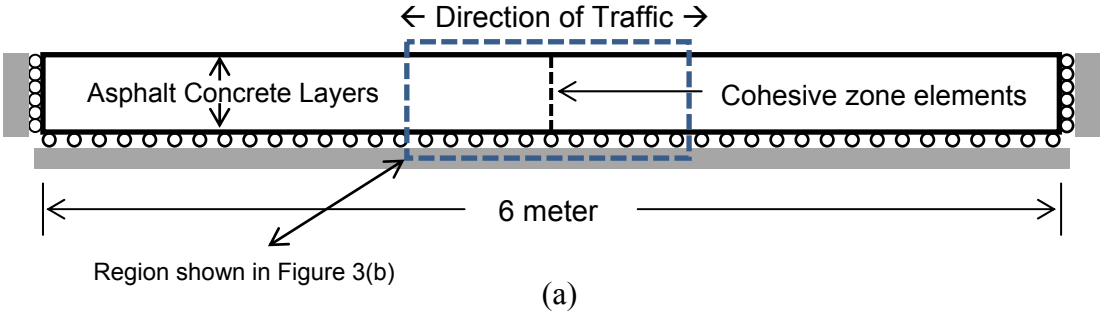


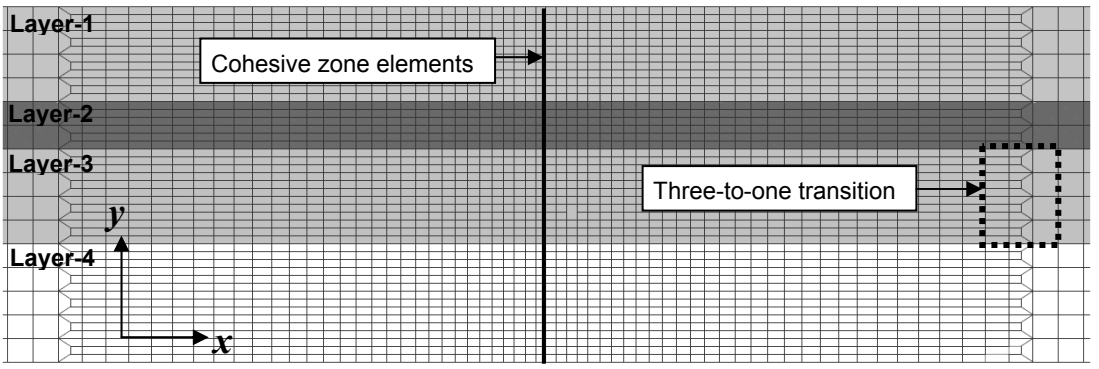

Figure 2. Flowchart of interaction between GUI and analysis modules

The first task of the Input File Generator is to develop a finite element mesh for the selected pavement geometry. The finite element mesh consists of coordinates of the nodal points, and an element connectivity table that links node numbers to their respective elements. During the first phase of low temperature pooled fund study an initial version of mesh generator was developed (Marasteanu et al., 2007). In the present work, this mesh generation code was significantly extended to develop full pavement models, perform checks for inconsistencies in the mesh, and automatically insert interfacial cohesive elements. Based on the recommendations and findings from previous studies (Paulino et al. 2006; Dave et al. 2007), the domain size and minimum element sizes were selected as 6 m and 4 mm respectively. The mesh generator utilizes an automatic transition scheme to reduce the computational cost of the problem. The code generates a finite element mesh using four node quadrilateral elements (Q4) and it automatically increases the element side lengths in the longitudinal direction of pavement (x-direction) until the relative difference in the

element side lengths reach 30%. At this point the mesh generator combines the smaller elements into one larger element using a three-to-one transition scheme. Figure 3 shows a typical pavement mesh, including the three-to-one transition, generated using the software. The code supports multiple lifts of asphalt concrete, each with distinct material properties and thicknesses. To insert cohesive interface elements, the code traverses the mesh and generates duplicate nodes along the potential crack path. Next, cohesive zone elements are inserted and attached to the duplicate nodes. The location of cohesive elements is also illustrated in Figure 3.




(a)




(b)

Figure 3. Finite element mesh from input file generator: (a) Illustration of a model geometry and boundary conditions (domain size: 0.18 m by 6 m), and (b) Close-up of the mesh in vicinity of cohesive zone elements

The second task of the Input File Generator is to create the material data file, which is primarily based on the information provided by the user. This file consists of viscoelastic (bulk) properties, the thermal expansion and contraction coefficient, and fracture properties. The list of properties utilized by the analysis code is shown in Figure 4.

At the current stage of the software, the user must provide the thermo-viscoelastic material properties in form of Prony series parameters (Generalized Maxwell model) and time temperature shift factors. The viscoelastic model coefficients can be determined by using creep testing of asphalt concrete following the AASHTO T-322 test procedure. Prior to full-scale deployment of the software, modifications will be made such that users can directly enter laboratory measured 1000 second creep test data from three temperatures. Tensile strength can also be determined using the AASHTO T-322 test procedure.

The fracture energy of asphalt concrete can be determined using a variety of test geometries, such as disk-shaped compact tension (DC[T]), semi-circular bend (SC[B]) and single-edge notched beam (SEN[B]) test. Currently, the model is anticipated to be calibrated and validated for the fracture energy obtained from the ASTM D7313 test procedure that utilizes DC[T] test geometry. Furthermore, the test is expected to be performed at crack mouth opening displacement (CMOD) rate of 0.0167 mm/s and at temperature of 10°C above the 95% reliability Superpave PG low temperature grade, as dictated by the project location.

Material Properties Required for Analysis Engine

1. Parameters for generalized Maxwell model (spring and dashpot coefficients) and reference temperature
2. Time-temperature shift factors for two temperatures other than reference temperature
3. Coefficient of thermal expansion and contraction
4. Fracture energy
5. Tensile strength

Figure 4. Material properties required by the LTC model

The user can either directly input the coefficient of thermal expansion and contraction (CTEC) or provide asphalt mixture volumetric properties. If volumetric properties are provided, the CTEC is estimated by using the same approximation equation that is utilized by the AASHTO MEPDG software.

Finally, the Input File Generator also provides the analysis engine with temperature loading conditions. The temperature profiles within the asphalt pavement are determined using the Integrated Climatic Model (ICM) developed by Larson and Dempsey (1997). ICM is also used in MEPDG for determining pavement temperature profiles. The ICM utilizes climatic information such as air temperature, wind velocity, cloud cover, precipitation, etc. to predict the pavement temperature through a finite difference heat flow analysis. Prediction of transient temperatures through the pavement thickness makes the ICM optimal for determining temperature boundary conditions, which are applied in the thermal cracking analysis. The finite element framework requires that temperature conditions be applied at each node, however the ICM outputs one-dimensional pavement temperatures with depths that do not necessarily correspond to those nodal locations. To overcome this challenge, the ICM is executed prior to the finite element analysis. A pre-processor uses the ICM output to generate nodal temperatures at the necessary locations and times, which are then passed to the finite element analysis engine.

Viscoelastic Finite Element Analysis

Finite element analysis is becoming increasingly popular in the design and analysis of pavements, for example, the current AASHTO design guide (MEPDG) utilizes finite element analysis for determination of critical pavement responses. The ability to model complex geometries and boundary conditions make finite element analysis well-suited for simulation of asphalt pavements. A typical response of asphalt concrete at low and intermediate temperatures is viscoelastic in nature. Figure 5 shows the creep compliance master-curve of asphalt concrete at two temperatures; a

Prony series model is fitted to the laboratory data. The change of loading time and temperature provides significant effect on the material behavior. In order to accurately capture the hereditary and temperature dependent behavior of asphalt concrete, a thermo-viscoelastic analysis procedure is needed. A variety of formulations have been proposed for thermo-viscoelastic finite element analysis (Yi and Hilton 1993; Zocher et al. 1997; Muliana and Khan 2008). In the present work, an incremental-recursive finite element formulation is utilized for thermo-viscoelastic finite elements (Yi and Hilton 1993; Zocher et al. 1997). This type of formulation has also been used for simulation of cracking in aged asphalt pavements (Dave et al. 2010). The analysis engine with thermo-viscoelastic finite elements is currently in the verification stage.

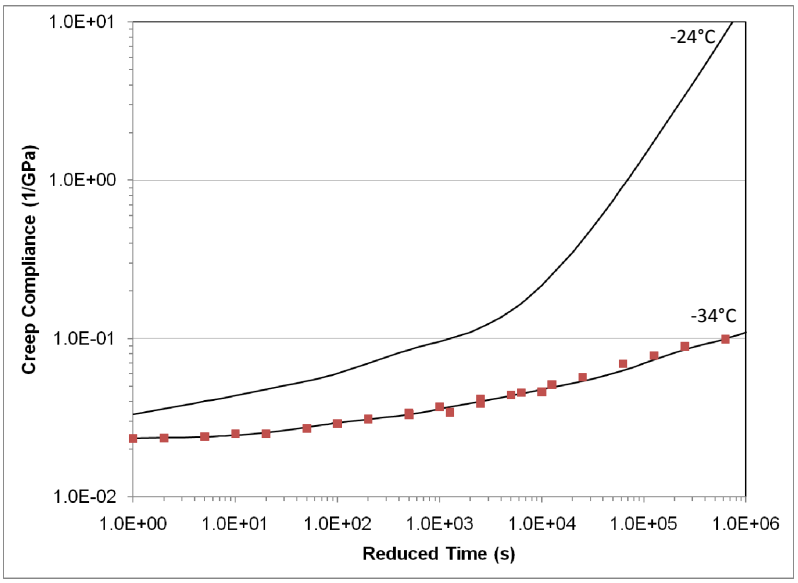



Figure 5. Creep compliance mastercurves for a asphalt concrete mixture at temperatures of -34 (Superpave PG low temperature grade, PG-LT) and -24°C (PG-LT+10°C)

Cohesive Zone Fracture Model

Asphalt concrete is classified as a quasi-brittle material because of the large nonlinear fracture process zone resulting from crack overlapping and branching, and from the weak interface between aggregates and asphalt binder. Such nonlinear fracture process zone is approximated by the cohesive zone model (Baranblatt 1959; Dugdale 1960). The cohesive zone model has been widely utilized to investigate a range of civil engineering materials such as Portland cement concrete (Hillerborg et al. 1976), reinforced concrete (Ingraffea et al. 1984), asphalt concrete (Song et al. 2006), and fiber reinforced concrete (Park et al. 2010), etc.

In the cohesive zone model, nonlinear cohesive traction is defined as a function of separation (or crack opening width) ahead of a macroscopic crack tip. A crack is initiated when the cohesive traction reaches the cohesive strength of the material. Note that further investigation is needed for crack initiation criteria. Then, as the

separation increases the cohesive traction decreases. Finally, when the separation is greater than a critical value, the material no longer has bearing capacity and the cohesive traction is zero. In this study, a linear softening model is employed, which is defined by the fracture energy ($G_F$) and the cohesive strength ($\sigma_{max}$). Additionally, an intrinsic cohesive zone modeling approach is used; hence a penalty stiffness (i.e. initial ascending slope) is introduced in the computational implementation. The initial penalty stiffness is determined on the basis of the numerical stability associated with the finite element implementation (Roesler et al. 2007).

## Use of LTC Model through GUI

This section briefly demonstrates the use of LTC Model through the GUI. Upon execution of the software, the user is greeted by the GUI that is organized into five sections: (1) Start, (2) Project Information, (3) Pavement Materials and Structure, (4) Run, (5) Results. For each section there are number of inputs required from the user. The flow of the program is described below in context of each of the five sections.

*(1) Start:* The user either opens an existing project or starts a new project. When existing project is opened all the inputs are pre-loaded into the GUI, but the user still has capability to alter or change any of the inputs. In case of new project, the user is required to provide all inputs.

*(2) Project Information:* The user inputs general information about the project including project name, location, length of analysis, etc., as shown in Figure 6. The location of analysis is necessary to select the pavement temperature profiles.

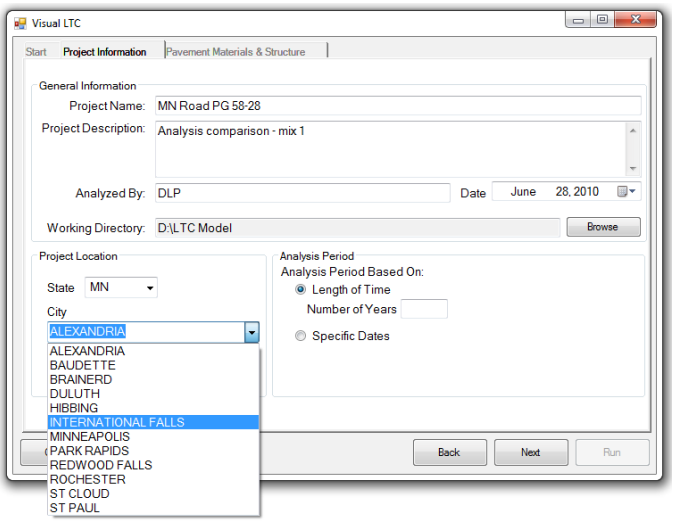



Figure 6. Project information

*(3) Pavement Materials and Structure:* The user builds the pavement structure by adding layers (i.e. asphalt layer, base layers, subgrade layer). Pavement structural characteristics and material properties can either be provided by the user or selected from a preexisting library. Figure 7 shows the third of the three steps required to add an asphalt layer.

*(4) Run:* Visual LTC executes the necessary analysis modules for pre-processing, finite element analysis and post-processing the results. As the analysis runs, the GUI informs the user of the runtime progress by indicating which stages of the analysis are complete and which are in queue to be executed.

*(5) Results:* The results from finite element analysis is converted to a user-friendly format and presented to user. Three sets of outputs are provided, namely: percent of fracture energy dissipated, extent of pavement thickness damaged and extent of pavement thickness cracked (shown in Figure 8). The outputs are available to users in both graphical and tabular formats with capability to export data in convenient comma separated value (CSV) format.

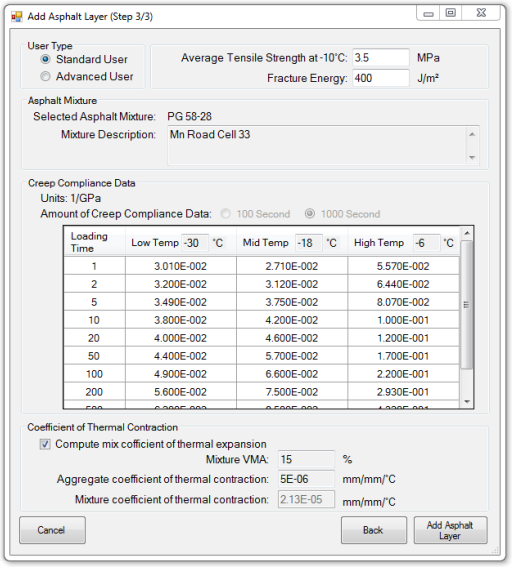



Figure 7. Step three of three to add an asphalt layer to the pavement structure

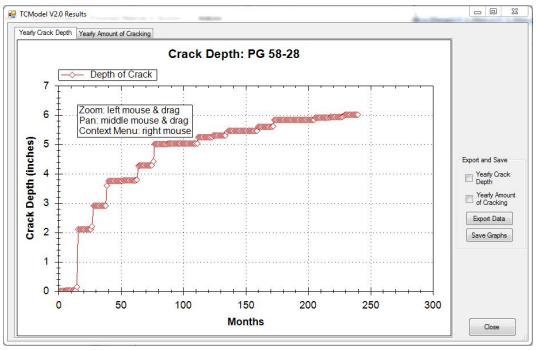



Figure 8: Depth of cracking in asphalt layer as graphical output from GUI

## Summary and Future Extensions

This paper describes the framework and organization of a low temperature cracking prediction model called "LTC Model". An overview of the various components of the model are presented along with in-depth details on the graphical user interface that acts as the pre- and post-processor to the finite element based analysis engine. The preprocessing step involves generation of finite element mesh, material properties and temperature boundary conditions. The finite element analysis is developed using a thermo-viscoelastic formulation for the bulk material response and a cohesive zone model for initiation and propagation of cracks. A brief demonstration is presented by showing use of the GUI for thermal cracking analysis. The development of LTC Model is presently in the final stages. Verification and calibration is currently being conducted on the finite element based analysis engine.

The thermal cracking prediction model is anticipated to be deployed to public and private transportation agencies in form of freely distributed software. Agencies participating in the low temperature cracking pooled fund study will form the first set of users to help fine tune the software and, if present, identify and resolve any

software bugs. There are several aspects of the model and the analysis procedure that can be extended to improve the capabilities and applicability of the model described in this paper, these includes:

- Critical climatic events that are likely to induce progressive damage to asphalt pavement should be identified. With critical climatic events identified only pavement in-service durations with such events will be simulated, thus reducing the computation costs.
- Tire loading conditions may be added to the analysis engine to include effects of traffic on the pavement cracking performance.
- Effects of loading rate and temperature on the cracking should be included to further increase the efficiency and reduce reliance on calibration. This could be achieved through extending the cohesive zone model to include aforementioned effects.
- Aging has significant effect on the bulk viscoelastic and fracture behavior of asphalt concrete. Aging models could be incorporated in the analysis engine such that, the mechanical properties are dependent on age, climate, thickness, volumetric properties, etc.

## Acknowledgements

Authors would like to acknowledge all the support and technical advice received from Professor William G. Buttlar and Professor Glaucio H. Paulino from the University of Illinois at Urbana-Champaign. We are grateful for the support provided by the sponsors and partners of Federal Highway Administration (FHWA) Pooled Fund Study TPF-5(132). Any opinions expressed herein are those of the writers and do not necessarily reflect the views of the sponsors.

## References

Barenblatt, G. I. (1959). “The formation of equilibrium cracks during brittle fracture: general ideas and hypotheses, axially symmetric cracks”. *Applied Mathematics and Mechanics*, 23(3), 622–636.

Dave, E. V., Song, S. H., Buttlar, W. G., and Paulino, G. H. (2007). "Reflective and Thermal Cracking Modeling of Asphalt Concrete Overlays.” *Proc., Int. Conf. on Advanced Characterization of Pavement and Soil Engineering Materials*, Athens, Greece, Taylor and Francis, 1241-1252.

Dave, E. V., Braham, A. F., Buttlar, W. G., Paulino, G. H., and Zofka, A. (2008). “Integration of Laboratory Testing, Field Performance Data, and Numerical Simulations for the Study of Low-Temperature Cracking.” *Proc. 6th RILEM International Conference on Cracking in Pavements*, Taylor and Francis, 369-378.

Dave, E. V., Buttlar, W. G., and Paulino, G. H. (2010). “Thermal Cracking Simulations of Aged Asphalt Pavements using Viscoelastic Functionally Graded Finite Elements.” *Proc., 11th International Conference on Asphalt Pavements, International Society of Asphalt Pavement*, Nagoya, Japan. (article in-press)

Dugdale, D. S. (1960). "Yielding of steel sheets containing slits." *Journal of the Mechanics and Physics of Solids*, 8(2), 100–104.

Hillerborg, A., Modeer, M., Petersson, P. E. (1976). "Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements". *Cement and Concrete Research,* 6(6), 773–781.

Ingraffea, A. R., Gerstle, W. H., Gergely, P., Saouma, V. (1984). "Fracture mechanics of bond in reinforced concrete." *Journal of Structural Engineering,* 110 (4), 871–890.

Larson G., and Dempsey B. J. (1997). "Enhanced Integrated Climatic Model: Version 2.0", Report No. DTFA MN/DOT 72114, Minnesota Road Research Project and Federal Highway Administration, Minneapolis, MN.

Li X., Marasteanu M. O., Iverson N., and Labuz J. F. (2006). "Observation of crack propagation in asphalt mixtures with acoustic emission", *Transp. Res. Rec.,* 1970, 171-177.

Marasteanu M., Zofka A., Turos M., Li X., Velasquez R., Li X., Williams C., Bausano J., Buttlar W., Paulino G., Braham A., Dave E., Ojo J., Bahia H., Gallistel A., and McGraw J. (2007). "Investigation of Low Temperature Cracking in Asphalt Pavements", Report No. 776, Minnesota Department of Transportation, Research Services MS 330, St. Paul, MN 55155.

Muliana A, and Khan K. A. (2008). "A Time-Integration Algorithm for Thermo-Rheologically Complex Polymers." *Computational Materials Science*, 41, 576-589.

Park, K., Paulino, G. H., Roesler, J. R., 2010. Cohesive fracture model for functionally graded fiber reinforced concrete. *Cement and Concrete Research,* 40 (6), 956-965.

Paulino, G. H., Buttlar, W. G., Blankenship, P. B., Wagoner, M. P., Song, S. H., and Dave, E. V. (2006). "Reflective Crack Control Treatment and. Design Procedures: A New Integrated Approach." *Final Report for CMS:0219566. National Science Foundation,* Washington, DC.

Roque R., Hiltunen D. R., and Buttlar W.G. (1996). "Thermal cracking performance and design of mixtures using Superpave", *Proc., Association of Asphalt Paving Technologists*, 65, 718-735.

Roesler J.R., Paulino G.H., K. Park, and Gaedicke C. (2007). "Concrete fracture prediction using bi-linear softening", Cement & Concrete Composites, 29, 300-312.

Song, S. H., Paulino, G. H., and Buttlar, W. G. (2006). "A bilinear cohesive zone model tailored for fracture of asphalt concrete considering viscoelastic bulk material." *Eng. Frac. Mech.,* 73:2829.

Wagoner, M. P., Buttlar, W. G., and Paulino, G. H. (2005) "Disk-Shaped Compact Tension Test for Asphalt Concrete Fracture." *Exp. Mech.,* 45, 270-278.

Yi S., and Hilton H. H. (1994). "Dynamic Finite Element Analysis of Viscoelastic Composite Plates in the Time Domain." *Int. J. Num. Meth. Eng.,* 37(12), 4081-4096.

Zocher, M. A., Groves, S. E., and Allen, D. H. (1997). "A Three-Dimensional Finite Element Formulation for Thermoviscoelastic Orthotropic Media." *Int. J. Num. Meth. Eng.,* 40(12), 2267-2288.

# Use of Fractional Order Viscoelastic Models to Characterize Asphalt Concrete

Samer W. Katicha, Ph.D., and Gerardo W. Flintsch, Ph.D., P.E.

## Abstract

The paper presents the application of fractional viscoelastic models to characterize viscoleastic properties of asphalt concrete. This implies the replacement of integer order derivatives in the constitutive equations with fractional derivatives. Integer order stress and strain derivatives lead to exponential relaxation and typically a large number of Maxwell or Kelvin elements are needed to characterize the full viscoelastic response range. In each case, the representation is not unique and the parameters cannot be linked to the composition of asphalt concrete. Fractional models lead to non-exponential relaxation making it possible to characterize the full viscoelastic response range with a small number of elements (typically 1 or 2). As such, the representation is unique and can be linked to the composition of asphalt concrete. Fractional models can also be used to construct the dynamic modulus master curve. As Witczak's sigmoidal model is a simple curve fitting it has no real physical meaning. Fractional models on the other hand have physical meaning, uniquely define the creep compliance and relaxation modulus, and allow better analysis of the physics of the relaxation process by considering the storage modulus, the loss modulus, and the phase angle. Using fractional models to analyze experimental asphalt concrete dynamic modulus results suggested two distinct relaxation processes; one at low temperatures and another at high temperatures. A possible explanation for this behavior can be attributed to the composition of asphalt concrete; at low temperatures, the binder behaves as a viscoelastic solid with aggregate particles more or less securely tied to the binder. The relaxation process is therefore restricted to the binder. At high temperatures, the binder behaves more as a viscoelastic fluid allowing aggregate particles to slide past each other which introduces another aspect of the relaxation process.

## Introduction

Over the last 30 years, fractional calculus has found extensive use in viscoelastic theory. The main advantage of fractional calculus over ordinary integer calculus is the ability to characterize the viscoelastic material response with only a few experimentally determined parameters (Soczkiewicz, 2002; Oeser et al., 2008). Initially, while mathematically sound, physical interpretation of fractional models was difficult, integer order derivatives offered the advantage of being equivalent to arrangements of springs and dashpots making them physically easy to visualize. However, in the 1990s Schiessel and Blumen (1993, 1995), Schiessel et al. (1995), and Heymans and Bauwens (1994) demonstrated that fractional order models can be realized physically through hierarchical arrangements of springs and dashpots. These models consist of a ladder, tree, or fractal structure, and in the limit of an infinite number of constituting elements, they are describe by fractional differential equations. More recently, Adolfsson et al. (2005) and Papoulia at al. (2010) have shown that fractional order models arise from the classical generalized Maxwell and generalized Kelvin models when the number of viscoelastic units tend to infinity. In other words, the Maxwell or Kelvin representations converge to corresponding fractional models in the limit as the number of units tends to infinity. These physical interpretations with springs and dashpots and the classical Maxwell and Kelvin models have provided fractional models with a link to classical models while at the same time keeping the

number of parameters of fractional models relatively small. Current efforts in fractional viscoelasticity focus on microscopic and molecular basis of fractional equations (Sharma and Cherayil, 2010).

Maxwell and Kelvin models that arise from integer differential equation models are extensively used to model viscoeasltic behavior of asphalt concrete (Park and Kim, 2001; Chehab et al., 2003; Al-Qadi et al., 2008; Katicha et al., 2008). The use of fractional viscoelastic models on the other hand has been limited to a few publications by Oeser and his collaborators (Oeser et al., 2008). The parameters of integer (Maxwell and Kelvin) models are generally determined from experimental data. It is known however, that the determination of these parameters is an ill-posed problem which has no unique solution (Honerkamp and Weese, 1989; Gerlach and Matzenmiller, 2005). Therefore, unless additional constraints are imposed, the identified parameters do not determine the exact solution, rather they represent an optimal solution that depends on a number of factors such as the number and distribution of relaxation times. As such, the same data set can be fit by very different model parameters. This does not cause problems for computational purposes of material response from applied loading as all models will practically lead to the same response (small differences will arise from numerical error). However, for purposes of relating the model to the mix physical properties (microscopic properties such as binder type, aggregate gradation etc.), Maxwell and Kelvin models are of little value. Fractional viscoelastic models on the other hand require a small number of parameters to characterize the viscoelastic properties of the material. As such determining the model parameters is a well-posed problem, and these parameters can theoretically be linked to material composition.

## Objective

The objective of the paper is to introduce fractional viscoelastic models using dynamic modulus test results. The models are used to construct the dynamic modulus master curve in a procedure similar to that proposed by Pellinen (1998), which uses a sigmoidal model. The advantages of the fractional viscoelastic models over the sigmoidal model are highlighted and it is shown how these models can point to mixture characteristics that are not detectible using the sigmoidal model.

## Background

There are a number of definitions of the fractional operator. The most widely used is the one based on the Riemann-Liouville fractional integral which associates with a real function $f$:R→R, its integral $D^{-\alpha}f$ of order $\alpha>0$ defined as

$$D^{-\alpha}f(x) = \frac{1}{\Gamma(\alpha)}\int_a^x f(t)(x-t)^{\alpha-1}dt \quad (1)$$

Where $\Gamma$ is the Gamma function and $a$ is an arbitrary fixed point (very often set to zero). For $\alpha$=1 Equation 1 gives the usual integral of the function and for positive integer values of $\alpha$, Equation 1 gives the usual Cauchy formula for repeated integration. The fractional derivative of order $\alpha$ can be defined as (among other possible definitions)

$$\frac{d^{\alpha}}{dx^{\alpha}}f = D^{\alpha-\lceil\alpha\rceil}\left(\frac{d^{\lceil\alpha\rceil}f}{dx^{\lceil\alpha\rceil}}\right) \quad (2)$$

Where $\lceil . \rceil$ represents the ceiling function that gives the smallest integer $\geq \alpha$. The fractional derivative in Equation 2 is referred to as the Caputo derivative which is the one used for fractional viscoelastic models along with a value of $a$=0. For more detail on fractional calculus the reader is referred to Oldham and Spanier (1974).

In viscoelasticity, the simplest fractional differential equation linking stress and strain is that of the springpot as described by Koeller (1984)

$$\sigma = E\tau^{\beta}\frac{d^{\beta}}{dt^{\beta}}\varepsilon \tag{3}$$

For $\beta$=0 the equation represents linear elastic behavior (Hookean spring) and for $\beta$=1 Newtonian viscosity (dashpot). For intermediate values of $\beta$, the springpot describes viscoelastic behavior. The springpot on its own has an infinite instantaneous frequency and a constant phase lag $\beta\pi/2$ independent of the frequency. This is clear from the complex modulus of the springpot

$$E^{*} = E(i\omega\tau)^{\beta} = \left(i\omega E^{-\beta}\tau\right)^{\beta} = (i\omega\tau_{E})^{\beta} \tag{4}$$

Note that in Equation 4 any combination of $E$ and $\tau$ is equivalent to a springpot with $E$=1 and the appropriate $\tau_E$. The reason for writing the equation with the constant $E$ is for convenience (see Equation 5). To describe real viscoelastic behavior, the springpot is associated with at least one elastic element. Adding a spring in series with the springpot would result in a model similar to the classical Maxwell model with the dashpot replaced by a springpot. In this case, compliances of each element can be added giving the complex modulus of the modified Maxwell element as (Heymans, 1996)

$$E^{*} = \left(\frac{1}{E} + \frac{1}{E(i\omega\tau)^{\beta}}\right)^{-1} = E\frac{(i\omega\tau)^{\beta}}{1+(i\omega\tau)^{\beta}} \tag{5}$$

The real (storage modulus) and imaginary (loss modulus) parts of the complex modulus can be obtained after some mathematical manipulation as

$$E' = E\frac{(\omega\tau)^{\beta}+\cos(\beta\pi/2)}{(\omega\tau)^{-\beta}+(\omega\tau)^{\beta}+2\cos(\beta\pi/2)} \tag{6}$$

$$E'' = E\frac{\sin(\beta\pi/2)}{(\omega\tau)^{-\beta}+(\omega\tau)^{\beta}+2\cos(\beta\pi/2)} \tag{7}$$

The storage and loss modulus for $E$=1, $\tau$=1 and different values of $\beta$ are presented in Figure 1 and Figure 2, respectively. The case of $\beta$=0 represents a purely elastic material while the case of $\beta$=1 represents the classical Maxwell model consisting of a spring in series with a dashpot. Note how the frequency dependence of the storage modulus changes with different values of $\beta$. The classical Maxwell model has a relatively short transition zone. To characterize real viscoelastic materials, a large number of such elements are therefore needed to cover the larger transition zone of typical materials. The fractional model on the other hand has a transition zone whose frequency range depends on $\beta$. Therefore, a relatively small number of elements are needed (typically one or two) to describe real materials. Finally, the model described in Equation 5 has an equilibrium modulus of zero (similar to the classical Maxwell model); an additional spring placed in parallel to the model can provide for a non-zero equilibrium modulus.

**Construction of Asphalt Concrete Dynamic Modulus Master Curve**

Because of the ability to have a longer transition zone using fractional viscoelastic models, it is possible to use these models to construct the dynamic modulus master curve in a procedure similar to the one proposed by Pellinen (1998) with the sigmoidal model replaced by the

fractional viscoelastic model. The fractional viscoelastic model is however different from the sigmoidal model in that it has a physical representation in terms of mechanical analogues and results from a power law kernel in the integral representation of viscoelasticity. Also the model describes both the storage and loss modulus whereas the sigmoidal model gives a description of the dynamic modulus (actually, the logarithmically transformed dynamic modulus). As will be illustrated using experimental results, considering both storage modulus and loss modulus can give better information about the actual viscoelastic response of the material.

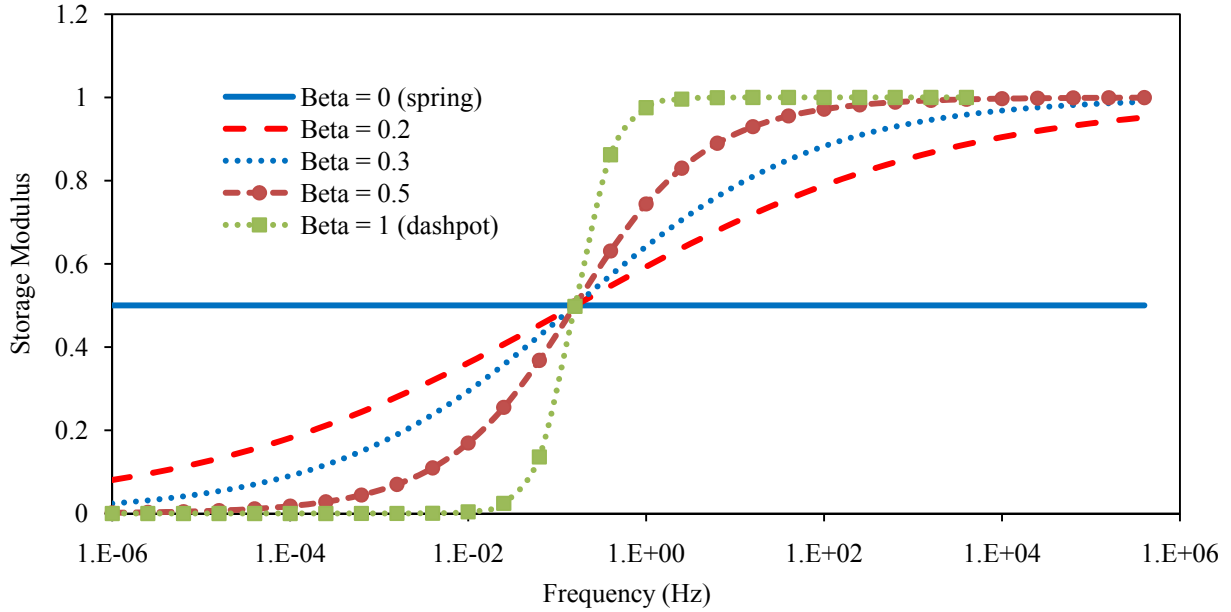



Figure 1. Storage modulus of the modified Maxwell model for different values of $\beta$.

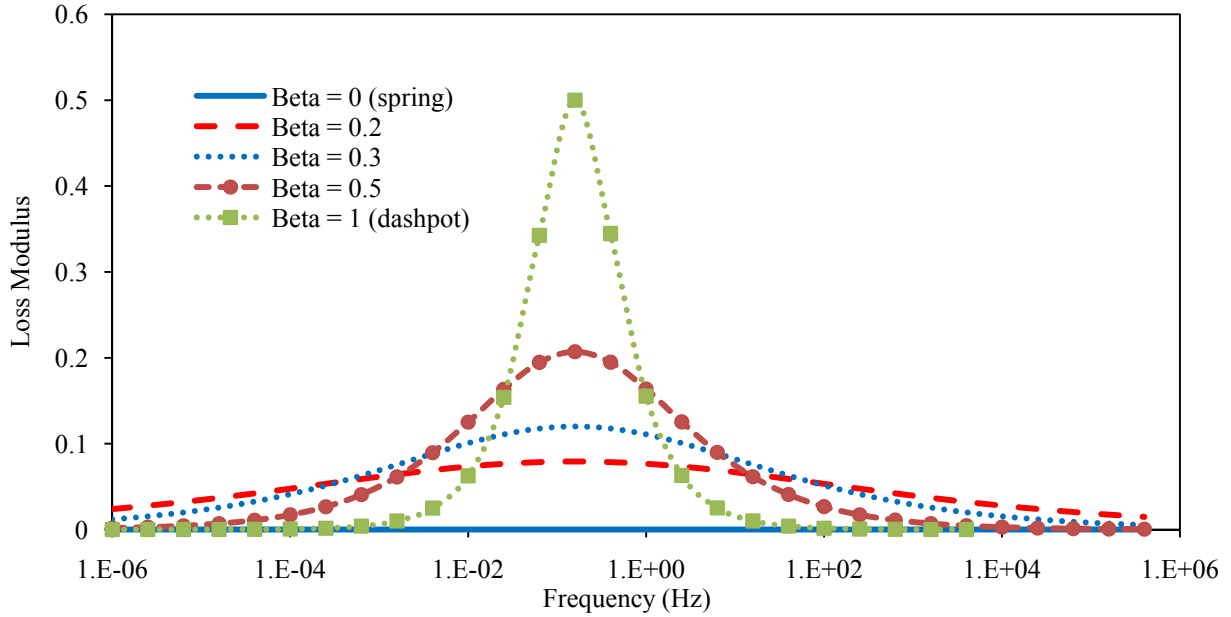



Figure 2. Loss modulus of the modified Maxwell model for different values of $\beta$.

Two fractional viscoelastic models were used to construct the dynamic modulus master curve. The first model consists of the modified Maxwell model placed in parallel with a spring. The second model consists of a similar arrangement with an additional modified Maxwell model placed in parallel. This model is therefore similar to a generalized Maxwell model consisting of two elements (along with the spring) with the dashpot replaced by a springpot. Note that there are other possible configurations of springs and springpots that can be used to generate different models. These can lead to the modified Kelvin model, the modified Zener model, the modified Burger model, and many more (Heymans, 1996). There are however thermodynamic constraints on the model parameters (Friedrich, 1991; Glöcke and Nonnrnmacher, 1991; Beris and Edwards, 1993; Palade et al., 1999; Heymans, 2003).

The master curve construction was performed in Excel using the Solver function by minimizing the error between the data and fractional model (storage modulus, loss modulus, and phase angle). Because of the large variation in dynamic moduli (range of three orders of magnitude), the error is taken on a logarithmic scale (except for the phase angle). Eleven typical mixes (3 surface mixes, 4 base mixes, and 4 intermediate mixes) used in Virginia were investigated. As the procedure for constructing the master curve is similar for all mixes, we only present the results of one mix. Dynamic modulus tests were performed at five temperatures (-12.2, 4.4, 21.1, 38.7, and 54.4°C), and 6 frequencies (25, 10, 5, 1, 0.5, and 0.1Hz) according to AASHTO TP62. Tests were performed from the lowest to the highest temperature and from the highest to lowest frequency. In each case three replicates were used and results represent the average of these three replicates. More details about the mixes and testing can be found in (Flintsch et al., 2007). Figure 3 shows the obtained master curve at the reference temperature of 21.1°C for a base mix (BM1) using the first model (i.e. one modified Maxwell element in parallel with a spring). Clearly, storage modulus and dynamic modulus test results at the lowest temperature (-12.2°C) seem under shifted and do not conform to the model. The reason for this is that further shifting of test results at -12.2°C will result in considerable error in the loss modulus. This is illustrated in Figure 4. In Figure 4 shifting was performed by minimizing the error in dynamic modulus between the test results and the fractional model as is done using the sigmoidal model. The resulting master curves show a very good agreement between the fractional model and test results for both the storage modulus and the dynamic modulus; however the agreement for the loss modulus is very weak at high reduced frequencies. Another evaluation of the validity of the model and the resulting master curve is to compare the phase angle from the model to the phase angle obtained from the dynamic modulus test. Figure 5 shows a significant difference between the model phase angle and the measured phase angle at reduced times between 0.001 and 1Hz. This difference is not obvious for the loss modulus and even less so for the storage modulus. The results presented show that the model with a single modified Maxwell element is inadequate to fully characterize the viscoelastic response of asphalt concrete. Note that the sigmoidal model cannot offer such a detailed analysis as the one presented here; it can only evaluate the model in terms of the dynamic modulus. This is however not sufficient to validate the model as Figure 4 shows that the model can adequately describe the dynamic modulus while at the same time poorly describe the viscoelastic response (loss modulus and phase angle). As a note, just like the sigmoidal model, the model with a single modified Maxwell element has four parameters; two parameters that describe the plateau maximum and minimum dynamic modulus values, one parameter that describes the slope of the model, and one parameter that describe the horizontal shift of the data. As such, the resulting master curve of Figure 4 is practically identical to the master curve obtained using the sigmoidal model.

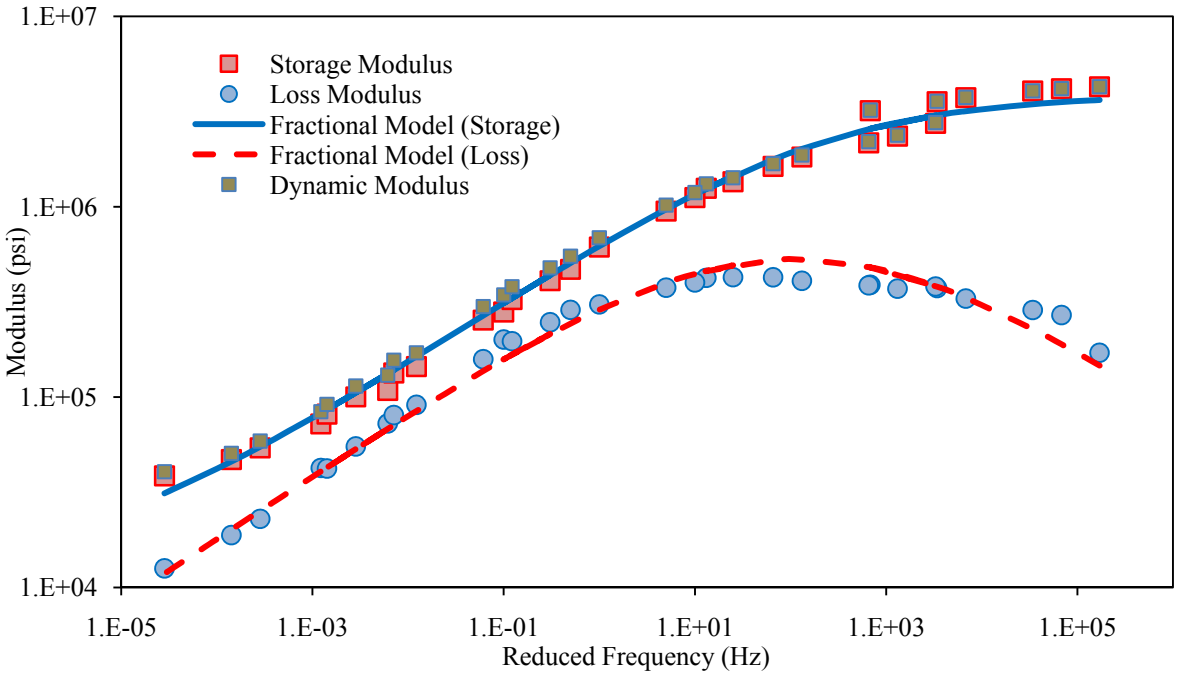



Figure 3. Master curve constructed with the one parameter model considering storage modulus, loss modulus, and phase angle.

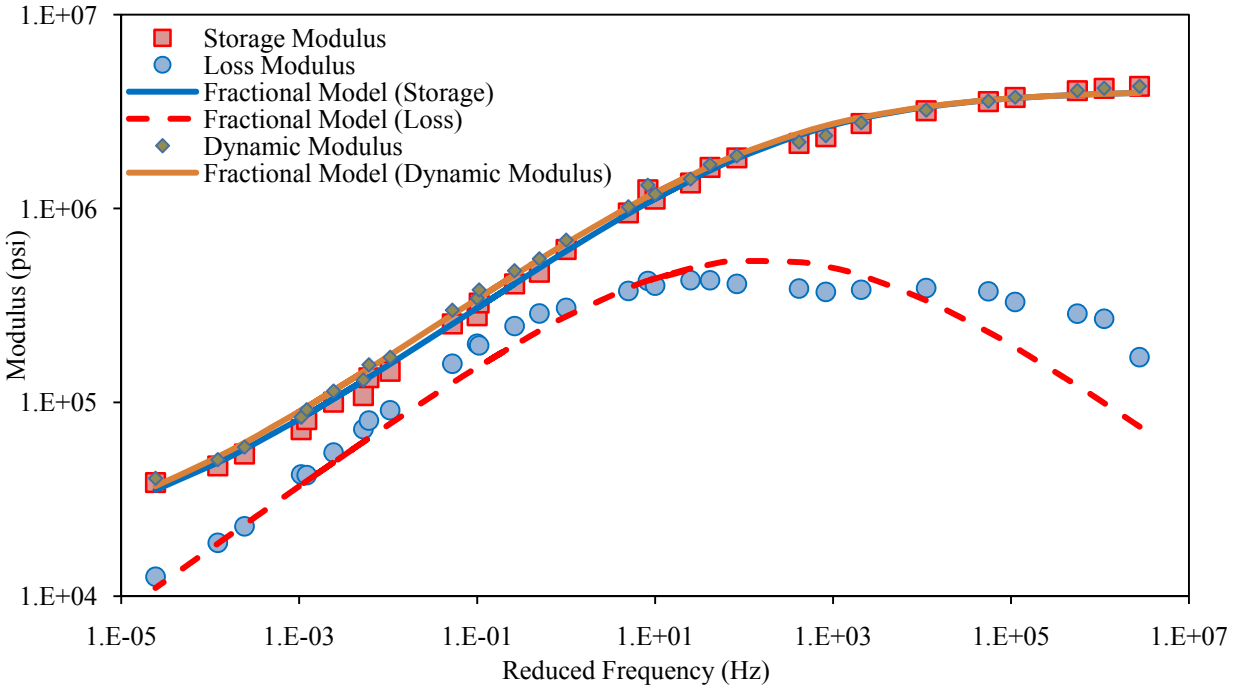



Figure 4. Master curve constructed with the one parameter model considering only the dynamic modulus.

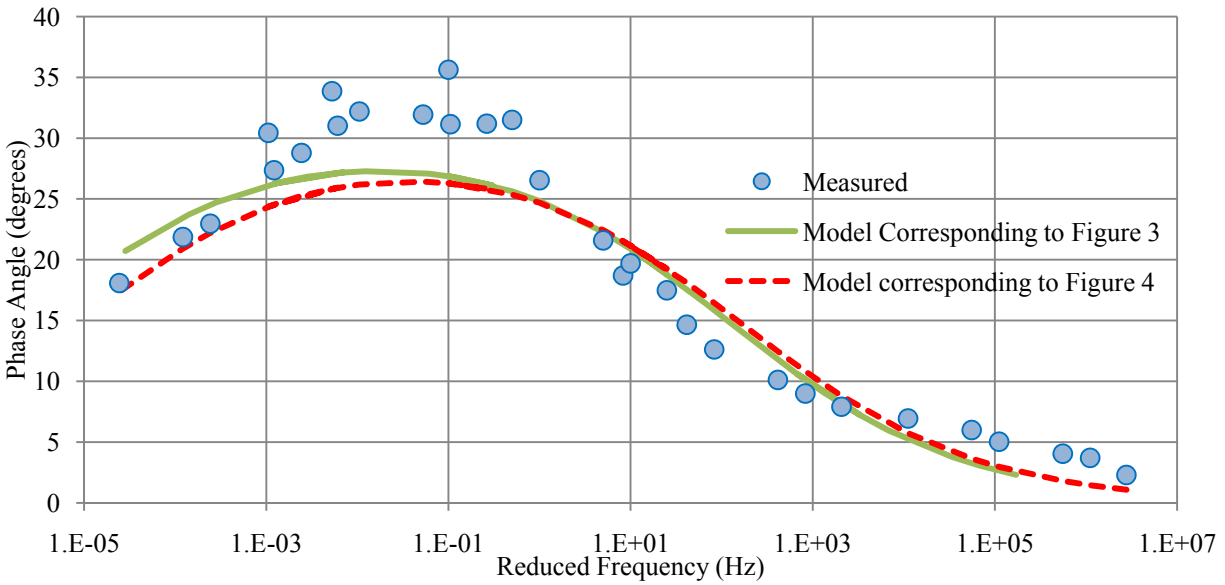



Figure 5. Phase angle using the one parameter model.

As the one parameter (one modified Maxwell element in parallel with a spring) fractional model fails to adequately represent the viscoelastic response of asphalt concrete, a two parameters (two modified Maxwell elements in parallel with a spring) fractional model was used to construct the master curve. The resulting master curve is presented in Figure 6. Agreement between the model and the test results for both storage and loss modulus is very good. The figure shows the two relaxation stages corresponding to each modified Maxell element. Although each stage extends over the whole reduced frequency range, one stage dominates the behavior at reduced frequencies above the 100Hz reduced frequency mark, while the other stage dominates the behavior at reduced frequencies below the 100Hz mark. Figure 7 presents the phase angle of the model and the experimental phase angle. The two parameters model shows a significant improvement between the model and test phase angle especially at reduced frequencies between 0.001 and 1Hz.

**Discussion**

The experimental results presented illustrate the advantage of fractional viscoelastic models compared to the sigmoidal model in their ability for detailed analysis of the viscoelastic response. They show that obtaining a good fit between the model dynamic modulus and the experimental dynamic modulus does not necessarily result in a good agreement of viscoelastic properties. Because the contribution to the dynamic modulus comes mainly from the storage modulus, the difference in viscoelastic properties is most evident in the loss modulus and phase angle. Furthermore, fractional viscoelastic models can be used to model the dynamic modulus master curve using mixture aggregate and binder properties in a similar manner the sigmoidal model is used in the Witczak predictive equation (Bari and Witczak, 2006). Witczak's model is purely empirical (Ceylan et al., 2009) and it has been argued that the model lacks accuracy at high temperatures and low frequencies (Schwartz, 2005; Donger et al., 2005; Al-Khateeb et al., 2006; Azari et al., 2007). Ceylan et al. (2009) obtained very good asphalt concrete dynamic modulus prediction using an artificial neural network (ANN) model. While their model gives

much improved predictions compared to the Witczak model it still suffers from the same limitations; that is it is empirical, does not give the storage modulus, loss modulus, or phase angle, and does not give any indication of the physical processes (even more so than the Witczak model) that control the dynamic modulus. One major reason why the ANN has such a good prediction capabilities is the large number of degrees of freedom (weights) that are adjustable. The fractional viscoelastic models have a physical and mathematical basis and the possibly larger number of degrees of freedom (7 for the two parameters model compared to 4 for the Witczak model) in the model can potentially improve prediction capabilities.

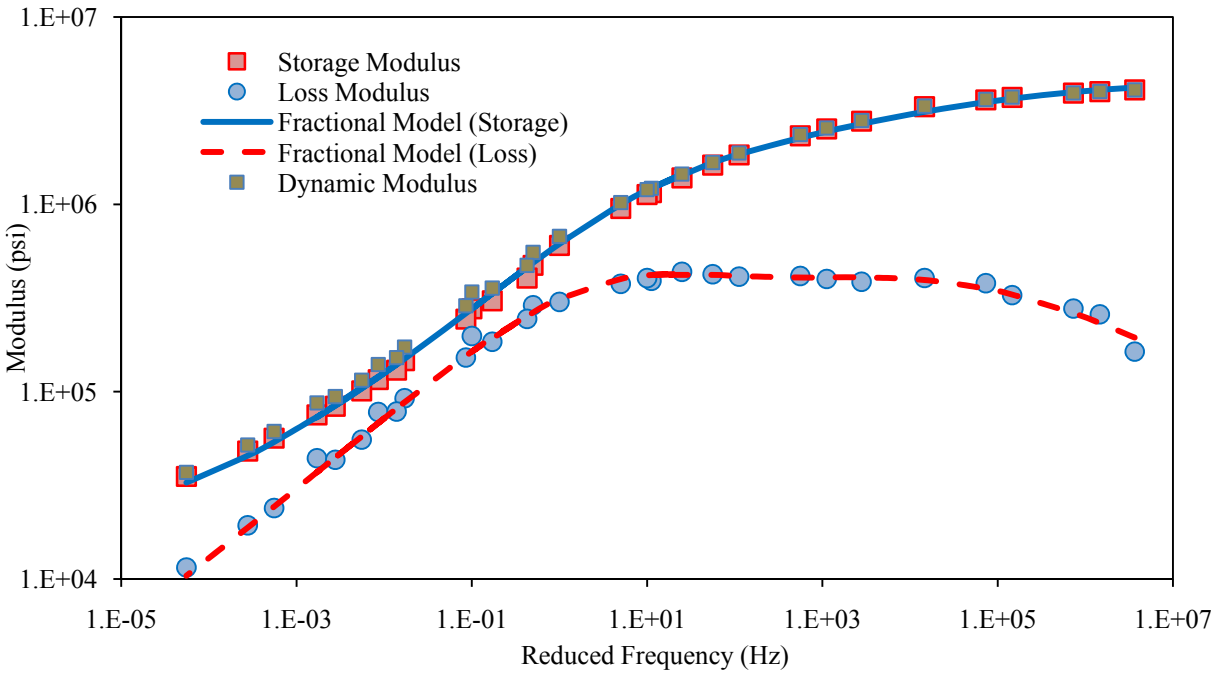



Figure 6. Master curve constructed with the two parameters model.

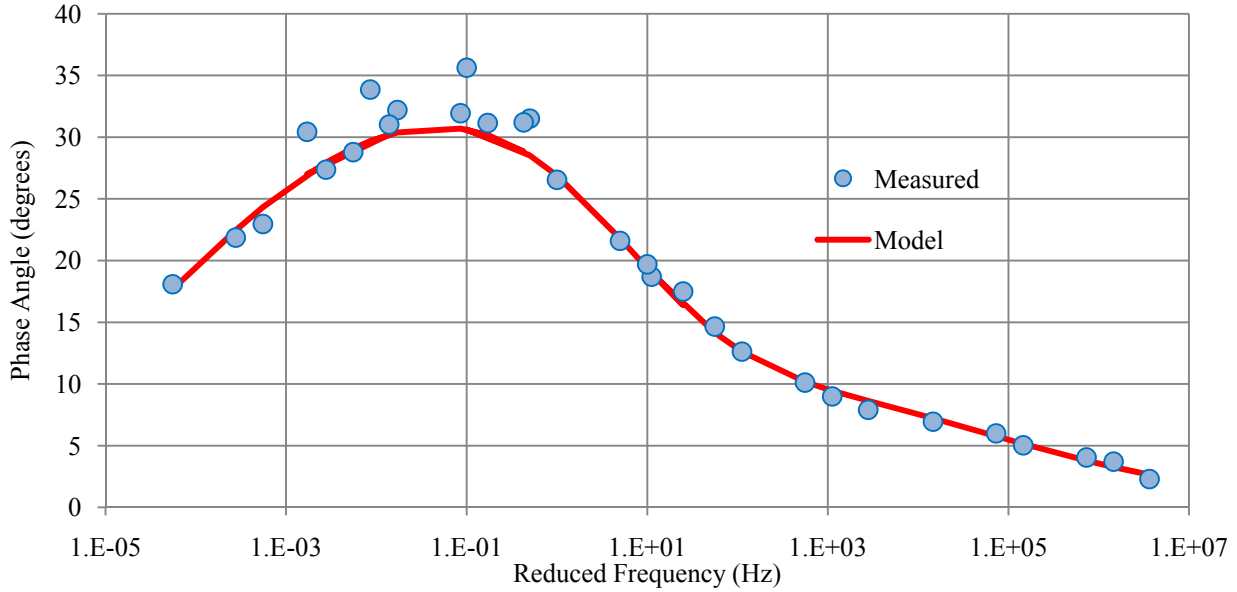



Figure 7. Phase angle from the two parameters model.

More importantly, fractional viscoelastic models have been to relate to physical characteristics of polymers. For example, Bagley (1983) showed that the Rouse model of chain dynamics leads to a constitutive equation with fractional derivative. Sharma and Cherayil (2010) obtained a molecular basis for fractional equations of viscoelasticity in terms of the generalized Langevin equation. Based on this success in polymer science, it would be interesting to investigate fractional viscoelastic models' relation to fundamental physical behavior that results from the individual properties and interactions of the main constituents that form asphalt concrete (binder, aggregates, and air voids). For example, at low temperature aggregate particles are strongly held by the stiff binder. Relaxation mainly occurs in the binder itself due to its viscoelastic nature. At high temperature, the binder is much softer and aggregates start to rotate and slip past each other. Aggregate interlock plays a much more important role in resisting deformation. This mechanism is physically different from the relaxation mechanism experienced at low temperatures which might explain why a two parameters model is needed to accurately represent the viscoelastic response of asphalt concrete. Each parameter will be a manifestation of one of the distinctive physical processes. This shows the potential to relate fractional viscoelastic models to physical characteristics of asphalt concrete.

## Conclusions

Fractional viscoelastic models have the potential to significantly further our understanding of asphalt concrete mechanical properties and how these properties relate to mixture components. The Witczak model is an empirical model that uses dynamic modulus results to construct the master curve. Fractional viscoelastic models have physical basis and show that considering only the dynamic modulus results in not sufficient to obtain a good representation of the viscoelastic properties of the material; tests results need to be analyzed in terms of storage modulus, loss modulus, and phase angle. This more detailed analysis allows for a better understanding of material behavior. The parameters of the fractional models have the potential to be related to aggregate and binder characteristics in a way similar to how the parameters of the Witczak model are, or more importantly, to fundamental physical properties that relate to microstructure theories. The increased detail in macroscopic properties provided by fractional models can lead to better understanding, calibration, and validation of microscopic models.

## Acknowledgments

We would like to thank the reviewers for their constructive comments.

## References

Al-Khateeb, G., Shenoy, A., Gibson, N., and Harman, T. (2006) "A new simplistic model for dynamic modulus predictions of asphalt paving mixtures" [CD-ROM]. *Journal of the Association of Asphalt Paving Technologists*, Vol. 75.

Al-Qadi, I.L., Elseifi, M.A., Yoo, P.J., Dessouky, S.H., Gibson, N., Harman, T., D'Angelo, J, and Petros, K. (2008) "Accuracy of complex modulus selection from vehicular load pulse: NCHRP Project 1-37A mechanistic-empirical pavement design guide" *Transportation Research Record*, Vol. 2087, pp. 81–90.

Azari, H., et al., (2007) "Comparison of measured STP E* of ALF mixtures with predicted E* using NCHRP 1-37 A and Witczak's new equations" [CD-ROM]. *Presented at 86th Annual Transportation Research Board Meeting*. Washington, DC: Transportation Research Board.

Bagley, R.L. (1983) "A theoretical basis for the application of fractional calculus to viscoelasticity" Journal of Rheology, Vol. 27, pp. 201–210.

Bari, J. and Witczak, M.W. (2006) "Development of a new revised version of the Witczak E* predictive model for hot mix asphalt mixtures" *Journal of the Association of Asphalt Paving Technologists*, Vol. 75, pp. 381–423.

Beris, A. N. and Edwards, B. J. (1993) "On the admissibility criteria for linear viscoelastic kernels" *Rheologica Acta*, Vol. 32, pp. 505–510.

Ceylan, H., Gopalakrishnan, K., and Kim, S. (2009) "Looking to the future: the next-generation hot mix asphalt dynamic modulus prediction models" *International Journal of Pavement Engineering*, Vol. 10 (5), pp. 341–352.

Chehab, G.R., Kim, Y.R., Schapery, R.A., Witczak, M.W., and Bonaquist, R. (2003) "Characterization of asphalt concrete in uniaxial tension using a viscoelastoplastic continuum damage model" *Journal of the Association of Asphalt Paving Technologists*, Vol. 72, pp. 315–355.

Dongre, R., Myers, L., D'Angelo, J., Paugh, C, and Gudimettla, J. (2005) "Field evaluation of Witczak and Hirsch models for predicting dynamic modulus of hot-mix asphalt" *Journal of the Association of Asphalt Paving Technologists*, Vol. 74, pp. 381–442.

Flintsch, G.W., Loulizi, A., Diefenderfer, S.D., Galal, K.A., and Diefenderfer, B.K (2007) "Asphalt materials characterization in support of implementation of the proposed mechanistic-empirical pavement design guide" *Contract Report VTRC 07–CR10*, Virginia Transportation Research Council, Charlottesville, Virginia.

Friedrich, Chr. (1991) "Relaxation and retardation functions of the Maxwell model with fractional derivatives" *Rheologica Acta* Vol. 30, pp. 151–158.

Gerlach, S., and Matzenmiller, A. (2005) "Comparison of numerical methods for identification of viscoelastic line spectra from static test data" *International Journal for Numerical Methods in Engineering*, Vol. 63, pp. 428–454.

Glöckle, W. G. and Nonnenmacher, T. F. (1991) "Fractional integral operators and Fox functions in the theory of viscoelasticity" *Macromolecules*, Vol. 24, pp. 6426–6434.

Heymans, N. (1994) "Fractal rheological models and fractional differential equations for viscoelastic behavior" *Rheologica Acta*, Vol. 33, pp. 210–219.

Heymans, N., and Bauwens, J.C. (1996) "Hierarchical models for viscoelasticity: dynamic behavior in the linear range" *Rheologica Acta*, Vol. 35, pp. 508–519.

Heymans, N. (2003) "Constitutive equations for polymer viscoelasticity derived from hierarchical models in cases of failure of time-temperature superposition" *Signal Processing* Vol. 83, pp. 2345–2357.

Honerkamp, J., and Weese, J. (1989) "Determination of the relaxation spectrum by a regularization method" *Macromolecules*, Vol. 22, pp. 4372–4377.

Katicha, S.W., Flintsch, G.W., Loulizi, A., and Wang, L. (2008) "Conversion of testing frequency to loading time applied to the mechanistic-empirical pavement design guide" *Transportation Research Record*, Vol. 2087, pp. 99–108.

Koeller, R.C. (1984) "Application of fractional calculus to the theory of viscoelasticity" *Journal of Applied Mechanics*, Vol. 51, pp. 299–307.

Oeser, M., Pellin, T., Scarpas, T., and Kasbergen, C. (2008) "Studies on creep and recovery of rheological bodies based upon conventional and fractional formulations and their application on asphalt mixture" *The International Journal of Pavement Engineering*, Vol.9(5), pp. 373–386.

Oldham, K.B., and Spanier, J. (1974) *The fractional calculus*. Academic Press, New York.

Palade, L. I., Attane, P., Huilgol, R. R., and Mena, B. (1999) "Anomalous stability behaviour of a properly invariant constitutive equation which generalises fractional derivative models" *International Journal of Engineering Sciences*, Vol. 37, pp. 315–329.

Papoulia, K.D., Panaskaltsis, V.P., Kurup, N.V., and Korovajchuk, I. (2010) "Rheological representation of fractional order viscoelastic material models" *Rheologica Acta*, Vol. 49, pp. 381–400.

Park, S.W., and Kim, Y.R. (2001) "Fitting Prony-series viscoelastic models with power law presmoothing" *Journal of Materials in Civil Engineering*, Vol. 13, Issue 1, pp. 26–32.

Pellinen, T. (1998) "The assessment of validity of using different shifting equations to construct a master curve of HMA" University of Maryland, Department of Civil and Environmental Engineering, College Park, MD.

Soczkiewicz, E. (2002) "Application of fractional calculus in the theory of viscoelasticity" *Molecular and Quantum Acoustics*, Vol. 23, pp. 397–404.

Schiessel, H., and Blumen, A. (1993) "Hierarchical analogues to fractional relaxation equations" *Journal of Physics A: Mathematical and General*, Vol. 26, pp. 5057–5069.

Schiessel, H., Metzler, R., Blumen, A., and Nonnenmacher, T.F. (1995) "Generalized viscoelastic models: their fractional equations with solutions" *Journal of Physics A: Mathematical and General*, Vol. 28, pp. 6567–6584.

Schiessel, H., and Blumen, A. (1995) "Mesoscopic pictures of the Sol-Gel transition: ladder models and fractal networks" *Macromolecules*, Vol. 28, pp. 4013–4019.

Schwartz,C.W. (2005) "Evaluation of theWitczak dynamicmodulus prediction model" [CD-ROM]. *Presented at 84th Annual Transportation Research Board Meeting*, Washington, DC: Transportation Research Board.

Sharma, R., and Cherayil, B.J. (2010) "Polymer melt dynamics: Microscopic roots of fractional viscoelasticity" *Physical Review E*, Vol. 81 (2), 021804 pp.1–6.

# Disk-shaped Compact Tension Test for Plain Concrete

A. Amirkhanian[1], D. Spring[1], J. Roesler[1], K. Park[1], G. Paulino[1]

[1]Dept. of Civil Eng., Univ. of Illinois at Urbana–Champaign, 205 N. Mathews Ave., Urbana, IL 61801

## *Abstract*

The estimation of concrete fracture properties is essential for an accurate cracking prediction of concrete pavement systems. The single-edge notched beam test has been used to characterize fracture parameters of concrete materials in the laboratory, but obtaining a field specimen with this geometry is not always practical. Currently, a standard exists, ASTM D7313, for the measurement of fracture energy in asphalt concrete using the disk-shaped compact tension (DCT) test. The benefit of this specimen geometry for both concrete and asphalt is that it can easily be fabricated in the laboratory or cored from the field. The total fracture energy ($G_F$) of the material is estimated by using the concept of the work-of-fracture. Additional properties, such as the initial fracture energy ($G_f$) and the critical crack tip opening displacement ($CTOD_C$), can be extracted from the same test through employing compliance measurements and the concept of an equivalent elastic crack model. In this pilot study, the DCT specimen is adopted for concrete materials with small changes to the hole and notch geometry and loading rate of the specimen relative to ASTM D7313. The initial DCT experimental results for concrete containing virgin limestone aggregate and recycled concrete aggregate have been consistent and repeatable. A finite element model (FEM) of the specimen was developed to check the published $K_{IC}$ equation for this geometry and to derive the $CTOD_C$ correction factor. A cohesive zone model was also successfully implemented to simulate the DCT specimen, which verified the validity of the calculated fracture properties from the DCT experiments.

## *Introduction*

Future improvements to mechanistic-based design methods for concrete pavements will require properties that account for the concrete material's crack growth resistance. Current pavement designs utilize the flexural strength of the concrete and fatigue relationships to completely describe the slab's failure. These empirical models do not adequately predict the service life of concrete pavements without significant calibration of the fatigue damage models (Beckett and Humphreys 1989; Falkner and Teutsch 1993; Roesler 1998; Ioannides 2005; Rao and Roesler 2004; Roesler et al. 2005). Recent progress has been made to quantify the flexural capacity of concrete slabs using fracture properties of the concrete in conjunction with a cohesive crack model in a finite element framework (Ioannides et al. 2006, Gaedicke et al. 2009a, Gaedicke 2009b).

The key inputs to the cohesive crack models are fracture and strength properties. The tensile strength can be estimated by the split tensile test or calculated by optimizing the softening curve to the test data. The most common specimen type to measure concrete fracture quantities has been a notched or unnotched three-point beam test (e.g., Jenq and Shah 1985, RILEM 1990, Guinea et al. 1994, Planas et al. 1999, Bažant et al. 2002a, Bažant 2002b, Cusatis and Schauffert 2009). However, because of its geometry and the

geometry of other specimen types (e.g., wedge split test by Brühwiler and Wittmann 1990), the use of these fracture tests are limited to laboratory cast samples. Recent work on the disk-shaped compact tension (DCT) specimen geometry has shown it to give accurate and reliable load-CMOD curves and fracture energy values for asphalt concrete materials at low temperatures (Wagoner et al. 2005, 2006; Kim, H. and Buttlar 2009, Kim, M. and Buttlar 2009, Zofka and Braham 2009). Some research has been done on compact tension (CT) specimens for concrete (Van Mier 1991; Issa et al. 2000a,b; Kumar and Barai 2009) but no studies were found on a DCT specimen. The main advantage of the DCT over the CT specimen geometry is it can easily be extracted from field concrete slabs with a coring machine, which is the primary motivation of this work.

***Research Objective***

The objective of this study is to evaluate the viability of DCT geometry, similar to ASTM D7313, for testing two types of concrete material fracture properties. In this initial work, the exact DCT specimen geometry is taken from Tada et al. (2000) because the $K_I$ and CMOD geometric factors are already available. The second objective is to extract concrete fracture properties and to simulate the fracture response of the two concrete materials (i.e. limestone coarse aggregate (LCA) and a recycled concrete aggregate (RCA) concrete) in the DCT specimen configuration using a finite element-based cohesive zone model.

***Experimental Procedure***

Multiple concrete slabs measuring 2.2m by 2.2m by 15cm thick were cast with either limestone coarse aggregate (LCA) or recycled concrete aggregate (RCA). The mixture proportions for the two types of concrete slabs are shown in Table 1.

Table 1. Concrete mix design for specimens.

| Material | Quantity [kg/m$^3$] |
|---|---|
| Cement | 246 |
| Class C Fly Ash | 61 |
| Coarse Aggregate: Limestone or RCA | 1019 |
| Fine Aggregate: Natural Sand | 774 |
| Water | 129 |
| HRWR | As needed |

After curing for three months, four cores (144 mm diameter) were taken from each LCA and RCA concrete slab and then cut to match the DCT geometry specified in Figure 1. Two DCT specimens could be fabricated from each 15cm thick core. Additional cores were taken to measure the concrete's compressive and tensile strengths (Table 2). The compressive strength of the LCA concrete higher than the RCA concrete but this trend was not seen in the split tensile test results. The elastic modulus was calculated according to the ACI Building Code (ACI 318) as a standard specimen could not be obtained from the slab specimens.

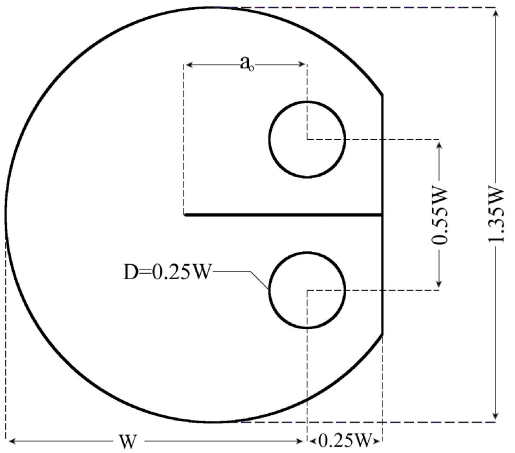
A servo-hydraulic testing frame was used to apply the load and a clip-on strain gage was mounted to control and measure opening deformations at the edge of the specimen as

shown in Figure 2. The specimens were seated with a 0.2 kN load and then loaded with a crack mouth opening displacement (CMOD) rate of 0.1 mm/min. Once the peak load was reached, the specimen was unloaded at 0.2 kN/sec until the seating load was reached. This loading and unloading was completed to calculate the initial fracture properties of the specimen based on the two-parameter fracture model (Jenq and Shah 1985). The specimen was reloaded at the initial rate until the concrete had softened to 0.2 kN.

Table 2. Concrete strength properties.

| Slab Specimen | Compressive Strength [MPa] | Split Tensile Strength [MPa] | Split Tensile Strength COV | Elastic Modulus* [GPa] |
|---|---|---|---|---|
| Limestone | 50.2 | 3.3 | 10.3% | 33.5 |
| RCA | 39.2 | 3.2 | 7.7% | 29.6 |

*Calculated from compressive strength




| Dimension | Value [mm] |
|---|---|
| W | 106.9 |
| D | 26.7 |
| $a_o$ | 44.3 |
| Thickness (B) | 51.0 |

Figure 1. DCT specimen dimensions, adapted from Tada et al. (2000).

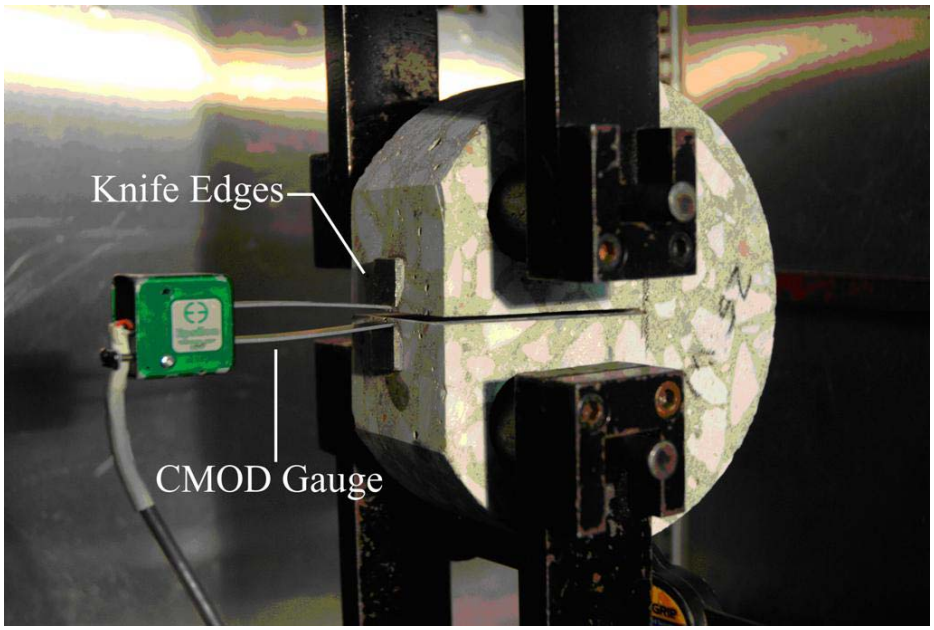



Figure 2. Specimen setup for the DCT test.

The load-CMOD plots for the LCA and RCA specimens are presented in Figures 3 and 4. A summary of the peak loads for the two material types can be found in Table 3 with an average of 1.39 kN (st. dev. = 0.16 kN) and 1.02 kN (st. dev. = 0.06 kN) for the LCA and RCA concrete, respectively. The initial load and CMOD results suggest the DCT specimen was a repeatable test with the average COV less than 15 percent.

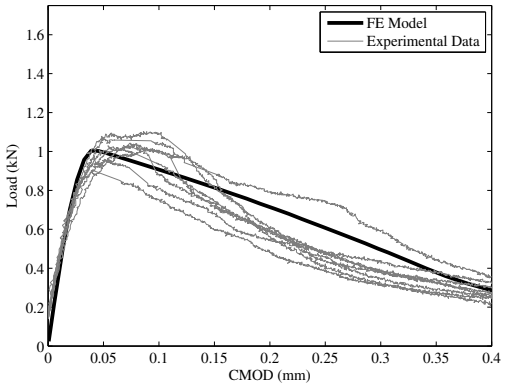



Figure 3. DCT experimental and FE model load versus CMOD results for concrete containing limestone coarse aggregate.

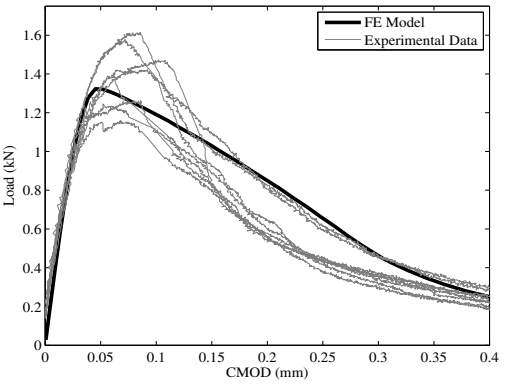



Figure 4. DCT experimental and FE model load versus CMOD results for concrete containing recycled concrete aggregate

***Concrete Fracture Properties***

The finite element modeling of the DCT configuration requires the concrete's fracture properties, such as the critical stress intensity factor ($K_{IC}$), critical crack tip opening displacement ($CTOD_C$), and the total fracture energy ($G_F$). The initial fracture properties ($K_{IC}$ and $CTOD_C$) can be determined from the Two-Parameter Fracture Model (TPFM) by Jenq and Shah (1985). Equations to calculate these quantities are needed to relate the critical crack length at the peak load to the specimen geometry. First, the initial loading concrete stiffness ($E_i$) can be calculated from equation 1.

$$E_i = \frac{WV_{CMOD}(\alpha_o)}{C_i} \tag{1}$$

The initial compliance, $C_i$, was taken to be the inverse slope of the initial P-CMOD curve from the seating load to 50 percent of the peak load. The CMOD geometric factor, $V_{CMOD}(\alpha_o)$, for the DCT geometry (Tada et al. 2000) is given in eq. 2. The initial notch depth ratio, $\alpha_o = a_o/W$ is corrected for the knife edge thickness, $h$, and shown in the general form (eq. 3).

$$V_{CMOD}(\alpha) = \exp\left(1.742 - 0.495\alpha + 14.71\alpha^2 - 22.06\alpha^3 + 14.44\alpha^4\right) \tag{2}$$

$$\alpha = \frac{a+h}{W+h} \tag{3}$$

The unloading compliance, $C_u$, can also be used to calculate the concrete stiffness ($E_u$) after the peak load has been reached. The unloading compliance was experimentally

determined as the inverse slope of the unloading P-CMOD curve from 80 percent of the peak load to 20 percent of the peak load.

$$E_u = \frac{WV_{CMOD}(\alpha_c)}{C_u} \tag{4}$$

In order to find the critical crack length ratio, $\alpha_c$, equations (1) and (4) are set equal, assuming the compliance change is only related to the crack extension. The critical crack length ($a_c$) can then be determined based on the specimen geometry, loading and unloading compliances, and initial notch length. The critical stress intensity factor, $K_{IC}$, at the peak load is then calculated using the geometric factor in Tada et al. (2000), $F(\alpha_c)$, and the applied nominal stress, σ:

$$K_{IC} = \sigma\sqrt{W}F(\alpha_c) \tag{5}$$

$$F(\alpha_c) = \frac{(2+\alpha_c)(0.76 + 4.8\alpha_c - 11.58\alpha_c^2 + 11.43\alpha_c^3 - 4.08\alpha_c^4)}{(1-\alpha_c)^{3/2}} \tag{6}$$

$$\sigma = \frac{P}{WB} \tag{7}$$

Since a crack opening displacement geometric factor for the initial crack tip is not available in the literature, a finite element analysis of the geometry was performed to obtain this geometric factor. The finite element model consisted of eight node quadrilateral (Q8) elements with a radial mesh at the notch tip. Figures 5a and 5b show the global mesh and the biased radial mesh towards the crack tip, respectively. The elements at the crack tip are collapsed Q8 elements with an average length of 0.03 mm and nodes at the quarter points. The quarter point elements enable the model to accurately represent the singular stress field at the crack front. The newly derived equation for $V_{CTOD}(\alpha_c)$ , the CTOD geometric correction factor, is given in eq. 9.

$$CTOD_c = \frac{\sigma W V_{CTOD}(\alpha_c)}{E} \tag{8}$$

$$V_{CTOD}(\alpha_c) = 0.009621 \cdot \exp(35.35\alpha_c - 58.99\alpha_c^2 + 36.54\alpha_c^3) - 11.45 \tag{9}$$

The initial fracture properties are calculated from these equations and are provided in Table 3. The mean $K_{IC}$ is greater for the LCA concrete (st. dev. = 0.119 MPa/m$^{1/2}$) relative to the RCA (st. dev. = 0.110 MPa/m$^{1/2}$) with the $CTOD_C$ values being similar. The total fracture energy, $G_F$, of the two concrete materials was calculated by dividing the area under the load-CMOD curve by the ligament area. The RCA concrete (st. dev. = 9.3 N/m) had slightly lower $G_F$ relative to the LCA concrete (st. dev. = 15.3 N/m), which has been reported previously in SEN(B) tests (Bordelon et al. 2009). The coefficient of variation of all the fracture quantities were well within expected ranges (Bažant and Becq-Giraudon 2002c) and thus the preliminary experimental results suggest the DCT is a viable geometry for determining field fracture parameters. One specimen anomaly that caused some concern was the crack deviation angles. These ranged from 9° to 20° for the limestone specimens and 5° to 30° for the RCA specimens. The likely reason for this was

that the nominal maximum aggregate size was 25 mm, which is about 50 percent of the fracture length. Nevertheless, the deviations did not significantly affect the peak loads or fracture parameters.

### *Computational Crack Growth Modeling*

A nonlinear finite element model was created to simulate the crack growth in the DCT specimens. Previous work had suggested that a cohesive zone model (Borst et al. 2004) with bilinear softening could adequately describe the fracture behavior of plain concrete (Petersson 1981, Gustafsson and Hillerborg 1985, Wittmann et al. 1988, Guinea et al. 1994, and Bažant 2002b, Roesler et al. 2007, Park et al. 2008).

Table 3. Experimental test results for all DCT specimens.

| Specimen | $P_c$ [kN] | $G_F$ [N/m] | $K_{IC}$ [MPa*m$^{1/2}$] | $CTOD_c$ [mm] | Averages |
|---|---|---|---|---|---|
| LCA-1 | 1.389 | 87.4 | 0.916 | 0.006 | |
| LCA-2 | 1.159 | 87.6 | 0.998 | 0.009 | $G_F$ |
| LCA-3 | 1.612 | 123.0 | 1.161 | 0.009 | 100.1 N/m |
| LCA-4 | 1.421 | 98.7 | 1.117 | 0.009 | $K_{IC}$ |
| LCA-5 | 1.244 | 89.4 | 0.945 | 0.008 | 1.090 MPa*m$^{1/2}$ |
| LCA-6 | 1.473 | 100.0 | 1.207 | 0.010 | $CTOD_c$ |
| LCA-7 | 1.262 | 94.4 | 1.179 | 0.011 | 0.009 mm |
| LCA-8 | 1.577 | 121.0 | 1.201 | 0.010 | |
| RCA-1 | 1.029 | 111.0 | 1.011 | 0.012 | |
| RCA-2 | 0.905 | 81.7 | 0.687 | 0.006 | $G_F$ |
| RCA-3 | 1.029 | 103.0 | 0.820 | 0.008 | 92.9 N/m |
| RCA-4 | 1.043 | 86.1 | 0.816 | 0.008 | $K_{IC}$ |
| RCA-5 | 1.008 | 87.9 | 0.807 | 0.008 | 0.816 MPa*m$^{1/2}$ |
| RCA-6 | 0.959 | 96.0 | 0.722 | 0.006 | $CTOD_c$ |
| RCA-7 | 1.064 | 92.0 | 0.750 | 0.006 | 0.008 mm |
| RCA-8 | 1.102 | 85.4 | 0.918 | 0.009 | |

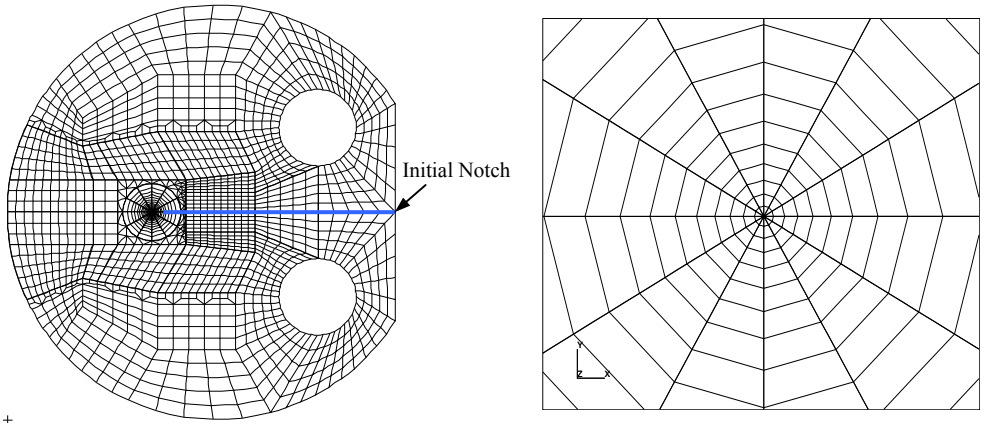



Figure 5. Example DCT finite element mesh: (a) global mesh, (b) biased radial mesh towards crack tip.

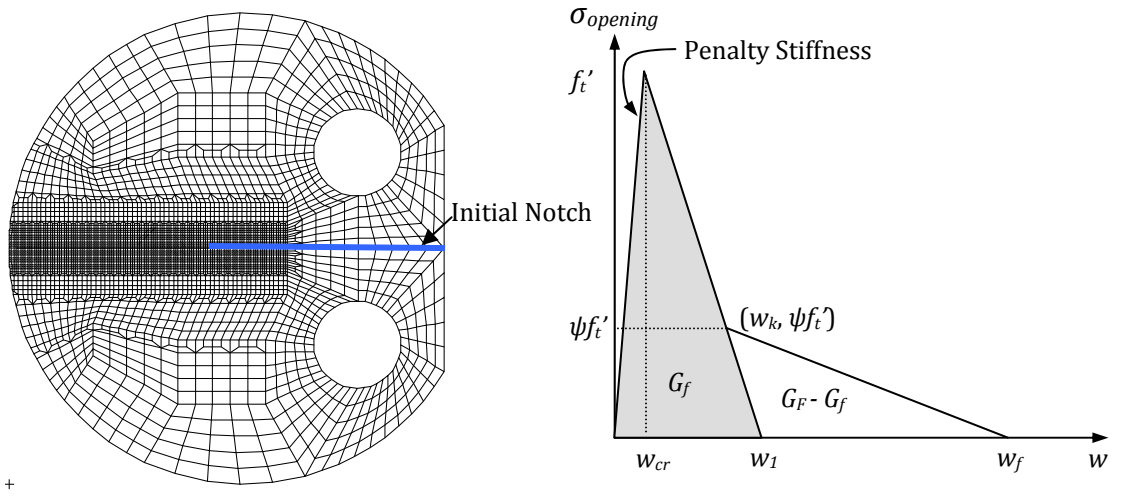



Figure 6. (a) DCT specimen with bulk and cohesive zone finite element mesh. (b) Cohesive stress-opening model with bilinear softening (Park et al., 2008).

*Cohesive Zone Model*

Figure 6a shows the finite element mesh used to simulate the mode I fracture of the DCT specimen. Linear elastic elements (Q4) were used in the model except in front of the crack tip. The cohesive zone model is implemented in commercial software (e.g. ABAQUS) as a user-defined subroutine. Cohesive elements were inserted in front of the initial crack tip and along the expected fracture path of the specimen. In order to model the fracture process accurately, the size of the cohesive crack elements was set to 0.7 mm, which is small enough to capture the nonlinear cohesive zone behavior. The bilinear softening model, shown in Figure 6b, was utilized to idealize the traction-separation relationship in front of the macro-crack tip.

Table 4. Bilinear cohesive model input parameters for limestone and RCA concrete.

| Specimen | $f_t$ [MPa] | $G_f$ [N/m] | $G_F$ [N/m] | CTODc [mm] | $\Psi$ [%] |
|---|---|---|---|---|---|
| Limestone | 3.3 | 26.50 | 100.1 | 0.0091 | 44.3 |
| RCA | 2.5 | 17.55 | 92.9 | 0.0079 | 45.0 |

The bilinear model inputs are the cohesive strength (e.g. tensile strength), $f_t$; the initial and total fracture energies, $G_f$ and $G_F$ respectively; the ratio of the kink point, $\psi$; and a parameter defining penalty stiffness, $p$. The experimentally determined input parameters for each material type are seen in Table 4. The tensile strength for the RCA concrete was lowered based on the additional split tensile strength tests, compressive strength results, and ultrasonic pulse velocity tests suggesting the RCA strength was approximately 25 percent lower than the LCA concrete. The initial fracture energy, $G_f$, was calculated from eq. 10.

$$G_f = \frac{K_{IC}^2}{E} \tag{10}$$

The kink point was defined (Park et al. 2008) at $(w_k, \psi f_t')$ where:

$$w_k = CTOD_c \tag{11}$$

$$\psi = 1 - \frac{CTOD_c f_t'}{2G_f} \tag{12}$$

The penalty stiffness was determined from the ratio of the opening displacement at the peak load, $w_{cr}$, to the final opening displacement, $w_f$, which was chosen to be equal to 0.01.

*Simulation Results*

The results of the finite element (FE) simulations are presented in Figures 3 and 4 for the LCA and RCA concrete. The cohesive zone model produced load versus CMOD curves that fell within the envelope of the experimental curves. There was a minor discrepancy observed between the experimental results and the computational model for the CMOD at the peak load. The experiments on the LCA specimens showed an average peak load of 1.39kN occurring at a CMOD of 0.067mm, while the FE model showed a peak load of 1.32kN occurring at a CMOD of 0.045mm. The experiments on the RCA specimens showed an average peak load of 1.02kN occurring at a CMOD of 0.072mm, while the FE model showed a peak load of 1.00kN occurring at a CMOD of 0.039mm.

## *Conclusions*

Characterizing the mode I fracture properties of in-situ concrete is difficult with most of the current fracture specimen configurations used in the laboratory. Fracture properties are essential to accurately predict cracking behavior and load capacity of any concrete structure including concrete pavements. DCT specimens, 144 mm diameter by 51 mm, were extracted from several 15 cm thick slabs in order to test its viability to represent the fracture properties of recycled concrete aggregate (RCA) and limestone coarse aggregate concrete. A DCT geometry provided in the literature was utilized initially to calculate the fracture properties including the stress intensity factor and the critical crack length. A geometric correction factor for the CTOD was derived based on the finite element analysis of the DCT geometry with several initial crack lengths. The extracted fracture properties were similar to values presented in the literature for recycled and virgin concrete. The initial and total fracture energies were lower for the RCA concrete as expected. The tensile strength for RCA was estimated from the splitting test and ultrasonic pulse velocity measurements. A finite element-based cohesive crack model defined by the measured fracture properties ($G_F$, $G_f$, $CTOD_C$, $f_t'$) was developed to simulate the load-CMOD curve of the two concrete materials. The numerical simulations reasonably matched the experimental data and correctly predicted the lower peak loads for the RCA concrete specimens. The results of this testing and modeling suggest that a DCT-type specimen can potentially provide useful mode I fracture properties of concrete.

***Acknowledgements***

We acknowledge the support from the National Science Foundation (NSF) through Grant CMMI #0800805. The information presented in this paper is the sole opinion of the authors and does not necessarily reflect the views of the sponsoring agency.

**References**

ASTM D7313-07b, 2007. Standard method for determining fracture energy of asphalt-aggregate mixtures using the disk-shaped compact tension geometry, ASTM International.

Bažant, Z. P., Qiang, Y., and Goanseup, Z. (2002a). "Choice of standard fracture test for concrete and its statistical evaluation." *Intl. J. Fracture*, 118, 303-337.

Bažant, Z. P. (2002b). "Concrete fracture models: Testing and practice," *Eng. Fracture Mech.*, 69, 165-205.

Bažant, Z. P., and Becq-Giraudon, E. (2002c). "Statistical prediction of fracture parameters of concrete and implications for choice of testing standard." *Cem. Con. Res.* 32 (4), 529-556.

Beckett, D. and Humphreys, J. (1989). *Comparative Tests on Plain, Fabric Reinforced and Steel Fibre Reinforced Concrete Ground Slabs,* Report No. TP/B/1, Thames Polytechnic School of Civil Engineering, Dartford, 33 pp.

Bordelon, A., Cervantes, V., Roesler, J. (2009). "Fracture properties of concrete containing recycled concrete aggregates." *Magazine of Concrete Research*, 61(9), 665-670.

Borst, R. de, Gutierrez, M.A., Wells, G.N., Remmers, J.J.C., Askes, H. (2004). "Cohesive zone models, higher-order continuum theories and reliability methods for computational failure analysis." *Int. J. for Numerical Methods in Eng.*, 60(1), 289–315.

Brühwiler, E., and Wittmann, F. H. (1990). "The wedge splitting test, a new method of performing stable fracture mechanics tests." *Eng. Fracture Mech.*, 35(1/2/3), 117-125.

Cusatis G., and Schauffert E. A., (2009). "Cohesive crack analysis of size effect." *Eng. Fracture Mech.*, 76, 2163-2173.

Darter, M. I. (1977). *Design of Zero-Maintenance Plain Jointed Concrete Pavement, Volume 1: Development of Design Procedures*. Federal Highway Administration Report No. FHWA-RD-77-III, Washington, DC.

Falkner, H., and Teutsch, M. (1993). *Comparative Investigations of Plain and Steel Fibre Reinforced Industrial Ground Slabs,* Institut Fur für Baustoffe, Massivbau und Brandschutz, Technical University of Brunswick, Germany, No. 102.

Gaedicke, C., Roesler, J. R., Shah, S. P. (2009a). "Fatigue crack growth prediction in concrete slabs." *Intl. J. Fatigue*, 31, 1309-1317.

Gaedicke, C. (2009b). "*Fracture-Based Method to Determine the Flexural Load Capacity of Concrete Slabs,*" Ph.D. thesis, University of Illinois, Urbana, Illinois.

Guinea, G. V., Planas, J., and Elices, M. (1994). "A general bilinear fit for the softening curve of concrete." *Mater. Struct.*, 27(2), 99–105.

Gustafsson, A., and Hillerborg, P. J. (1985). *Improvements in concrete design achieved through application of fracture mechanics*. In: Shah SP, editor. Application of fracture mechanics to cementious composites. The Netherlands: Dordrecht. p. 639–80

Ioannides, A. M. (2005). "Stress Prediction for Cracking of Jointed Plain Concrete Pavements, 1925–2000: An Overview." *Transportation Research Record; Journal of the Transportation Research Board, No. 1919*, Transportation Research Board of the National Academies, Washington, DC, pp.47-53.

Ioannides, A. M., Peng, J. and Swindler, J. R. (2006). "ABAQUS Model for PCC Slab Cracking." *Int. J. Pavement Eng.,* 7(4), 311-321.

Issa, M. A., Islam, M. S., and Chudnovsky, A. (2000a). "Size effects in concrete fracture: Part I, experimental setup and observations." *Int. J. Fracture*, 102, 1-24.

Issa, M. A., Islam, M. S., and Chudnovsky, A. (2000b). "Size effect in concrete fracture – Part II: Analysis of test results." *Int. J. Fracture*, 102, 15-42.

Jenq, Y. S., and Shah, S. P. (1985). "Two-Parameter Fracture Model for Concrete." *J. Eng. Mech.*, 111(4), 1227-1241.

Kim, H., and Buttlar, W. G. (2009). "Finite element cohesive fracture modeling of airport pavements at low temperatures." *Cold Regions Science and Technology*, 57, 123-130.

Kim, M., Buttlar, W. G., Baek, J., and Al-Qadi, I. L. (2009). "Field and laboratory evaluation of fracture resistance of Illinois hot-mix asphalt overlay mixtures," *Transportation Research Record; Journal of the Transportation Research Board, No. 2127*, Transportation Research Board of the National Academies, Washington, DC, pp. 146-154.

Kumar, S., and Barai, S. V. (2009). "Determining double-K fracture parameters of concrete for compact tension and wedge splitting tests using weight function." *Eng. Fracture Mech.*, 76, 935-948.

Park, K., Paulino, G. H., and Roesler, J. R. (2008). "Determination of the kink point in the bilinear softening model for concrete." *Eng. Fracture Mech.*, 75, 3806-3818.

Petersson, P. E. (1981). *Crack growth and development of fracture zone in plain concrete and similar materials.* Report No. TVBM-1006, Division of building materials, Lund Institute of Technology, Lund, Sweden.

Planas, J., Guinea, G. V., and Elices, M., 1999, "Size Effect and Inverse Analysis in Concrete Fracture." *Intl. J. Fracture*, 95, 367-378.

Rao, S., and Roesler, J. R. (2004). "Cumulative Fatigue Damage Analysis of Concrete Pavement using Accelerated Pavement Testing Results." Second International Conference on Accelerated Pavement Testing, September 25-29, 2004, Minneapolis, Minnesota.

RILEM, 1990, "Determination of Fracture Parameters ($K_{IC}$ and $CTOD_C$) of Plain Concrete using Three-Point Bend Tests," RILEM Committee on Fracture Mechanics of Concrete-Test Methods, *Mat. Struct.*, 23, 457-460.

Roesler, J. R. (1998), *"Fatigue of Concrete Beams and Slabs*," Ph.D. thesis, University of Illinois, Urbana, IL.

Roesler, J. R., Hiller, J. E., and Littleton, P. C. (2005), "Large-Scale Airfield Concrete Slab Fatigue Tests," 8th International Conference on Concrete Pavement, August 13-18, 2005, Colorado Springs, CO, 23 pp.

Roesler, J. R., Paulino, G. H., Park, K. and Gaedicke, C. (2007). "Concrete fracture prediction using bilinear softening." *Cem. Con. Comp.*, 29, 300–312.

Song, S. H., Paulino, G. H., and Buttlar, W. B (2006). "Simulation of Crack Propagation in Asphalt Concrete Using a Cohesive Zone Model." *J. Eng. Mech.*, 123(11), 1215-1223.

Tada, H., Paris, P. C., and Irwin, G. R. (2000). The stress analysis of cracks handbook /New York: ASME Press

Van Mier, J. G. M. (1991). "Mode I Fracture Of Concrete: Discontinuous Crack Growth And Crack Interface Grain Bridging." *Cem. Conc. Res.*, 21(1), 1-15.

Wagoner, M. P., Buttlar, W. G., and Paulino, G. H. (2005). "Disk-shaped compact tension test for asphalt specimens." *Exp. Mech.*, 45(3), 270-277.

Wagoner, M. P., Buttlar, W. G. , Paulino, G. H., and Blankenship, P. I. (2006). "Laboratory Testing Suite for Characterization of Asphalt Concrete Mixtures Obtained from Field Cores." *J. Assoc. Asphalt Paving Technologists*, 75, 815-852.

Wittmann, F. H., Rokugo, K., Brühwiler, E., Mihashi, H., Simopnin, P. (1988). "Fracture energy and strain softening of concrete as determined by compact tension specimens." *Mater. Struct,*, 21(1), 21–32.

Zofka, A., and Braham, A. (2009). "Comparison of low-temperature field performance and laboratory testing of 10 test sections in the Midwestern United States," *Transportation Research Record; Journal of the Transportation Research Board, No. 2127*, Transportation Research Board of the National Academies, Washington, DC, pp. 107-114.

# Influence of alkali-silica reactions on cracking in concrete pavements

R. Breitenbücher[1] and C. Sievering[2]

[1]Instiute for Building Materials, Department for Civil and Environmental Engineering, Ruhr-University Bochum, Universitätsstr. 150, 44801 Bochum, Germany; PH (+49) 234-3222649; FAX (0049) 234-3214113; email: rolf.breitenbuecher@rub.de
[1]Instiute for Building Materials, Department for Civil and Environmental Engineering, Ruhr-University Bochum, Universitätsstr. 150, 44801 Bochum, Germany; PH (+49) 234-3222671; FAX (0049) 234-3214113; email: christoph.sievering@rub.de

## ABSTRACT

In the last few years cracking in concrete pavements was observed in several highway sections in some regions of Germany. Within the scope of extensive investigations no definite single cause of cracking could be observed in the majority of these cases. Besides incremental load-induced stresses by increasing heavy traffic, within this scope especially load-independent stresses due to hygral and thermal changes are of relevance. Several investigations additionally substantiate reaction products of an alkali silica reaction (ASR). Because of this multiplicity of potential causes for cracking it has not been definitely clarified up to now to which extent especially the ASR contributes to cracking in concrete pavements. Rather it seems that superposition and/or interactions of different mechanisms are responsible.

## INTRODUCTION

In current discussions cracking in concrete pavements is often associated with an alkali silica reaction (ASR). However, even if relevant reaction products were detected in appropriate samples, cracks result in the rarest cases from an ASR solely. Rather it has to be assumed, that cracks were caused by a superposition of several stress impacts. Besides restraint stresses due to disabled thermal and hygral self-deformations, which mainly are raised by restraining of warping concrete pavements mainly are also subjected to traffic loads.

However, the influence of the alkali silica reaction in concrete pavements cannot be neglected. In such constructions especially the supply of alkalis from external sources has a substantial importance. Alkaline de-icing agents applied in the winter period penetrate more or less intensively into the concrete structure, which leads to a continuous increase of the alkali potential. In case of high traffic volume this penetration process is intensified by the following vehicles. The intrusion of the alkalis furthermore is particularly forwarded by already existing microcracks. Due to these complicated interactions in concrete pavements more extensive measures are necessary to avoid damaging alkali silica reactions in new concrete pavements than in other conventional concrete constructions.

## CRACKING IN CONCRETE PAVEMENTS

### Differentiation of crack types

The durability of concrete pavements is strongly impaired by cracks. Therefore a comparatively high flexural strength of at least 5.5 N/mm² is generally required in plain concrete slabs to minimize the risk of cracking. Nevertheless cracking in concrete

pavements cannot be avoided completely. However, not each crack results generally in an impairment of serviceability or leads to a hazard. In this context it should be considered that only 1 to 2 percent of the total 3,600 kilometres of concrete pavements are really damaged in Germany nationwide.

For the characterization of cracks in concrete pavements it must be differentiated between various formations. In particular a differentiation between map cracking (Fig. 1) and clearly recognizable longitudinal cracks (Fig. 2) is essential.

Figure 1. Map cracking in a concrete pavement

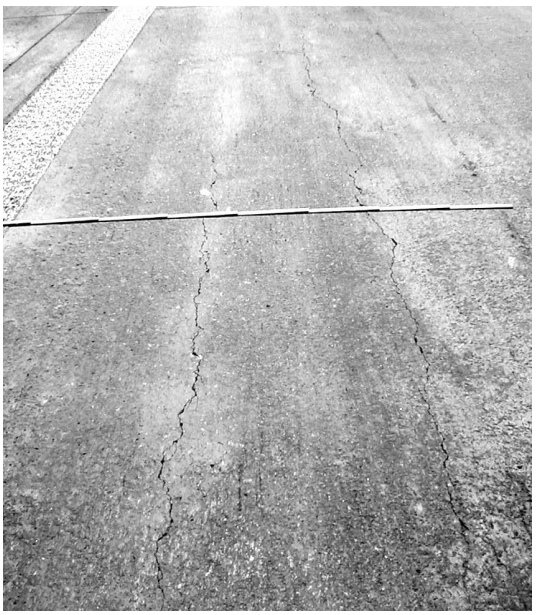


Figure 2. Longitudinal cracks in a concrete pavement

As long as the crack width remains sufficiently small (crack width < 0.3 - 0.4 mm) the serviceability of the pavement usually is not impaired. A critical stage has to be stated, when the edges of the cracks break out , particularly in the surroundings of joints, due to the over-rolling traffic. In this way the crack widths are expanding continuously by themselves which finally will result in a complete destruction of the material in this area. If larger particles are bursted out, the following traffic will be seriously endangered. By this the general conditions of the concrete pavement will be drastically worsened.

In order to be able to define adequate technological measures for avoiding of such critical cracking, the most significant causes have to be identified before. In this context it should be considered that cracks in concrete pavements in the rarest cases are raised only by one reason, in the majority a superposition or rather interactions between different mechanisms are responsible (Breitenbücher 2006b).

## Restraint and residual stresses

Due to the endless extension in longitudinal direction, deformations in pavements are practically completely restrained. In case of non-load-induced deformations, e.g. example by thermal or hygral changes, which extent constant over the complete cross-section, longitudinal restraint stresses are generated. Although in transverse direction a movement is enabled to some extent, restraint stresses also cannot be excluded completely in this case.

If a thermal or a hygral gradient yield over the slab thickness, which is the standard case, the slab would tend to warp and curl. These deformations are restrained by the dowels and tie bars resp. as well as by the dead load of the slab to a large extent. In addition the slabs are pressed down by traffic loads. In consequence appropriate flexural stresses with the tensile zone on the "cold" and/or „dry" side are generated.

While due to longitudinal restraint stresses transversal cracks with nearly constant crack width are formed, flexural stresses result in wedge-shaped cracks with the opening on the dry or cold side. Furthermore residual stresses due to non-linear distributions of deformations are generated, which can result in map cracking.

Longitudinal cracks (fig. 2) often can be observed in the lanes of the vibrators of the slip form paver. However, the latter are only responsible for the location of these cracks (influenced by an accumulation of mortar in the lanes) and not for cracking itself.

## Impact of traffic on microcracking

Within a case study of cracking in concrete pavements both load-independent influences (thermal and hygral) as well as load-dependent influences (i.e. traffic) have to be considered. Even if the permissible axle loads have not significantly increased in the past decades in most countries, the volume of the overall traffic and in particular of the heavy traffic has drastically grown in this period. Between 1970 and 2005 the transported cargo volume in Germany increased from 80 billion ton kilometers up to 400 billion ton kilometers, i.e. by five times. Besides the static loads especially the dynamic / cyclic loads can affect the microstructure of the concrete.

In appropriate tests the behavior of the concrete under cyclic loadings was investigated. It could be proved by a reduction in the dynamic E modulus, that already at upper stress levels between about 45 and 70 % of the concrete strength degradation processes take place in the microstructure far before fatigue failure is exceeded (Fig. 3) This degradation process is attended by the formation of microcracks in the concrete microstructure (Fig. 5) (Breitenbücher et al. 2006a).

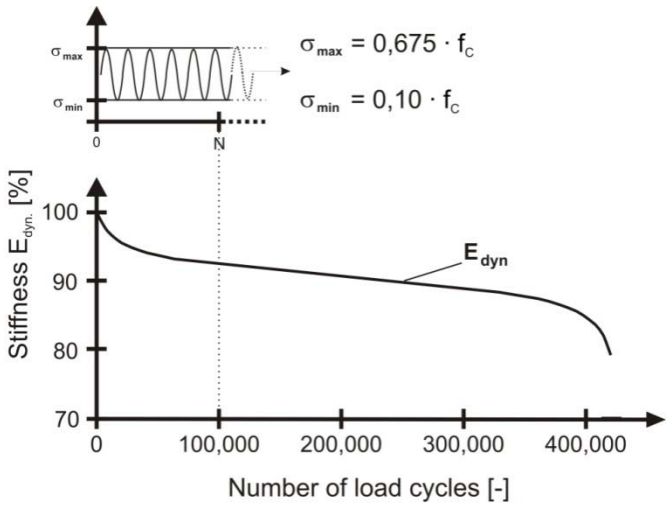



Figure 3. Decreasing stiffness of concrete under cyclic load (Breitenbücher et al. 2006a)

Such degradation processes can also be detected in concrete pavements by comparing the dynamic E-modulus between various lanes. In appropriate tests ultrasonic velocities of Rayleigh waves were measured by the Institute for Building Materials of Ruhr-University Bochum at shoulder, near side lane and passing lane of various German highways. From these US-velocities

then the dynamic E modulus of each lane was calculated. Figure 4 shows relative dynamic E-moduli of five highway sections, in which the dynamic E moduli of the near side lane and passing lane was related to the E modulus of the shoulder (same concrete in all lanes).

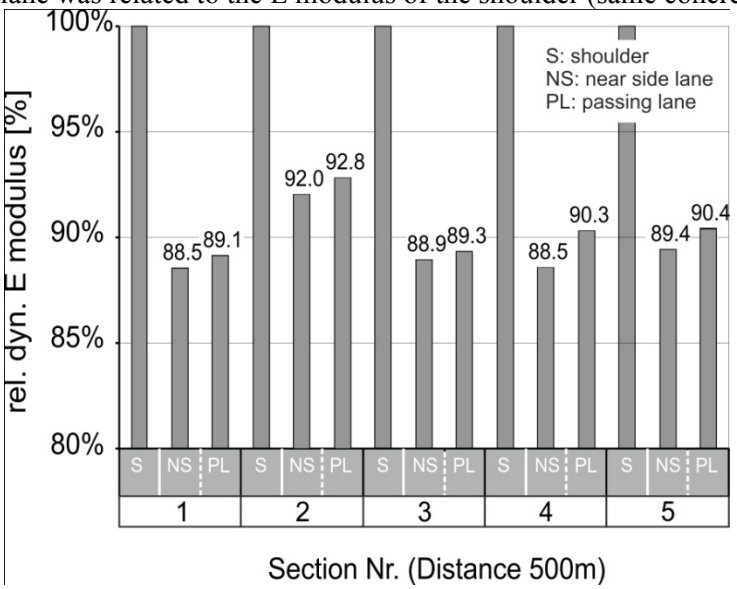



Figure 4. Relative dynamic E moduli of different highway lanes / sections

It demonstrates that the relative dynamic E moduli of the traffic lanes are approximately 10 percent lower than the E modulus of the shoulder, which is assumed to be nearly unloaded within the investigated period and thus was used as reference-basis. This loss of stiffness can be related to cyclic loading due to traffic.

The changes in mechanical properties indicate that degradation processes were also proceeding in the microstructure. However, such degradations are the basis of microcracking, which can be shown by increasing microcracks due to increasing load cycles N (Fig. 5). Due to such increasing mircocracks the penetration of moisture and other liquids into the concrete structure is enhanced to a large extent.

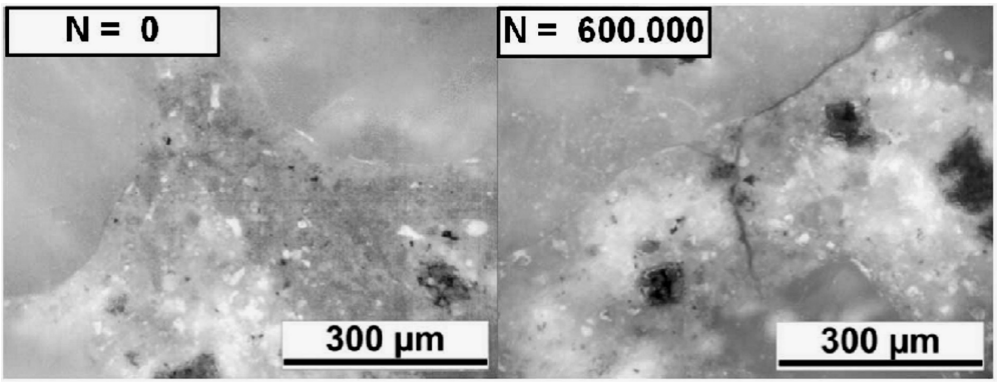



Figure 5. Increasing microcracking due to cyclic loading (Breitenbücher et al. 2006a)

## Expansion due to Alkali-Silica-Reaction

In the well-known alkali silica reaction amorphous silica ($SiO_2$) reacts with alkali hydroxide ions (NaOH, KOH), when simultaneously an appropriate amount of humidity is available. The reaction product is an alkali silica gel, which expands dramatically due to absorption of water (Fig. 6).

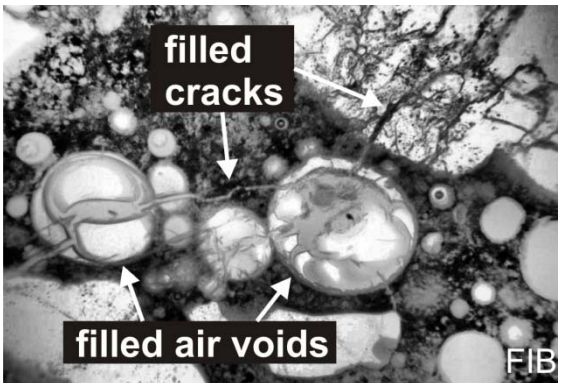



Figure 6. Voids and cracks in the matrix / aggregates filled with ASR products

The expansion caused by interior swelling is restrained by the surrounding concrete to a large extend. Thus an internal pressure occurs, which can accumulate up to 20 MPa (Stark et al. 2001). Such stresses can easily exceed the tensile strength of the concrete with the consequence of internal cracking. Cracks due to an ASR are usually map- distributed and netlike. They are not limited only to the visible surface area but extend over the complete concrete structure.

In the last few years cracking in concrete pavements often is associated with alkali silica reactions (ASR) (Breitenbücher 2006b). However, even if relevant ASR-products are detected in appropriate samples, cracks result only in the rarest cases from that. In this context it should be considered, that ASR is the only cracking cause, which can be identified later by their reaction products. Other stresses, e.g. thermal and hygral influences or traffic loads do not leave their marks subsequently.

## Alkalis from external sources

The production of the swellable gel in the ASR is mainly influenced by solved alkalis in the pore solution of the concrete. In the case of concrete pavements not only internal sources but also external sources should be considered. As shown in Figure 9b below the external alkali supply / de-icing agents) plays a decisive role to the extent of an ASR in concrete pavements. In Figure 7 the amount of applied de-icing agents on German highways is illustrated.

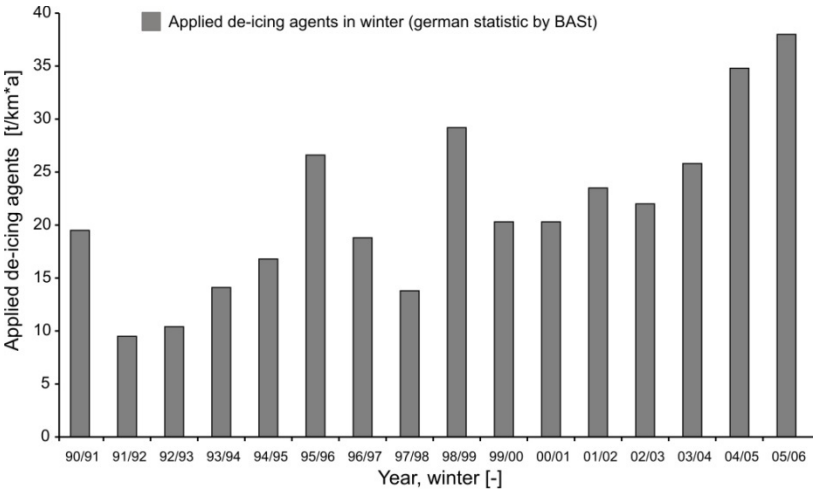



Figure 7. Applied de-icing agents in winter in Germany

It demonstrates that the amount of applied de-icing agents has successively increased in the last decade, which was caused by more and more prophylactic application in this period. However, due to the prophylactic application of sodium chloride nowadays this salt is pressed quite more intensive into the concrete by the over rolling traffic and hence forces the ASR subsequently. The intrusion of the alkalis is especially forwarded by the already existent cracks and microcracks respectively.

## INVESTIGATIONS ON CRACKED PAVEMENTS BY ANALYIS OF CONSTRUCTION DOCUMENTS AND ADDITIONAL TEST

### Thermal Influences on cracking in concrete pavements

Within the scope of extensive investigations highway sections (with and without cracking, map cracking as well as longitudinal cracks) were selected for studies referring to the causes for the cracking. In order to quantify these cracks, typical characteristics, e.g. the average quantity of cracks per slab, the average width of cracks per slab und the average cracking opening area, which is defined by the product of the total crack length and width per slab were determined.

In addition to the evaluation of the crack formation all available construction documents, i.e. daily construction records, reports on initial tests and quality control tests, weather records etc. were compiled and evaluated. In this context also the air temperatures during the concrete casting were compiled (Fig. 8).

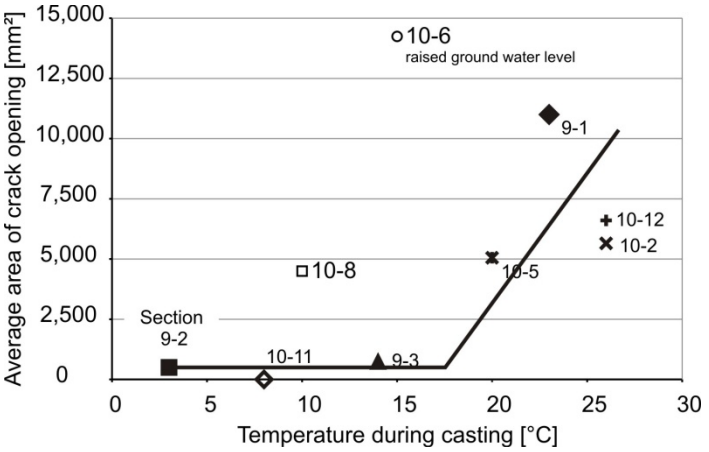



Figure 8. Influence of the air temperature at concrete casting on crack opening area

In these analyses a significant influence of the temperature at concrete casting on the cracking could be proved in general (Fig. 8), whereat it can be assumed that the concrete temperature correlates more or less with the air temperature at casting. Except two subsections (10-6 and 10-8), which were additionally damaged by intensive hygral effects (rising ground–water), it could be verified, that nearly no cracking was observed in the concrete pavements after an average age of 14 years, when the temperature during casting was below 15°C to 20°C. Exceeding this temperature range, i.e. concreting in the summer season, cracking increases progressively.

### Specific ASR- Tests on Concrete Cores at 60°C – with and without external alkali impact

To evaluate the ASR-potential within the inspected fields, specimens (cores) were sampled, on which the residual potential of ASR was determined by specific ASR-tests in the lab. To obtain information as soon as possible, the tests were performed in special climates at 60°C to accelerate the ASR. In this context also the influence of an external alkali supply was considered. So for this purpose the cores were divided into two parts. One subspecimen was stored constantly at 60°C above water according to the "Alkali-Guideline (2006)" (DAfStB 2006), which is very similar to the "RILEM Recommended Test Method TC 191-ARP AAR-4" (Rilem 2000). The other subspecimen was exposed to following cyclic procedure: 6 days above water at 60°C, 1 day at 20°C, 5 days at 60°C in a dry air and 2 days immerged in a sodium chloride solution with a concentration of 0.6 mol/l. With this procedure the test specimens were loaded for 16 cycles. During the complete test phase the deformations of the specimens were determined under both storing conditions.

Even when no ASR takes place deformations to some extent will occur in the samples only due to thermal and hygral effects within the test procedure. In the tests without external alkali supply these can lead to an expansion up to about 0.3 to 0.4 mm/m, in case of an external alkali supply they can increase up to 0.5 mm/m. Thus, only deformations exceeding these values may be linked with an ASR and can be considered within the evaluation.

In Figure 9 the results of the deformation measurements without (Fig. 9a) and with (Fig. 9b) external alkali supply are documented.

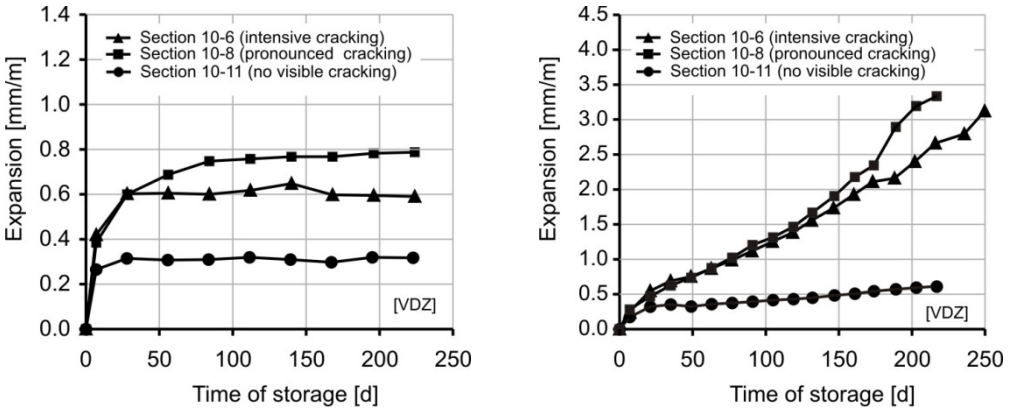



Figure 9. Expansion of drilled cores due to storage at 60°C
a) left: without external alkali supply
b) right: with external alkali supply

The comparison of Figure 9a and Figure 9b demonstrates that the expansions increased significantly when alkalis are additionally supplied by external sources. This has been known also from other investigations (Stark et al. 2006). Furthermore it could be proved, that the degree of expansion mainly is impacted by the degree of degradation / microcracking within the concrete microstructure.

## LABORATORY INVESTIGATIONS ON ALKALI-PENETRATION

In order to determine the impact of the degree of degradation in the concrete on the penetration of de-icing agents (sodium chloride) into the concrete structure, further laboratory investigations were carried out. Therefore beams (140 x 40 x 27 cm³) made of typical pavement concrete were cyclically loaded. The load applied simulated a superposition of thermal caused stresses and stresses due to traffic loads. This means that the maximum stress $\sigma_{max}$ corresponded to thermal restraint stresses, which appear due to casting the concrete pavement in summer months. The amplitude of the cyclical load ($\sigma_{max}$ - $\sigma_{min}$) characterizes the traffic load, which appears due to an over rolling truck with an axle load of 10 tons. Based on these configurations the concrete beams were cyclically loaded up to 5 million load cycles N with a frequency of 7 Hz in order to accelerate the test.

After various load cycles the dynamic E modulus was determined by measuring the ultrasonic speed of the Rayleigh wave at the tensile zone of the concrete beams. All dynamic E modulus were furthermore related to the E modulus of the beams before the first load was applied. On this basis a relative dynamic E modulus was determined in order to describe the deterioration of the concrete beam due to cyclic loading (Fig. 10).

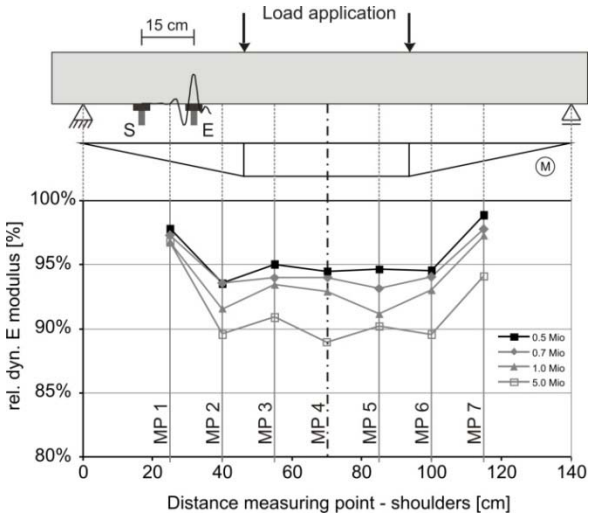



Figure 10. Decrease of the relative dynamic E modulus due to cyclic loading

Figure 10 demonstrates that the relative dynamic E modulus decreases continuously due to the applied cyclic loading. So the relative dynamic E Modulus decreases up to 90 percent after 5 million load cycles for instance, which can be related to an adequate degradation of the concrete microstructure.

In order to determine the influence of such a degradation on the penetration of a sodium chloride solution (3.0 w % sodium chloride) further experiments were conducted. For this purpose a sodium chloride solution was pressed into the tensile zone of variously pre-damaged beams (0, 2 million and 5 million load cycles N) by tire passes (TP), whereat the wheel was additionally loaded with a weight of 1 ton (Fig. 11).

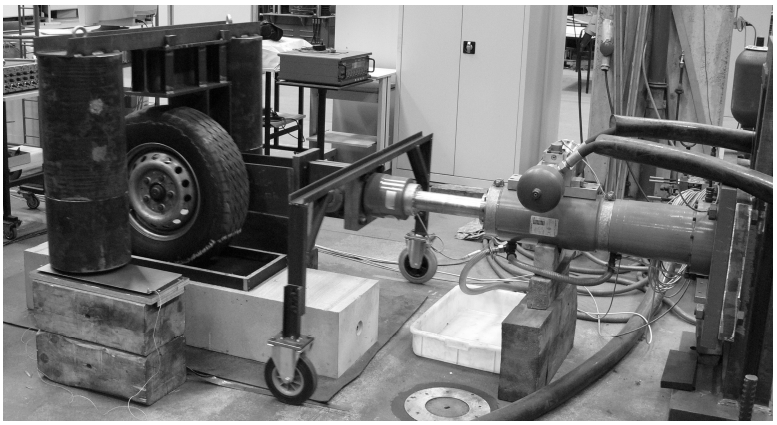


Figure 11. Penetration-test of sodium chloride solution due to over rolling movements of a tire

After 1 million tire passesconcrete specimens (40 x 10 x 10 cm³) were cut transversally out of the beam. These specimens then were split in longitudinal direction (transversal axis of the beam) thereupon. Afterwards silver nitrate was sprayed on the fracture surface to detect the depth of chloride penetration by a color change (cd = 0.07 mol / l) (Fig. 12).

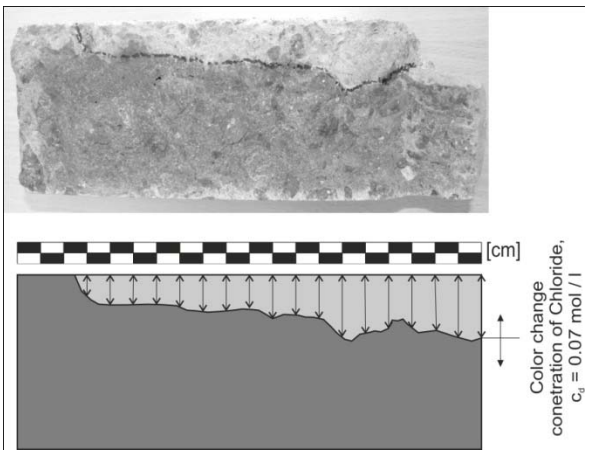



Figure 12. Depth of color change due to spayed silver nitrate

Even if the silver nitrate test identifies only the presence of chlorides – the presence of sodium, which mainly influences the ASR, cannot be identified by such rapid tests – , this test is suitable for scanning large concrete specimens in a very short time as it is assumed, that the penetration of chloride and sodium are similar or even in line. At this point it should be mentioned either, that the effect of chromatography was disregarded so far.

However, after measuring the penetration depth in steps of 1 cm a two-dimensional mapping was generated. In Figure 13 three penetration maps are shown, which differ in the applied load cycles (N) before the tire passes (TP) started.

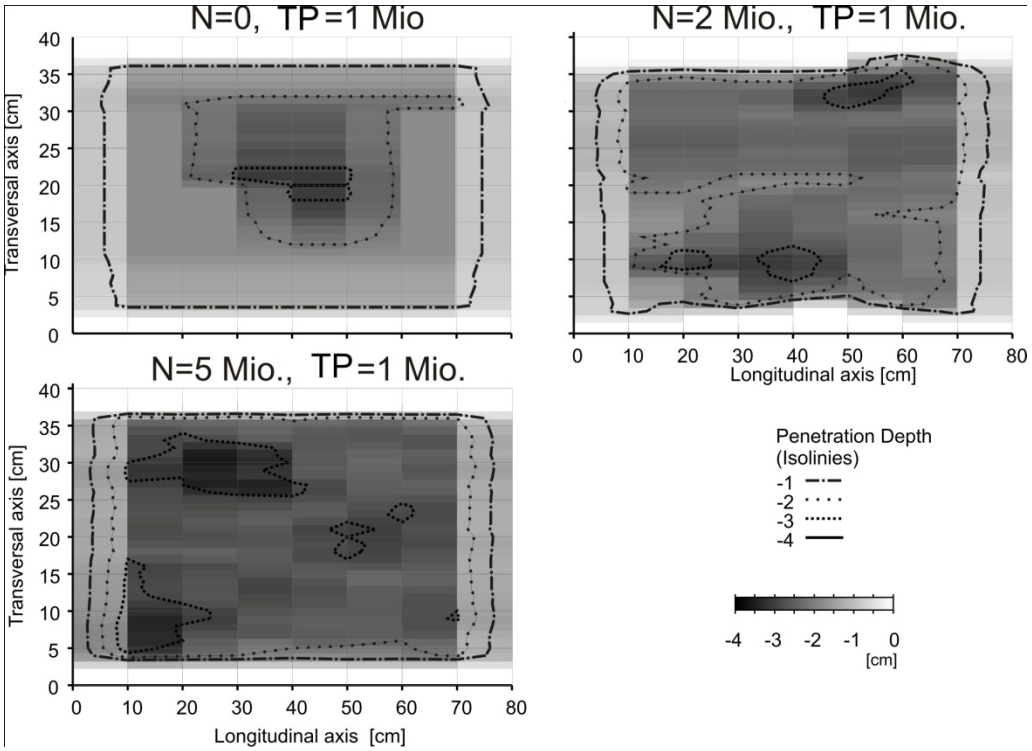



Figure 13. Penetration depth of the sodium chloride solution

If the concrete beams were not damaged before tire passes started, the sodium chloride penetrates 19 mm (mean value) into the concrete due to 1 million tire passes. Because of the increasing load cycles before the tire passes the penetration depth of the sodium chloride increases as well (22 mm after 2 million load cycles and 28 mm after 5 million load cycles). This means that the penetration depth of the sodium chloride increases noticeably by increasing degradation within the concrete microstructure.

## FINAL DISCUSSION AND CONCLUSION

Because of the multiplicity of potential causes for pavement cracking all possible influences have to be considered in order to reveal the relevant causes. In the minority of the investigated sections only one single reason could be found as responsible for the crack formation.

As presented in Figure 5, the temperature at concrete casting has significant influence on the initial cracking of concrete pavements. With increasing pouring temperature an increase of the average area of crack opening could be observed in the majority of the investigated highway sections. As a result of the interpretation of construction documents it can be assumed that the initial cracking in concrete pavements basically is influenced to a larger extent by thermal and hygral circumstances.

If microcracks due to cyclic traffic loads in combination with restraint stresses have already been formed in a concrete pavement, moisture as well as external supplied alkalis can easily penetrate deeper into the concrete microstructure. As demonstrated in Figure 9 in consequence such pre-existing degradations or damages play a decisive role for the further development of ASR-expansions, provided that simultaneously adequate portions of reactive aggregates are present in the concrete.

If the concrete has not been impaired in any way before, the impact of external alkalis plays only a minor role for the further ASR-expansion. Thus the extent of

premature degradations, e.g. raised by thermal / hygral effects and traffic loads, has been proved to be the most important impact for the formation of an ASR in concrete pavements in cases of external alkali supply.

## ACKNOWLEDMENTS

The authors are grateful to the Federal Ministry of Transport, Building and Urban Affairs (Germany) for financial support of this research project. The authors wish to acknowledge the cooperation with the Finger-Institute for Building Materials Science and the Research Institute of the German Cement Industry whose laboratory ASR tests were an important part of this project.

## REFERENCES

Breitenbücher, R. & Ibuk, H. 2006a. Experimentally based investigations on the degradation-process of concrete under cyclic load. *Materials & Structures* 39: 717–724.

Breitenbücher, R. 2006b: Potentielle Ursachen der Rissbildung in Betonfahrbahndecken. Proceedings of the 16. ibausil, Weimar, Germany, Sept. 20-23 2006, Vol. 1: 1239–1254.

DAfStB 2006. Vorbeugende Maßnahmen gegen schädigende Alkalireaktion im Beton (Alkali-Richtlinie). Deutscher Ausschuss für Stahlbeton (ed.), Berlin: Beuth.

Rilem 2000. Detection of Potential Alkali-Reactivity– Accelerated method for testing aggregate combinations using concrete prisms. Recommended Test Method TC 191-ARP AAR-4. *Materials and Structures 33:* 290-293.

Stark, J., Freyburg, E., Seyfarth, K. & Giebson, C. 2006. AKR-Prüfverfahren zur Beurteilung von Gesteinskörnungen und projektspezifischen Betonen. *Beton 12*: 574-581

# Comparative Performance of Concrete Pavements with Recycled Concrete Aggregate (RCA) and Virgin Aggregate Subbases

Sunghwan Kim[1], Halil Ceylan[2], Kasthurirangan Gopalakrishnan[3], David J. White[4], Charles T. Jahren[5], and Thang Huu Phan[6]

[1]Postdoctoral Research Associate, Department of Civil, Construction and Environmental Engineering (CCEE), Iowa State University, Ames, IA 50011-3232; sunghwan@iastate.edu
[2]Assistant Professor, Department of CCEE, Iowa State University, Ames, IA 5011-3232; hceylan@iastate.edu
[3]Research Assistant Professor, Department of CCEE, Iowa State University, Ames, IA 50011-3232; rangan@iastate.edu
[4]Associate Professor, Department of CCEE, Iowa State University, Ames, IA 5011-3232; djwhite@iastate.edu
[5]Associate Professor, Department of CCEE, Iowa State University, Ames, IA 5011-3232; cjahren@iastate.edu
[6]Research Assistant, Department of CCEE, Iowa State University, Ames, IA 5011-3232; thphan@iastate.edu

## ABSTRACT

Using recycled concrete aggregate (RCA) as base/subbase for road construction is alternative way to manage construction waste and reduces the need for virgin aggregates. This paper discusses the long-term performance of concrete pavements on RCA granular layer in comparison to concrete pavements on virgin aggregate granular layer. 18 representative RCA pavement sections across Iowa were primary selected considering state wide location and pavement age. Detailed visual distress surveys were conducted to identify the current pavement surface condition information and the primary distress on RCA granular layer pavement. The pavement information data for the surveyed RCA subbase sections and the corresponding virgin aggregate subbase sections in similar condition were extracted from the Iowa DOT's Pavement Management Information System (PMIS). The performance of the RCA and the virgin aggregate sections was investigated and compared in terms of the pavement condition index (PCI) and the International Roughness Index (IRI). The results indicate that RCA granular layer provides performance compared to the virgin aggregate granular layer in Iowa pavements and are performing adequately.

## INTRODUCTION

Waste construction materials from rehabilitation and reconstruction have become global concerns. Historically, the most common method of managing this material has been through disposal in landfills. However, the use of these materials in

landfills becomes more restrictive and the need to seek alternative uses of the waste material increases (FHWA, 2004).

Using recycled aggregate in pavement construction can reduce the need for virgin or natural aggregates, preserves the environment, and saves landfill space. In spring of 2002, Federal Highway Administration (FHWA) released a memorandum emphasizing the FHWA interest in using recycled material products in the national highway system (FHWA, 2002).

At present, recycled concrete aggregate (RCA) is mostly used in pavement granular layer, shoulder concrete, and porous granular fill, and, to a limited extent, as aggregates in new concrete pavement (Chini et al, 2001). Approximately 68 % of the RCA produced in the US is as a granular layer of base/subbase under pavements (Wilburn and Goonan, 1998). The role of pavement bases and sub-bases is to provide uniform support of pavement surface layers and adequate drainage during the lifetime of the pavement (Huang, 2003). To meet these requirements, the materials used in unbound pavement layers must meet specific particle size distributions with adequate stiffness, good durability (e.g. freeze-thaw performance), high permeability, and resistance to permanent deformation (e.g. particle crushing).

Several studies have been conducted to understand the RCA properties and specify for their use in unbound pavement layer (Miyagawa, 1991; Snyder, 1995; Chini et al., 1998; Chini et al, 2001; White et al., 2004). However, most of the previous research doses not address the long-term field performance of concrete pavement with RCA granular layer regarding engineering properties, which is required to develop proper guidelines and mechanistic-empirical design procedures.

The objective of this study is to evaluate the long-term performance of RCA granular layer pavements. For this purpose, 18 representative RCA pavement sections were selected considering state wide location and pavement age. Detailed visual distress surveys were conducted to identify the current pavement surface condition information and the most common distresses on concrete pavements that are supported with a RCA granular layer. The pavement information data for the 10 of surveyed RCA subbase sections and the corresponding virgin aggregate subbase sections in similar condition were extracted from the Iowa DOT's Pavement Management Information System (PMIS). The performance of the RCA and the virgin aggregate sections was compared in terms of the pavement condition index (PCI) and the International Roughness Index (IRI). The procedure and the results of the analysis are discussed in this paper highlighting the important findings regarding the long-term performance of concrete pavements that are supported by a RCA granular layer.

## PAVEMENT PERFORMANCE DATA COLLECTION

A field experiment was carried out in 2007 to assess the long-term performance of concrete pavement with RCA granular layer. 18 representative RCA pavement sections listed in Table 1 were primary selected considering state wide location and pavement age in Iowa. The ages of most pavements selected are at least older than 10 years. The detailed current pavement performance data for the selected

sections in this study were collected through the visual distress survey. The exiting pavement performance information was also extracted from the Iowa DOT's PMIS.

**Table 1. RCA granular layer concrete pavement sections evaluated**

| Section No. | Location | | | PCC Thick (mm) | Subbase Thick (mm) | ADTT[a] | Construction Year |
|---|---|---|---|---|---|---|---|
| | County | Mile Post No. | Dir. | | | | |
| RCA/1 | I-35 in Story | 119.95 to 120.05 | S | 305 | 152 | 5,074 | 1999 |
| RCA/2 | I-35 in Hamilton | 140.75 to 140.80 | N | 267 | 203 | 4,657 | 2003 |
| RCA/3 | I-80 in Jasper | 165.0 to 165.05 | E | 330 | 127 | 8,883 | 1996 |
| RCA/4 | I-80 in Jasper | 165.20 to 165.25 | E | 330 | 178 | 8,870 | 1994 |
| RCA/5 | I-35 in Hamilton | 131.40 to 131.45 | N | 267 | 203 | 4,709 | 1983 |
| RCA/6 | I-80 in Pottawattamie | 10.55 to 10.60 | W | 330 | 279 | 5,421 | 1999 |
| RCA/7 | I-80 in Pottawattamie | 10.55 to 10.65 | E | 305 | N/A[b] | 5,421 | 2003 |
| RCA/8 | I-80 in Cass | 65.10 to 65.20 | E | 305 | N/A | 7,506 | 1988 |
| RCA/9 | I-80 in Cass | 65.80 to 65.90 | W | 305 | N/A | 7,506 | 1987 |
| RCA/10 | I-80 in Polk | 128.50 to 128.55 (Travel) | E | 356 | 229 | 13,322 | 1994 |
| RCA/11 | I-80 in Cedar | 269.00 to 269.10 | E | 330 | 152 | 10,777 | 1991 |
| RCA/12 | I-80 in Cedar | 272.30 to 272.40 | E | 330 | 229 | 10,919 | 1992 |
| RCA/13 | I-80 in Cedar | 272.55 to 272.65 | E | 330 | 229 | 10,919 | 1992 |
| RCA/14 | I-80 in Cedar | 269.30 to 269.40 | W | 318 | 216 | 10,847 | 1992 |
| RCA/15 | I-80 in Cedar | 269.10 to 269.20 | W | 330 | 203 | 10,847 | 1992 |
| RCA/16 | I-80 in Cedar | 275.70 to 275.75 | W | 330 | 203 | 10,847 | 1992 |
| RCA/17 | I-80 in Cedar | . 275.90 to 275.95 | W | 330 | 203 | 10,847 | 1992 |
| RCA/18 | I-80 in Cedar | 276.60 to 276.70 | E | 343 | 229 | 10,920 | 1991 |

a. ADTT: Average Daily Truck Traffic
b. N/A: Not Available

## Visual distress survey

Visual distress surveys were conducted to gather current detailed pavement surface condition information including the extent and severity of the distress. Core samples were taken during the visual distress survey to confirm that the selected pavement sites were indeed concrete pavements with RCA granular layer. Typical RCA granular layer condition after coring is illustrated in Figure 1. The distress survey methodology used in this study was the one described in the Strategic Highway Research Program's (SHRP) "Distress Identification Manual for the Long-Term Pavement Performance (LTPP) Project." (Miller and Bellinger, 2003). The distress types and severity levels were identified using the Distress Identification Manual and recorded on the distress map sheets with the symbols. Visual survey distress map was prepared for individual test section.

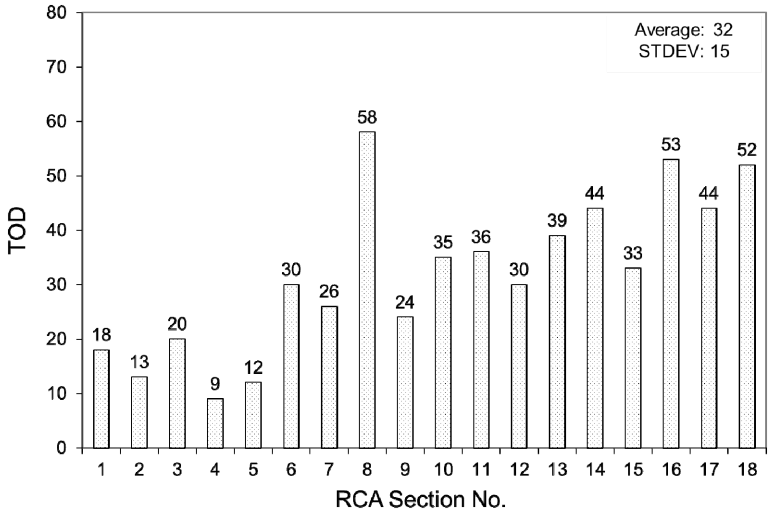
**Figure 1. RCA subbase condition underneath concrete slab after coring.**

### Pavement performance information from PMIS

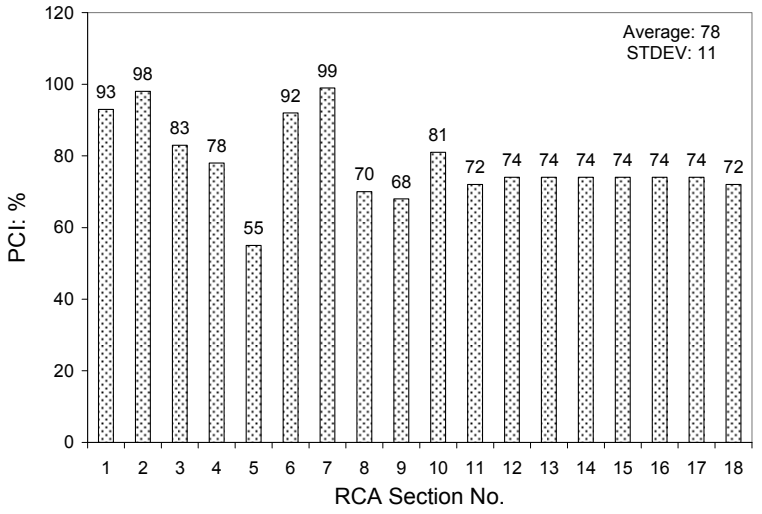
The most recent pavement performance information of 2006 for the surveyed test sections was also extracted from the Iowa DOT's PMIS. The PCI and the IRI were used as the objective measure representing the pavement performance information in this study. The PCI is a numerical index, ranging from 0 for a failed pavement to 100 for a pavement in perfect condition, to provide an index of the pavement's structural integrity and pavement surface condition. The IRI represents the severity of roughness on pavement surface computed from the measured longitudinal pavement profile.

## PAVEMENT PERFORMANCE OF CONCRETE PAVEMENTS WITH A RCA GRANULAR LAYER

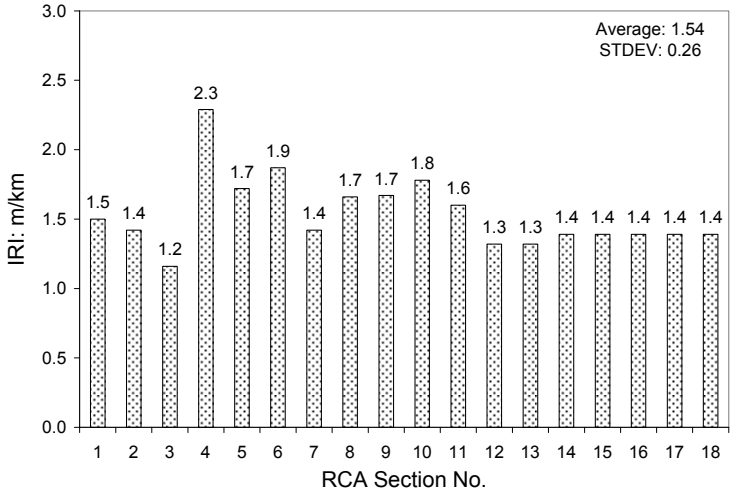
The pavement performance results for the surveyed test sections are summarized in Figure 2. The total number of distress (TOD) shown in Figure 2 (a) is the sum of number of distresses identified through the visual distress survey. Although a few popouts were observed in certain sections, these numbers were not added in the calculation of TOD since popouts do not result from subbbase problem. The values of PCI and IRI in Figure 2 (b) and (c) represent the most current measurements (2006). The average TOD, PCI and IRI value of RCA sections are 32, 78%, and 1.54 m/km, respectively. These indicate fairly good of pavement surface condition.




(a)




(b)




(c)

**Figure 2. Summary of pavement performance results.**

Figure 3 show the pavement surface condition on test section of RCA/15. Although a large number of pictures were taken as part of the visual distress survey for individual test sections to provide the types and severities of the distress, some representative pictures are included here which are indicative of the overall conclusion.

As seen in Figure 3, few longitudinal and transverse cracks were observed in the surveyed field test sections. The predominant distress exhibited along the all surveyed test sections are joint sealing damage and spalling. Especially, the lane-to-shoulder separation and/or the lane-to-shoulder drop off as shown in Figures 3 and 4 are often observed in the most RCA test sections. These observations are consistent with those reported by Rollings et al. (2006), who concluded that the cause of the distresses between lane and shoulder is probably due to sulfate attack on the RCA used as fill and granular layer.

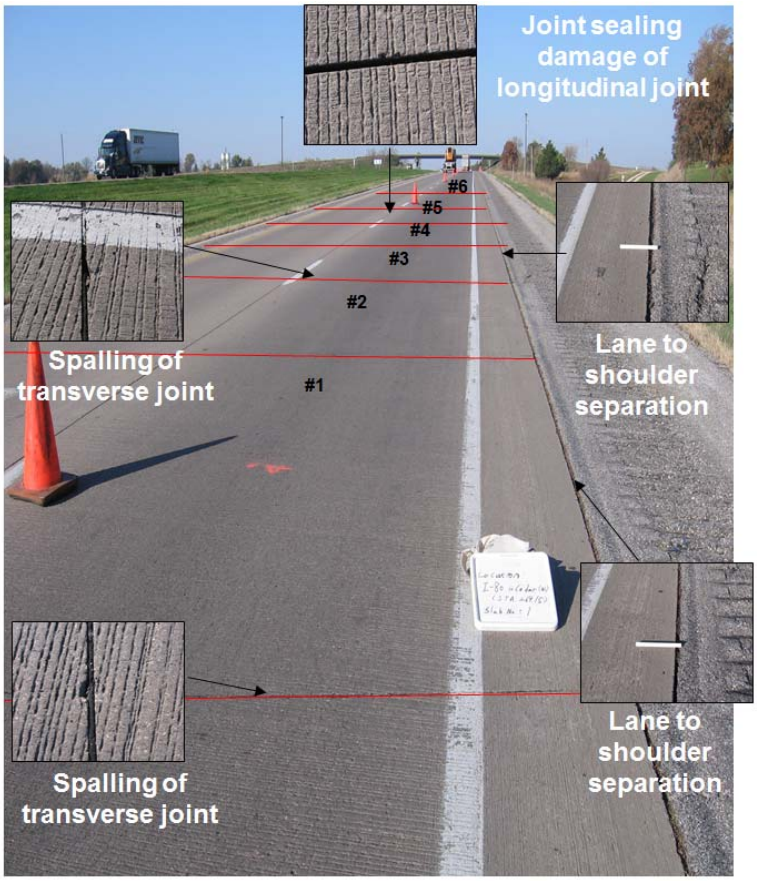



**Figure 3. Pavement surface condition in test section of RCA/15 (I-80 in Cedar county mile post 269.10 to 269.20).**


**Figure 4. Lane to shoulder drop off in test section of RCA/5 (I-35 in Hamilton county mile post 131.40 to 131.45).**

## COMPARISON OF PAVEMENT PERFORMANCES OF RCA AND VIRGIN AGGREGATE SECTIONS

The virgin aggregate granular layer sections corresponded to the surveyed RCA granular layer sections in similar conditions were identified from the Iowa DOT's PMIS and listed in Table 2 for comparisons. The identified virgin aggregate section counterpart had similar traffic volume, construction year, and location but different granular aggregate layer type (virgin as opposed to recycled) corresponding to the surveyed RCA section.

**Table 2. Virgin aggregate granular layer concrete pavement sections identified**

| Section No. | Location | | | PCC Thick (mm) | Sub-base Thick (mm) | ADTT* | Construction Year |
|---|---|---|---|---|---|---|---|
| | County | Mile Post No. | Dir. | | | | |
| Virgin/1 | I-35 in Story | 117.09 to 121.48 | N | 292 | 229 | 5,069 | 1988 |
| Virgin/2 | I-35 in Hamilton | 126.04 to 131.03 | S | 300 | 260 | 4,762 | 1999 |
| Virgin/3 | I-80 in Jasper | 151.48 to 156.28 | E | 305 | 229 | 8,837 | 1993 |
| Virgin/4 | I-80 in Jasper | 149.89 to 151.48 | W | 292 | 229 | 8,848 | 1990 |
| Virgin/5 | I-35 in Hamilton | 126.04 to 130.60 | N | 254 | 203 | 4,763 | 1985 |
| Virgin/6 | I-80 in Pottawattamie | 21.70 to 28.04 | W | 300 | 260 | 5,306 | 1998 |
| Virgin/7 | I-80 in Pottawattamie | 20.70 to 28.04 | E | 300 | 260 | 5,331 | 1999 |
| Virgin/8 | I-80 in Cass | 49.71 to 55.33 | E | 292 | 229 | 7,285 | 1989 |
| Virgin/9 | I-80 in Cass | 55.33 to 59.90 | W | 305 | 229 | 7,478 | 1992 |
| Virgin/10 | I-80 in Polk | 137.81 to 141.58 | E | 318 | 152 | 11,445 | 1994 |

*ADTT: Average Daily Truck Traffic in 2005

The vales of PCI and IRI in 2006 for the surveyed RCA subbase sections and the corresponding virgin aggregate subbase sections were extracted from the Iowa

DOT's PMIS. Figures 5 and 6 present comparison results of the RCA/1 to RCA/10 sections in Table 1 with those of virgin/1 to virgin/10 in terms of PCI and IRI.

The average PCI of RCA aggregate sections are 82 % with 14% of standard deviation while those of virgin aggregate are 74% with 13% of standard deviation. The average IRI of RCA aggregate sections are 1.65 m/km with 0.31 m/km of standard deviation while those of virgin aggregate are 1.48 m/km with 0.20 m/km of standard deviation. The paired t-tests were performed to evaluate if the PCI and IRI obtained under the RCA sections are statistically different from those under the virgin aggregate sections. A paired t-test uses the mean of difference between the observations in one group and the matched observations in the other group (SAS, 2005). In the Figures 5 and 6, the symbol '$\mu_D$' indicates the mean of differences between the PCI or IRI in the RCA sections and those in the corresponding virgin aggregate section; thus accepting null hypothesis, i.e., the PCI or IRI values in both of test sections are not significantly different. These results indicate that the RCA granular layer provides at least similar, if not better performance compared to the virgin aggregate layer in Iowa pavements and are performing adequately.

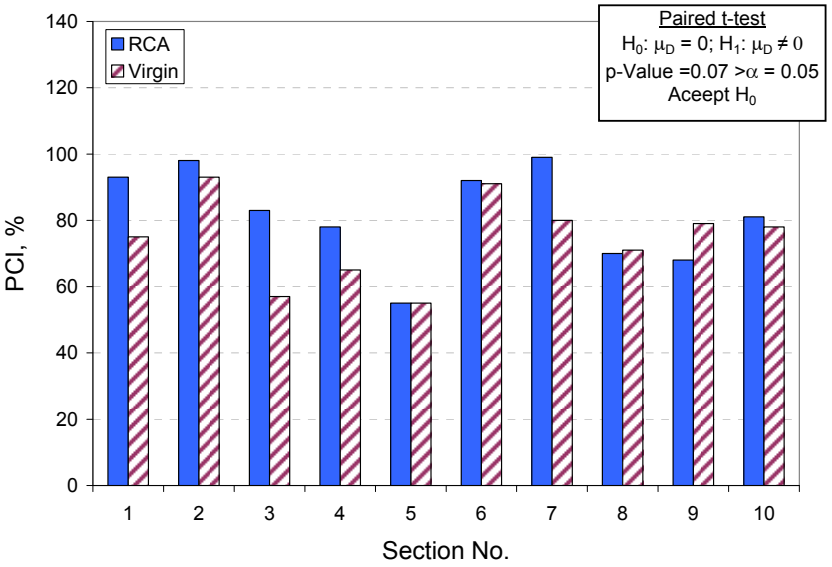



**Figure 5. PCI values for the RCA and the virgin aggregate sections.**

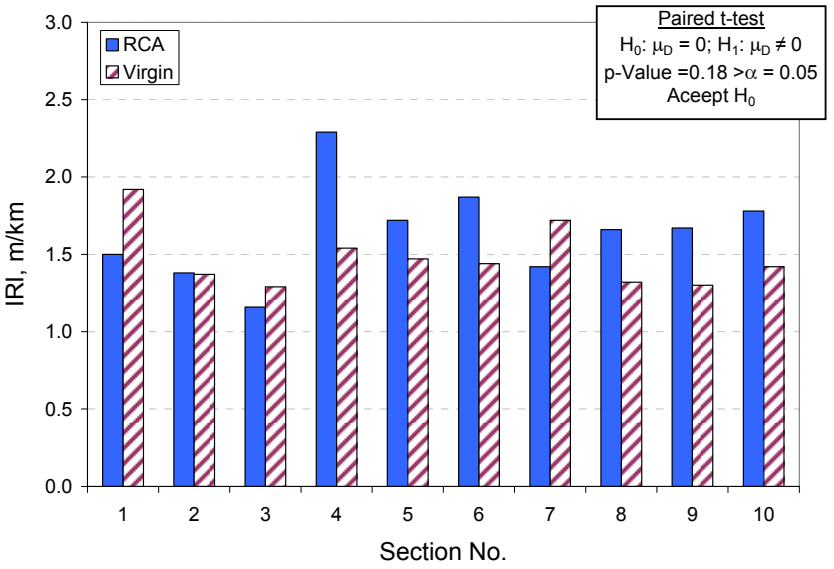



**Figure 6. IRI values for the RCA and the virgin aggregate sections.**

## CONCLUSIONS

This study investigated the long-term field performance of concrete pavement with Recycled Concrete Aggregate (RCA) granular layer. Based on the results of this study, the following observations were drawn:

- Based on the evaluation of representative RCA pavement sections, a RCA granular layer provides performance compared to a virgin aggregate granular layer in Iowa pavements and is performing adequately.
- Few structural distresses such as longitudinal and transverse cracks were observed on all RCA pavement test sections evaluated in this study. The featured distresses on RCA are the lane-to-shoulder separation and lane-to-shoulder drop off, which are consistent with the findings reported by previous researchers.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Iowa Department of Transportation (Iowa DOT) and the Iowa Highway Research Boards (IHRB) for supporting this study. The contents of this paper reflect the views of the authors who are responsible for the facts and accuracy of the data presented within. The contents do not necessarily reflect the official views and policies of the Iowa DOT. This paper does not constitute a standard, specification, or regulation.

## REFERENCES

Chini, A. R., Kuo, S. S., Duxbury, J. P., Monteiro F. M. B., and Mbwambo W. J. (1998). *Guidelines and specifications for the use of reclaimed aggregates in*

*pavement*, Final Report for Contract BA 509, Florida Department of Transportation, State Materials office, Gainesville, FL.

Chini, A. R., Kuo, S. S., Armaghani, J. M., and Duxbury, J. P. (2001). "Test of recycled concrete aggregate in accelerated test track." *American Society of Civil Engineering Journal of Transportation Engineering*, 127(6): 486-492.

FHWA. (2002). Formal policy on the use of recycled materials, http://www.fhwa.dot.gov/legsregs/directives/policy/recmatmemo.htm (Accessed in March 2008), U.S. Department of Transportation, Washington, D.C.

FHWA. (2004). *Transportation applications of recycled concrete aggregate*, U.S. Department of Transportation, Washington, D.C.

Huang, H. Y. (2003). *Pavement analysis and design*, 2nd ed., Pearson Prentice Hall, NJ.

Miller, J. S., and Bellinger, W. Y. (2003). *Distress identification manual for the Long-Term Pavement Performance (LTPP) project*, FHWA-RD-03-031, 4th Ed, Federal Highway Administration, Mclean, Virginia.

Miyagawa, F. K. (1991). *Permeability of granular subbase materials*, Interim Report for Project MLR-90-4, Iowa Department of Transportation, Ames, IA.

Rollings, R. S., Rollings, M. P., Wong, S., and Gutierrez, G. (2006). "Investigation of heaving at Holloman air force base, New Mexico." *American Society of Civil Engineering Journal of Performance of Constructed Facilities*, 20(1): 54-63.

SAS Institute, Inc. (2005). *JMP 6.0 Statistics and Graphics Guide*, Cary, NC.

Snyder, M. B. (1995). *Use of crushed concrete product in Minnesota pavement foundations*, Final Report MN/RD-96/12, Minnesota Department of Transportation, St. Paul, MN.

White, D. J., Jahren, C. T., Cackler, E. T., and Vennapusa, P. (2004). *Determination of the optimum base characteristics for pavements*, Final Report for Project TR-482 and CTRE Project 02-119, Iowa Department of Transportation, Ames, IA.

Wilburn, D. R., and Goonan, T. G. (1998). *Aggregate from natural and recycled sources, economics assessment of construction applications – a materials flow analysis*, United State Geological Circular 1176, United State Department of Interior, Washington, D.C.

# Green Initiatives at MnROAD

Timothy R. Clyne, P.E[1], Benjamin J. Worel, P.E.[2], and Melissa K. Cole, P.E.[3]

[1] Minnesota Department of Transportation, Office of Materials & Road Research, 1400 Gervais Ave., Maplewood, MN 55109; PH (651)366-5473; FAX (651)366-5461; email: tim.clyne@state.mn.us

[2] Minnesota Department of Transportation, Office of Policy Analysis, Research & Innovation, 395 John Ireland Blvd, St. Paul, MN 55155; PH (651)366-3757; FAX (651)366-3589; email: ben.worel@state.mn.us

[3] Minnesota Department of Transportation, Office of Materials & Road Research, 1400 Gervais Ave., Maplewood, MN 55109; PH (651)366-5432; FAX (651)366-5461; email: melissa.cole@state.mn.us

## Abstract

The transportation network and its supporting infrastructure have significant economic, social, and environmental impacts. Using more sustainable methods to design, construct, and preserve our roads will better protect the environment and meet our ongoing transportation needs. The Minnesota Department of Transportation (Mn/DOT) and our partners in government, industry, and academia have been researching and implementing ways to make our roads greener while maintaining and improving roadway quality.

The Minnesota Road Research Facility (MnROAD), an accelerated pavement test facility near St. Paul, MN, is conducting research on several “green” pavement technologies that reduce environmental impacts, reuse roadway materials, or use recycled materials in pavement applications.

Green pavement technology should maintain or improve the state of the practice of pavement engineering. This ensures the green product will be market-driven, cost-effective, and sustainable.

## Background

The Minnesota Road Research Project (MnROAD) was constructed by the Minnesota Department of Transportation (Mn/DOT) in 1990-1993 as a full-scale accelerated pavement testing facility, with traffic opening in 1994. Located 40 miles northwest of St. Paul, MN, MnROAD is one of the most sophisticated, independently operated pavement test facilities of its type in the world. Its design incorporates thousands of electronic in-ground sensors and an extensive data collection system that provide opportunities to study how traffic loadings and environmental conditions affect pavement materials and performance over time. MnROAD consists of two unique road segments located parallel to Interstate 94:

- A 3.5-mile Mainline interstate roadway carrying "live" traffic averaging 28,500 vehicles per day with 12.7% trucks.
- A 2.5-mile closed-loop Low Volume Road carrying a MnROAD-operated 18-wheel, 5-axle, 80,000-lb tractor-semi-trailer to simulate the conditions of rural roads.

Over time, many of the original test sections (cells) have deteriorated. Several new research opportunities have been constructed at MnROAD since 2007. Included in these research opportunities are many projects looking at "greener" pavements. Presented in this paper are the Green Initiatives at MnROAD.

## Reduce, Reuse, Recycle

MnROAD is conducting research on the following "green" pavement technologies that reduce environmental impacts, reuse roadway materials, or use recycled materials in pavement applications:

REDUCE
- Warm mix asphalt – reduce energy costs and greenhouse gas emissions
- Pavement noise – reduce roadway noise pollution with innovative surface textures
- Permeable pavements – reduce storm water runoff

REUSE
- Full depth reclamation – reuse in-place roadway materials
- Taconite aggregates – reuse high quality aggregate

RECYCLE
- Asphalt and concrete pavements – recycle old roadways into new pavement surfaces
- Base materials – recycle old concrete and asphalt roadways as unbound base materials
- Shingles – recycle both tear off and manufactured shingles

These pavement techniques are being analyzed at MnROAD to determine their potential to improve the quality and cost-effectiveness of pavements, while also providing environmental benefits.

## Warm Mix Asphalt

Warm mix asphalt (WMA) technologies allow asphalt to be produced and placed at lower temperatures. Typical warm mix is heated to around 250°F, which reduces temperatures by 50 to 100°F compared to typical Hot Mix Asphalt (HMA) temperatures of 330°F. Such drastic reductions cut fuel consumption and decrease the production of greenhouse gases and fumes. Potential engineering benefits include better compaction, less thermal cracking, and the ability to pave in cooler weather.

In 2008, MnROAD constructed six Mainline test cells to demonstrate WMA's constructability and evaluate its long-term performance. The mix used was a level 4 Superpave (3-10 million ESALs) with PG 58-34 binder and 20 percent Recycled Asphalt Pavement (RAP) from MnROAD millings.

Five cells were constructed consisting of 5-inch WMA (3-inch wear, 2-inch non-wear) over a 12-inch recycled aggregate base, 12-inch aggregate subbase, 7-inch select granular, and clay subgrade. A single asphalt mixture was used to cover five cells for a recycled unbound base study, which will be described later in this paper. The sixth cell is a 3" WMA overlay of an existing HMA pavement, representing a "typical" Minnesota rehabilitation strategy. The contractor chose to use a liquid chemical additive in all the mixes. MnROAD also included a control cell on the Low Volume Road. The control cell has the same mix design as the WMA, but is produced at typical HMA temperatures without the additive (Johnson et al, 2009).

Several samples were collected for further analysis in the laboratory by MnROAD. Our research partners tested for mixture and binder properties. Initial laboratory tests have shown that WMA performs as well as, or better than, similar HMA mixtures.

Observations by the contractor during construction confirmed WMA handled similar to HMA, had comparatively lower energy costs, and noticeably reduced asphalt fumes. In addition, density was achieved with a faster roller speed, less vibration, and less passes. As the technology matures in future years and producers and contractors become comfortable with it, we expect to see more projects being constructed using WMA. Mn/DOT's construction specifications have been updated to allow WMA on all projects, and it has been required on a few projects in 2010 (Mn/DOT, 2010). Several contractors have recently installed a WMA foaming nozzles on their plants, and others have used chemical additives successfully.

## Pavement Noise

Pavement surfaces need a balance of texture and smoothness to be safe in diverse weather conditions and provide a smooth, quiet ride. Researchers have been developing improved surface textures for new and existing pavements and methods to monitor in-place pavement performance as it relates to ride quality, friction, quietness, and texture. One such texture involves diamond grinding an existing concrete pavement to restore and/or improve its surface characteristics.

At typical freeway speeds, tire-pavement noise is the dominant traffic noise source. In 2007, MnROAD commenced On-Board Sound Intensity (OBSI) testing—a method to measure tire-pavement noise using sophisticated microphones mounted on a vehicle's rear wheel near the pavement surface. Minnesota is one of only five states that adopted this technology (Izevbekhai, 2007).

MnROAD has partnered with other states, industry, academia, and FHWA on three distinct, but related, research projects that investigate surface characteristics of both new and rehabilitated asphalt and concrete pavements. The overall goal of these projects is to optimize ride quality, quietness, pavement durability, and safety from hydroplaning and splash and spray.

As part of these projects, two concrete test sections on the Mainline were diamond ground. One section was a conventional diamond grind, while the other was an innovative grind first developed at Purdue University. Figure 1 shows the results of several rounds of noise testing on these two ground surfaces as well as traverse tined concrete, which was Mn/DOT's typical surface texture in the 1990's. The figure shows that conventional diamond grinding a pavement surface reduces the noise level by approximately 3dBA, which is equivalent to either doubling the distance between noise source and receptor or reducing the traffic that generated the noise by half. The innovative grind reduced the noise level by an additional 2dBA, causing further benefits perceived by the driving public.

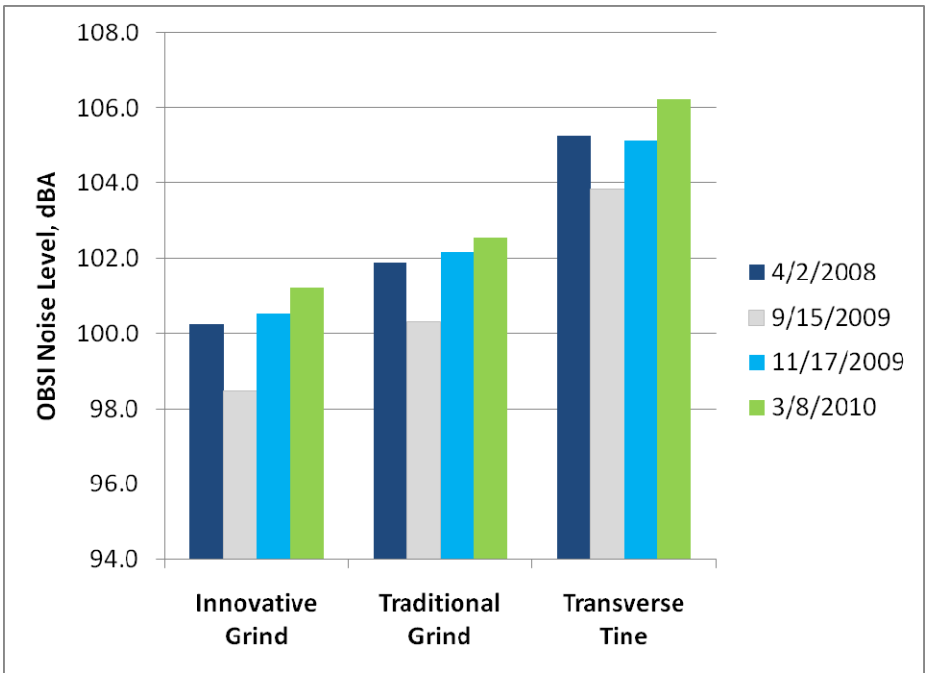



Figure 1. OBSI Data for 3 Types of Concrete Surfaces

**Permeable Pavements**

The increase in impervious surfaces has been an issue of concern for every community as housing and retail development continues to progress. Run-off from impermeable surfaces has been known to impact streams, wetlands, and other water resources in a negative way. Metro area communities have made various attempts to encourage infiltration by constructing permeable pavement on porous bases. Permeable pavements can be made from either asphalt or concrete. Additional benefits of permeable pavements include noise and splash and spray reduction.

Mn/DOT wanted to determine if these types of pavements perform well in northern climates. These climates experience cold winters and hot summers. By adequately evaluating permeable pavements in this climate, changes in porosity and infiltration can be monitored over time under measurable traffic loads, environmental conditions, and deicing operations.

The Minnesota Local Road Research Board (LRRB) recently funded two permeable pavement projects at MnROAD; one on porous asphalt and one on pervious concrete. To support this research, MnROAD constructed porous asphalt and pervious concrete sections, each on a clay and sand subgrade, as well as a conventional asphalt pavement (Johnson et al, 2009). The pavement designs called for a porous/pervious surface layers above and open graded aggregate base that could store storm water until it infiltrated into the subgrade. Each of the 5 test cells contains a traverse drain that collects the surface runoff if and when the pavement gets clogged. Each of the drains outlets to a collection tank, and when these tanks become full of runoff, they tip to empty into the ditch. Each tip is approximately 16 gallons. Figure 2 shows the number of tips from 2 similar rain events over the past year. The porous asphalt and pervious concrete sections show a significant reduction in storm water runoff compared to the conventional asphalt pavement sections.

MnROAD also partnered with industry and academia to build a pervious concrete overlay over a jointed concrete pavement. This technology allows water to flow through the pervious overlay laterally to the shoulders. It is also expected to reduce noise while providing a smooth, safe driving surface.

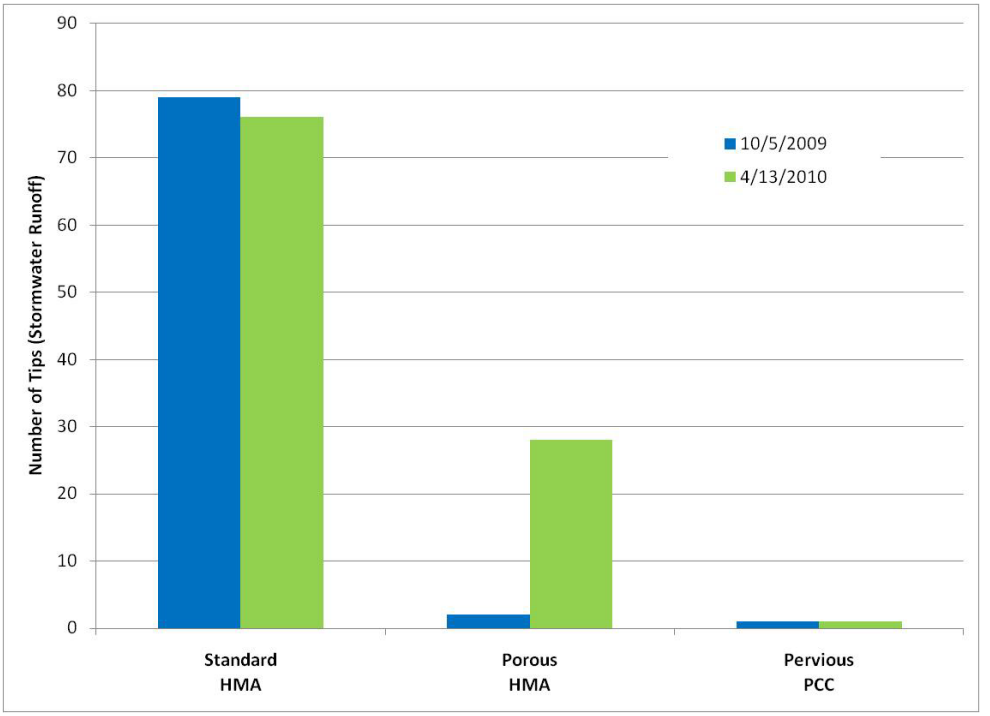



Figure 2. Water Runoff in Pervious Concrete and Porous Asphalt

**Full Depth Reclamation**

Full depth reclamation (FDR) is a construction technique in which the full asphalt pavement section and a predetermined portion of the underlying base materials are uniformly crushed, pulverized, and blended, resulting in a stabilized base course. Pavements that experience base failures or high amounts of cracking are considered ideal candidates for FDR. Further stabilization may be accomplished for these roads through available additives, including cement, fly ash, lime, foamed asphalt, and asphalt emulsions. A new surface course is then placed on the recompacted base. The advantage of this construction method is that an agency can reuse all of the in-place materials without having to haul in new granular materials, saving both material and trucking costs.

MnROAD has constructed a number of test sections studying this technology. In 2000, a 6" HMA test section on the Low Volume Road was reclaimed. MnROAD learned the hard way any work that blends clay subgrade material during the FDR process needs to be further stabilized with an additional additive (Mulvaney et al, 2004). This test section failed after only 38 truck passes because of low shear strength in the FDR + clay layer.

In 2007, three full depth reclamation test sections were constructed on the Low Volume Road through a fly ash stabilization study. Many coal power plants in the United States are looking for ways to use their fly ash (a byproduct of electricity production) to avoid land filling. Fly ash contains cementitious properties and is

frequently used to enhance mixtures of Portland cement concrete. While High Carbon Fly Ash cannot be used in the concrete paving industry, laboratory testing has shown it to be a viable stabilizing agent for unbound materials. MnROAD constructed three test sections consisting of an asphalt surface over fly ash treated FDR, non-treated FDR, or a crushed stone base as the control. The FDR + fly ash section proved to be the most stable construction platform during a wet construction season (Clyne and Palek, 2008).

In 2008, MnROAD constructed three FDR test sections on the Mainline in partnership with private industry to demonstrate alternate contracting methods and to evaluate the properties and performance of FDR using asphalt emulsion stabilization. This project was conducted to determine the optimal percentage blend of HMA and aggregate base for full depth reclamation with engineered emulsion. The three blends of HMA and aggregate base were at 50/50, 75/25, and 100/0. Researchers also evaluated how to best rehabilitate full depth HMA on clay pavements. These test sections rely primarily on the stabilized FDR layer to support traffic loading, since only a thin surface of asphalt was placed as the wearing course (Johnson et al, 2009).

Each of the FDR studies mentioned above demonstrates rapid, cost-effective construction experiences, especially when accompanied by additional stabilization. To date the 2007 and 2008 FDR sections are performing quite well, with the stabilized sections showing a measurable increase in strength over pavements with more conventional aggregate bases.

**Taconite Aggregate**

MnROAD has developed a number of partnerships with state and federal agencies since 2004 to research the use of taconite materials in roadways. An initial partnership helped construct two test sections (asphalt and concrete) using taconite aggregates at MnROAD. The most recent partnership uses federal funds to research Minnesota Iron Range aggregate materials in transportation applications. Currently aggregate material from the Iron Range is considered “waste product.” However, the goal is to assess the available aggregate resource that could potentially supply an abundant, high-quality, low-cost aggregate for roadway use.

The use of fine to coarse taconite aggregate materials in hot mix asphalt, concrete pavement, and aggregate base layers is being studied. Laboratory work on taconite aggregates is being carried out in parallel at Mn/DOT and the University of Minnesota.

In 2008, MnROAD built one test section on the Mainline to compare permeable large stone base materials to traditional finer non-permeable aggregate base materials. It will allow researchers to demonstrate:

- The benefits of a higher quality, stronger base layer
- Its effect on hot mix asphalt pavement performance
- How to account for these materials during design and construction

A second Mainline section was constructed to demonstrate the use of a 4.75-mm Superpave HMA pavement using two different sources of taconite tailings and locally available manufactured sand. This mixture, made with sand-size aggregate particles, makes a durable pavement surface, as well as provides superior surface characteristics such as friction, noise abatement, texture, ride quality, and splash and spray reduction. To date, we have shown at MnROAD that pavements constructed with taconite aggregates have performed as well as or better than those using conventional, locally available materials (Clyne et al, 2009).

**Recycled Asphalt and Concrete Pavements**

Using reclaimed asphalt pavement (RAP) or recycled concrete aggregate (RCA) pavements in new pavement construction has numerous advantages to the environment, owner agencies, and contractors.

Environmental benefits include a reduction of the product's carbon footprint and any of its end uses, conservation of natural resources, and conservation of landfill space. From an economic standpoint, the recycling of materials provides an opportunity to stabilize material prices, which fluctuate as the economy and demand for raw materials change. The technology for using increased amounts of RAP in asphalt mixtures has improved significantly in terms of mix design and material processing and handling.

In 2008, MnROAD built three test sections designed around recycled asphalt pavements, including a number of control sections. The RAP used in this project was recycled from the MnROAD Mainline, crushed to size, and reincorporated into the new HMA pavement (Johnson et al, 2009).

Test sections consisted of:

- 30% Non-Fractionated RAP (PG 58-28 wear and nonwear)
- 30% Fractionated RAP (PG 58-28, 20% Fine RAP, 10% Coarse RAP, wear and nonwear)
- 30% Fractionated RAP (PG 58-34, 20% Fine RAP, 10% Coarse RAP, wear and nonwear)

In addition, in 2010 MnROAD constructed three test sections as part of a SHRP II Composite Pavement Study. Two of the sections incorporated RCA as a partial course aggregate substitution in the lower concrete layer. Each section received a 3" overlay. One cell received a typical asphalt mixture and in the other cells a high quality concrete surface was placed. We were able to demostrate the viability of constructing concrete pavements that used recycled concrete as an aggregate. Figure 3 below is from the concrete composite pavement construction. Early performance data shows that these sections are performing well.


Figure 3. 2010 Composite Pavement Construction, Wet on Wet Concrete

**Recycled Unbound Base Materials**

Minnesota has a long history of using recycled materials in pavement construction. Mn/DOT's current construction specification allows salvaged or recycled asphalt, concrete, and glass to be used as part of the granular base materials. However, their material properties (strength, stiffness, unsaturated properties, etc.) are not well understood. New pavement design procedures require more detailed material properties to accurately predict pavement performance.

Mn/DOT initiated a pooled fund study to construct several test sections at MnROAD using recycled materials in the granular base layers. The material properties were monitored during construction and throughout the pavement life to determine their effects on pavement performance. The properties will be used to verify mechanistic-empirical design inputs, especially their variation with changing seasons and moisture regimes. Each of the test sections for this effort had 5" warm mix asphalt, 12" recycled base (see below), 12" aggregate subbase, and 7" select granular over clay subgrade.

The four test sections consisted of:

- Recycled concrete base
- 50% recycled concrete + 50% class 5 granular base
- Recycled asphalt pavement (RAP) base
- Class 5 aggregate base (control – Mn/DOT standard virgin base material)

Figure 4 shows a snapshot of the Falling Weight Deflectometer (FWD) testing on the asphalt pavement surface from April 5, 2010. This late spring measurement is an average of 5 test points at a 9,000 lb load level. The plot shows that the virgin

aggregate base has higher deflections than the three recycled base material. The RAP has the lowest deflection of the four materials.

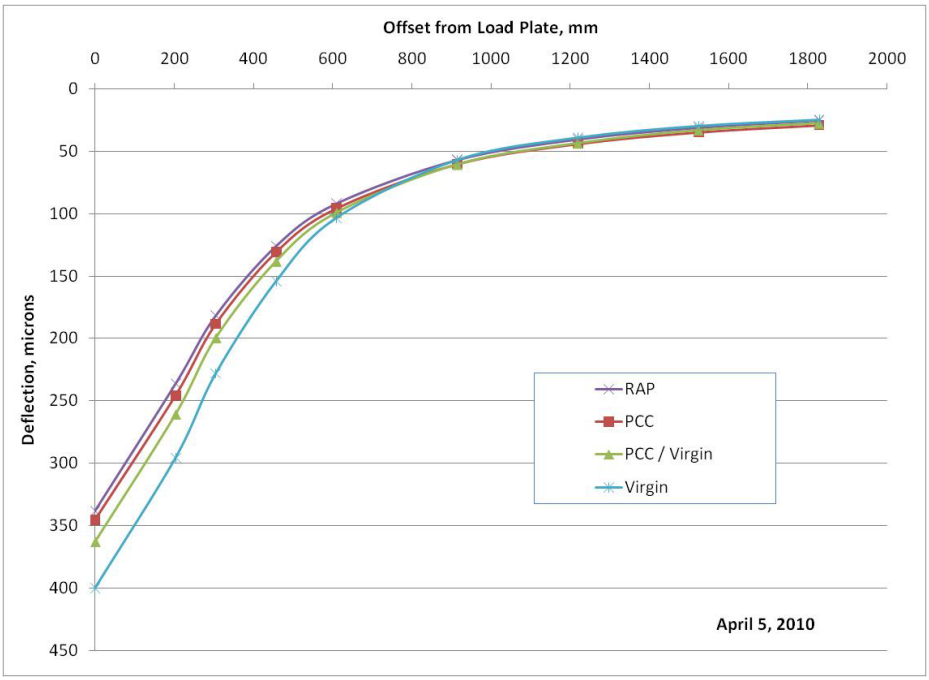



Figure 4. FWD Testing Results for Recycled Unbound Base Materials

**Shingles**

Manufacturer waste shingles (MW) and tear-off shingles (TO) from homes are being used nationwide as beneficial additions to asphalt pavements. MW shingles are obtained from a shingle manufacturing facility as scrap. They contain approximately 20% asphalt binder by weight. Shingle asphalt binder is significantly harder than paving grades of asphalt.

TO shingles are old shingles that have been removed from a roof. Granules are often lost; therefore TO shingles contain approximately 25% aged hardened asphalt. Mn/DOT is working closely with regulatory agencies and industry recyclers to develop quality control measures to allow shingles to be incorporated into asphalt mixes in an environmentally-sound fashion without impacting the quality or service life of the pavement. Recycling these shingle materials into pavements will dramatically reduce the amount of shingles that are landfilled every year. Some local companies are exploring ways to process the recycled shingles to create an even better HMA paving product.

In 2008, MnROAD used both manufactured waste and tear-off shingles on the Mainline. Tear off shingles were used on the shoulders on 10 different test sections and the manufactured shingles were used on the shoulders for 5 test sections. Each mix included 5% shingles with no other recycled asphalt pavement (RAP), using a PG 58-28 binder (Johnson et al, 2009).

A concurrent laboratory study performed at Mn/DOT investigated a large number of mixtures containing RAP and/or shingles (Johnson et al, 2010). Based on the results of this study and the field work performed at MnROAD, Mn/DOT's construction specifications were modified to allow both MW and TO shingles at up to 5% (Mn/DOT, 2010).

## Conclusions

Through a variety of research projects at MnROAD, researchers are discovering that some "green" solutions create an improvement in pavement performance, help the environment, and are cost-effective. Many products that were previously considered waste are now being used to build better roads. Old pavements are being recycled to be used as new base material or as new pavements saving on virgin materials, trucking costs, and fewer trucks mean less fuel emissions into the environment. Noise can be better controlled with a new diamond grinding technique or other innovative surface textures.

As the test sections discussed in this paper continue to age, researchers will better understand these green solutions for building roads and will be able to implement some of these technologies into future construction projects.

## References

Clyne, T.R., and Palek, L.E. (2008). *2007 Low Volume Road & Farm Loop (Cells 33, 34, 35, 77, 78, 79, 83, 84) Construction Report*. Minnesota Department of Transportation, St. Paul, MN.

Clyne, T.R., Johnson, E.N., and Worel, B.J. (2009). *Use of Taconite Aggregates in Pavement Applications*. Minnesota Department of Transportation, St. Paul, MN.

Izevbekhai, B.I. (2007). *Report of Diamond Grinding on Cells 7 & 8 MnROAD Mainline Interstate Highway I-94*. Minnesota Department of Transportation, St. Paul, MN.

Johnson, A., Clyne, T.R., and Worel, B.J. (2009). *2008 MnROAD Phase II Construction Report*. Minnesota Department of Transportation, St. Paul, MN.

Johnson, E., Johnson, G., Dai, S., Linell, D., McGraw, J., and Watson, M. (2010). *Incorporation of Recycled Asphalt Shingles in Hot-Mixed Asphalt Pavement Mixtures*. Minnesota Department of Transportation, St. Paul, MN.

Minnesota Department of Transportation (Mn/DOT), (2010). "2360 Plant Mixed Asphalt Pavement, Gyratory Design Specification," *Standard Specifications for Construction*. Minnesota Department of Transportation, St. Paul, MN.

Mulvaney, R., Worel, B., and Zerfas, W. (2004). *2004 Cell 26 Reconstruction*. Minnesota Department of Transportation St. Paul, MN.

# Impact of Mixed Nitrogen Dioxide ($NO_2$) and Nitrogen Oxide (NO) Gases on Titanium Dioxide Photodegradation of $NO_x$

**Marion Schmitt[1], Heather Dylla[2], Marwa M. Hassan[3], Louay N. Mohammad[4], Tyson Rupnow[5], and Earle Wright[6]**

**Abstract:** Nitrogen oxides ($NO_x$) emitted from vehicle exhausts are associated with adverse health effects on the public. Self-cleaning, air-purifying concrete pavement is a rapidly emerging technology that can be constructed using photocatalyst such as titanium dioxide ($TiO_2$). The main objective of this study was to evaluate the environmental effectiveness of $TiO_2$ coating in photodegrading mixed $NO_2$ and NO gases from the atmosphere. Results of the experimental program determined that increasing the flow rate and $NO_2/NO_x$ ratio negatively affect the effectiveness of the photocatalytic process. However, within the evaluated range, the titanium content and aggregate gradation had little impact on $NO_x$ removal efficiency. The highest photodegradation rate was observed at 25% relative humidity, which balances the availability of hydroxyl radicals at the surface with $NO_x$ contact with the photocatalytic surface.

## Introduction

$NO_x$ is a group of highly reactive gases, nitric oxide (NO) and nitrogen dioxide ($NO_2$), originating from high temperature fuel combustion (EPA 2010). Many negative impacts are associated with $NO_x$ emissions such as formation of tropospheric ozone, urban smog, acid rain, and fine particle pollution, which can cause many adverse health effects (Jimenez et al. 1999). As a result, $NO_x$ is one of the six major pollutants listed as a "criteria pollutant" as part of the ambient air quality standards required by the 1970 Clean Air Act (Tao et al. 2010).

The Clean Air Act has been successful in decreasing $NO_2$ concentration by 48% from 1970 to 2002 (Tao et al. 2010). However, in spite of the overall reduction in

[1] Research Assistant, Department of Construction Management and Industrial Engineering, Louisiana State University
[2] Graduate Research Assistant, Department of Construction Management and Industrial Engineering, Louisiana State University
[3] Assistant Professor, Department of Construction Management and Industrial Engineering, Louisiana State University, 3218 Patrick F. Taylor, Baton Rouge, LA 70803, e-mail: marwa@lsu.edu
[4] Irma Louise Rush Stewart Distinguished Professor , Department of Civil and Environmental Engineering , Director Engineering Materials Characterization Research Facility , Louisiana Transportation Research Center
[5] Concrete Research Engineer, Louisiana Transportation Research Center
[6] President/Sr. Scientist, Technical Monitoring Services Incorporated (TMSI) , 6310 Westwind Avenue, Baton Rouge, LA 70817

$NO_x$ emissions nationwide, high short-term concentration of this pollutant is raising new concerns. Mobile sources such as traffic emissions contribute up to 58% of $NO_x$ pollution emitted in the US (EPA 2010). This figure increases in areas with high traffic density (Jimenez et al. 1999). Compared to ambient concentrations, in-vehicle microenvironment $NO_x$ concentrations are 2-3 times higher, and near-roadway microenvironment $NO_x$ concentrations are 30-100% times higher (EPA 2010). Numerous studies have shown that even these short-term high concentrations can have negative impacts on human health (McConnel et al. 2010).

Self-cleaning, air-purifying concrete pavement is a rapidly emerging technology that can be constructed using air-cleaning agents with photocatalytic capabilities such as titanium dioxide ($TiO_2$). This technology is especially promising for city roads where trends of urban development and increase in traffic volumes add to emissions while street canyon conditions inhibit their dispersion resulting in high ground level pollutant concentrations. Recent research has shown that a thin surface coating is able to remove a significant portion of NO pollutants from the atmosphere when placed as close as possible to the source of pollution (Dylla et al. 2010). However, the effectiveness of photocatalytic compounds such as titanium dioxide in degrading mixed $NO_2$ and NO gases has not been evaluated.

The objective of this study was to evaluate the environmental effectiveness of $TiO_2$ coating in photodegrading mixed $NO_2$ and NO gases from the atmosphere. To achieve this objective, an experimental program was conducted in which the effects of the ratio of $NO_2$ and NO gases, relative humidity level, pollutants' flow rate, and mix design parameters including contents of $TiO_2$ and aggregate sizes were investigated.

## Background

Several methods have been investigated to reduce vehicle emissions such as lean $NO_x$ traps and Selective Catalytic Reduction (SCR). However, drawbacks exist with both techniques such as potential of producing secondary pollutants and thermal durability. An alternative solution is heterogeneous photocatalysis in which a semiconductor activated by light at a particular wavelength creates hydroxyl radicals, (*OH) and superoxides ($O_2^-$) that accelerate the natural decomposition process of organic pollutants (Fujishima et al. 2000). In contrast to other methods, pollutants are decomposed to nonhazardous waste products with little energy requirements and little selectivity (Zhao and Yang 2003). A proposed mode of oxidation of nitrogen oxides via hydroxyl radical intermediates in the presence of this photocatalyst is presented in Figure 1.

Various studies have incorporated photocatalytic oxidation in building materials for $NO_x$ reduction. Beelden (2008) studied photocatalytic concrete blocks' ability to decompose $NO_x$ pollutants. The concrete blocks consisted of an upper cement coat containing $TiO_2$. Poon and Cheung (2007) incorporated $TiO_2$ into several different concrete surface mix designs by changing the aggregates design. Hassan et al. (2010) evaluated three methods of application for $TiO_2$ to concrete pavement, a cementitious-based ultra-thin coating, a water-based titanium dioxide solution (PT), and sprinkling $TiO_2$ to the fresh concrete surface before hardening. The

environmental benefits of applying an ultra-thin $TiO_2$ coating were greater than the other two methods; however, the durability of the layer may have been comprised.

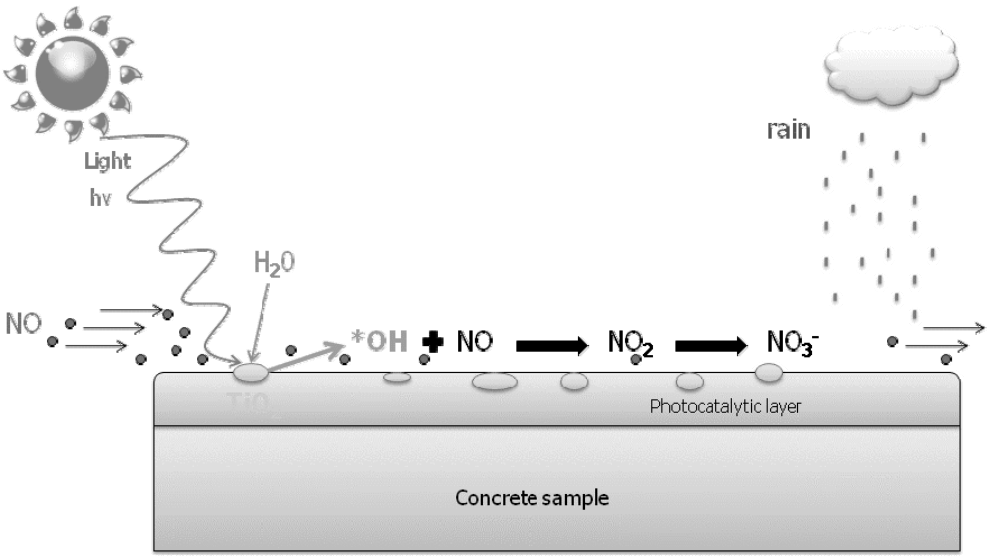



**Figure 1.** Illustration of the photocatalytic process

Although photocatalytic oxidation was proven effective in reducing $NO_x$ emissions, the impact of $NO_2$ pollution and its impact on overall $NO_x$ reduction efficiencies have not been evaluated. Since $NO_2$ is a highly unstable gas, it is more difficult to study. For example, reactions such as heterogeneous hydrolysis of $NO_2$ convert $NO_2$ into $HNO_3$ if water on a surface is available. This reaction has been noted to occur on many urban surfaces including concrete (Ramazan et al. 2004). Meanwhile, other studies described losses of $NO_x$ due to alkaline oxides forming in the cement matrix, which aids in $NO_x$ adsorption and would be an added benefit to use cement as a substrate for heterogeneous photodegradation (Cassar et al. 2003).

## Experimental Program

The objective of the experimental program was to evaluate the effects of different $NO_2/NO_x$ ratios on the effectiveness of photocatalytic concrete pavement to purify the air from $NO_x$ gases. For this purpose, five different factors were tested: $NO_2/NO_x$ ratio, concentration of $TiO_2$, mix design, humidity level, and flow rate. The test setup used to quantify $NO_x$ removal efficiency was modified from the Japanese Industrial Standard (JIS TR Z 0018 "Photocatalytic materials – air purification test procedure") in order to accommodate larger samples. This setup included a photocatalytic testing device, referred to a photoreactor, which created an enclosed controlled environment where the UV lights and the atmosphere can be simulated. The setup consists of a pollutant source (gas cylinder of NO), zero air source, adjustable valves, humidifier, calibrator, photoreactor, and chemilumiscent $NO_x$ analyzer.

The calibrator requires a pollutant source and a clean air source. A Thermo 146i Gas Calibrator supplied a constant polluted air stream at a controlled flow rate. The Thermo 146i included the Gas Phase Titration (GPT) option. This option allows $NO_2$

to be created in a reaction chamber using the rapid gas phase reaction between the NO and $O_3$ based on the following chemical reaction:

$$NO + O_3 \rightarrow NO_2 + O_2$$

The amount of $NO_2$ created can be changed by varying the amount of ozone created by adjusting the drive level of the ozonator in percentage when in the GPT mode (Thermo Scientific 2008). The amount of $NO_2$ created was measured by the Thermo 42i $NO_x$ analyzer, which was also used to measure $NO_x$ removal efficiency. The resulting GPT percentages used to simulate $NO_2/NO_x$ ratios of 0.7, 0.45, 0.2, and 0 were 25%, 20%, 15%, and 0, respectively. These ratios were all determined at a NO concentration of 550 ppb and a flow rate of 3 l/min.

To quantify the effectiveness of the photocatalytic concrete layer under different operating conditions, the JIS standard was used. The procedure calls for each sample to be tested for a total time of five hours under UV-irradiation that is started after at least 30 minutes to ensure equilibrium concentrations. The environmental efficiency is calculated by analyzing the concentrations of NO, $NO_2$, and $NO_x$ with the "lights on" compared to the difference in concentrations with the equilibrium at "lights off". Figure 2 illustrates the variation of $NO_x$ concentration during the testing process.

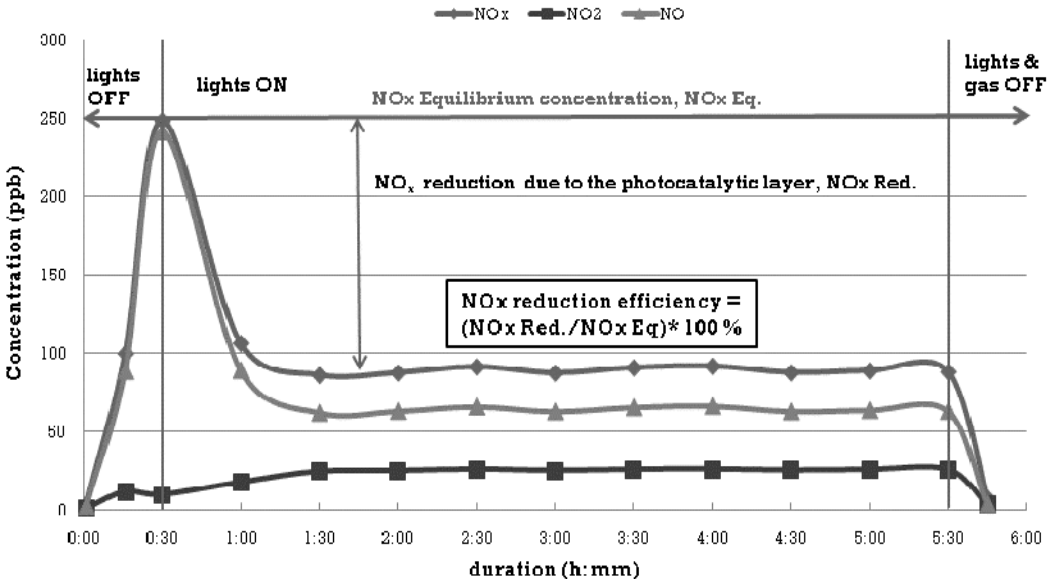



**Figure 2.** Variation of $NO_x$ concentration during the experiment

Table 1 presents the test factorial conducted in this study. It is noted that the flow rate was kept constant at 3 l/min, while the relative humidity was varied between five levels: 15%, 25%, 35%, 50%, and 80%.

### *Laboratory Samples and Materials Tested*

Prepared samples consisted of a concrete base and a photocatalytic $TiO_2$ surface coating. The base was made of concrete mix design typically used in Louisiana in order to achieve a compressive strength of 41 MPa. The concrete base layers were

cast in a wood mold of 310 mm x 381mm x 40 mm. A second removable layer was attached to the mold for the pouring of the photocatalytic concrete layer of 10 mm. The photocatalysis surface coating was applied after the concrete initial setting. The photocatalytic coating consisted of ultrafine titanium dioxide, cement, aggregate filler, and water. The titanium dioxide nanomaterial used was a commercially available *Cristal Millennium* PC105. The selected filler was sand with a maximum nominal size of 1.18 mm. The test matrix consisted of four different surface mix designs, two different $TiO_2$ contents, and two different aggregate gradations (with fines and without fines). The surface mixes produced with no fines were prepared by removing all fines 300 μm or smaller from the sand filler. The surface mixture was prepared at a water-cement ratio of 0.6. Three replicates were produced of each sample to account for variability.

**Table 1.** Test factorial for the $NO_2/NO_x$ experiment

| Variables | Content | Levels |
|---|---|---|
| Percentage of $TiO_2$ | 3% and 5% | 2 |
| Aggregate Gradation | with and without fines | 2 |
| $NO_2/NO_x$ ratio | 0, 0.2, 0.45, and 0.70 | 4 |
| Humidity Level (%) | 15, 25, 35, 50, and 80 | 5 |
| Flow Rate (l/min) | 3 and 8 | 2 |

**Results and Analysis**

The influence of operational and design factors ($TiO_2$ content, aggregate gradation, $NO_2/NO_x$ ratio, and flow rate) on the $NO_x$ removal efficiency was quantified. The effect of a given factor was determined based on the following equation:

$$e_f = \frac{\sum_{i=1}^{n} \pm f_i R_i}{4} \quad (1)$$

where $e_f$ = effect of factor f =1, 2, 3, or 4; n = number of experimental design run n = 8, R = $NO_x$ reduction efficiency per run; and ±: sign used to describe low and high levels for each factor. The effect of each factor is shown in Table 2. It is evident from the near zero main effects calculated, $e_1$ for the titanium content and $e_2$ for the aggregate design, that the mix design factors had little impact on $NO_x$ removal efficiency. In contrast, the main effects, $e_3$ and $e_4$, representing the $NO_2/NO_x$ ratio and the flow rate respectively, indicate that there was a decrease in efficiencies with higher $NO_2/NO_x$ ratios and higher flow rates. Furthermore, the flow rate had a larger average negative impact on the $NO_x$ reduction efficiencies compared to the $NO_2/NO_x$ ratio. As a result, further analysis was conducted to better understand the relationship between the $NO_2/NO_x$ ratio and flow impacts on the $NO_x$ reduction efficiency. Results of this analysis are presented in the following sections.

**Table 2.** Main effects of operational and design factors

| Factor | Effect | Rank[1] |
|---|---|---|
| $TiO_2$ Content | -0.08 | 4 |
| Aggregate Gradation | 1.2 | 3 |
| $NO_2/NO_x$ Ratio | -5.0 | 2 |
| Flow Rate | -8.8 | 1 |

[1]: rank of the influence on $NO_x$ reduction efficiency

### *$NO_2/NO_x$ Detailed Analysis*

As it was previously mentioned, an increase in $NO_2/NO_x$ ratio results in a decrease in $NO_x$ reduction efficiencies. To further understand this trend, the $NO_x$ removal efficiency was measured at four different $NO_2/NO_x$ ratios (see Table 1). In this part of the experiment, the flow rate was kept constant at 3 l/min and the relative humidity at 20% (± 5%). Under these conditions, the average reduction was as high as 75% when no $NO_2$ was present and as low as 42% with 70% $NO_2$, see Figure 3. As expected from the main effect results, the $NO_x$ reduction efficiencies did not change substantially between the four mix designs. Meanwhile, it is clear from Figure 3 as the ratio of $NO_2/NO_x$ increases, the overall $NO_x$ reduction efficiencies decrease. When increasing the $NO_2/NO_x$ ratio from 0 to 70%, the $NO_x$ reduction efficiencies dropped 32%, representing 56% of the total efficiency.

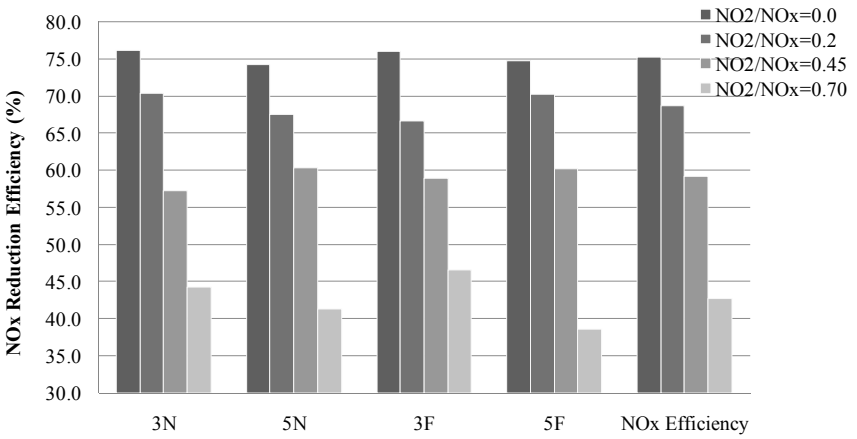



**Figure 3.** Impact of the $NO_2/NO_x$ ratio on $NO_x$ reduction efficiency

One may hypothesize that the reduction in $NO_x$ removal efficiencies is the result of the increase in the $NO_2/NO_x$ ratio. However, as the $NO_2/NO_x$ ratio increased, the initial equilibrium concentration of the $NO_x$ also decreased as shown in Figure 4. The lower initial concentrations are a consequence of the instability of $NO_2$. One explanation of the loss of $NO_2$ is through a reaction between $NO_2$ and the surface of the concrete sample such as heterogeneous hydrolysis. To verify that the initial concentration was not a factor in the decreasing $NO_x$ reduction efficiencies, the initial

concentration was lowered while the ratio was kept constant. This was tested for two different $NO_2/NO_x$ ratios, one without $NO_2$ and one with 20% $NO_2$ under the same test conditions. Figure 5 reveals that the NOx reduction efficiencies remained mostly constant regardless of the initial pollutant concentration.

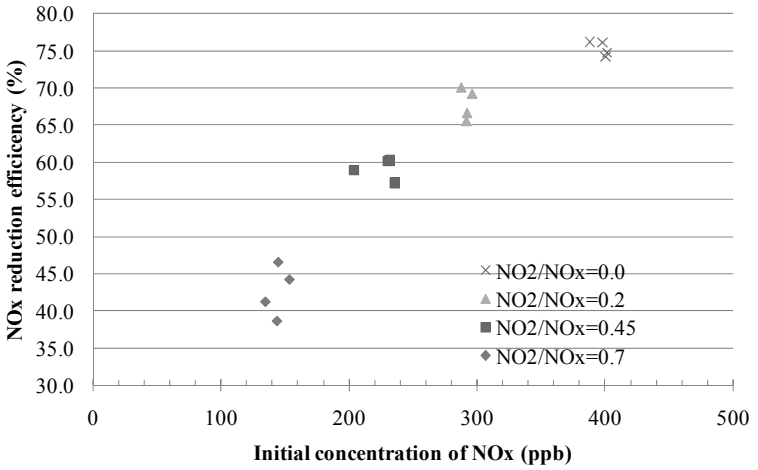



**Figure 4.** Impact of the $NO_2/NO_x$ on the initial concentration on the $NO_x$ reduction efficiency

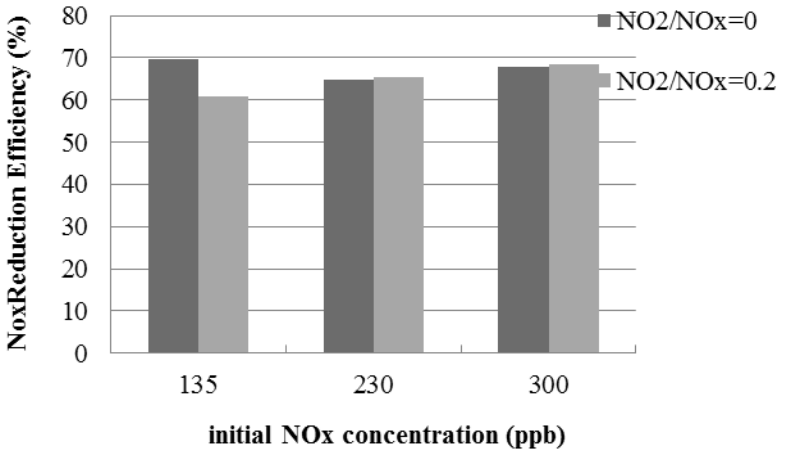



**Figure 5.** Impact of the initial concentration on the $NO_x$ reduction efficiency

### *Flow Detailed Analysis*

As it was previously mentioned, the flow rate had a significant impact on $NO_x$ reduction efficiencies. With faster flow rates, there is less contact time for a reaction to occur, which results in lower reduction efficiencies. To evaluate this trend when $NO_2$ is introduced, the flow was varied from 1, 3, and 8 l/min at four different $NO_2/NO_x$ ratios. Figure 6 shows the relationship between flow rates, $NO_2/NO_x$ ratio and $NO_x$ reduction efficiency. As shown in this figure, the increase in flow rates and $NO_2/NO_x$ ratios resulted in a decrease in $NO_x$ reduction efficiency. It is important to emphasize that the equilibrium concentrations changed when the flow rate changed.

The amount of $NO_2$ reduced from the equilibrium concentration was greater at slower flow rates. In fact, when the inlet stream was completely $NO_2$, the majority of the $NO_2$ was scrubbed at the concrete surface without photocatalytic degradation. For this reason, no $NO_x$ photocatalytic reduction was observed and this point has been omitted in Figure 6. This further supports the theory of a reaction such as heterogeneous hydrolysis taking place between the unstable $NO_2$ and the concrete surface.

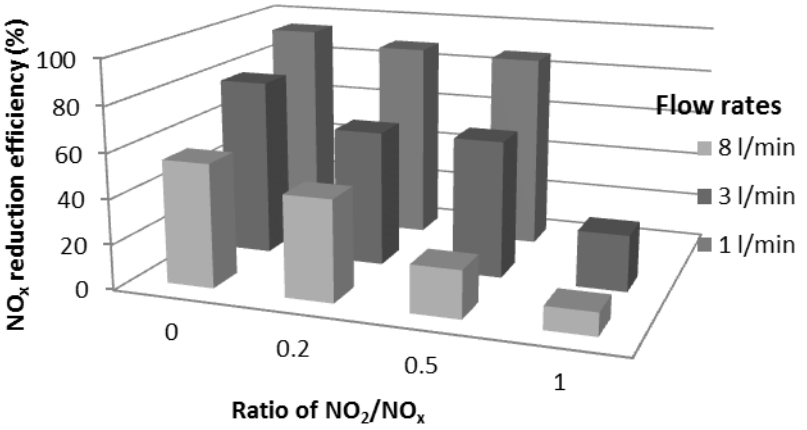



**Figure 6.** Impact of flow on the $NO_x$ reduction efficiency at various $NO_2/NO_x$ ratios

***Effect of Relative Humidity***

The impact of relative humidity on $NO_x$ reduction efficiency was investigated. Figure 7 illustrates the variation of $NO_x$ reduction efficiency with the relative humidity. As shown in this figure, humidity can have both a negative and positive impact on $NO_x$ reduction efficiency where the optimum relative humidity is near 25%. At relative humidity lower than 25% humidity, the lack of water molecules required for the hydroxyl radicals hinders the photocatalytic oxidation. At higher relative humidity, the water molecules interfere with $NO_x$ contact to the $TiO_2$ active sites on the surface (Dylla et al 2010).

***Discussion of the Experimental Results***

In spite of the rapid development of photocatalytic compounds, current applications of this technology are limited to building facades and gateway elements of bridges not subjected to traffic. One promising method to expand the reach of this technology is to apply titanium dioxide to concrete pavement surface to result in air-purifying concrete pavement. However, the effects of a number of operational parameters are still unclear.

Since past studies have determined that $NO_2/NO_x$ ratio typically ranges from 10 to 30% in urban areas, it is important to determine the photocatalytic efficiency in degrading $NO_2$. While results of the experimental program determined that increasing $NO_2/NO_x$ ratio reduce the photocatalytic efficiency, most of the $NO_2$ is being scrubbed through a heterogeneous hydrolysis reaction occurring at the concrete

surface. The increase in relative humidity above 25% was also found to hinder the efficiency of photocatalytic activity.

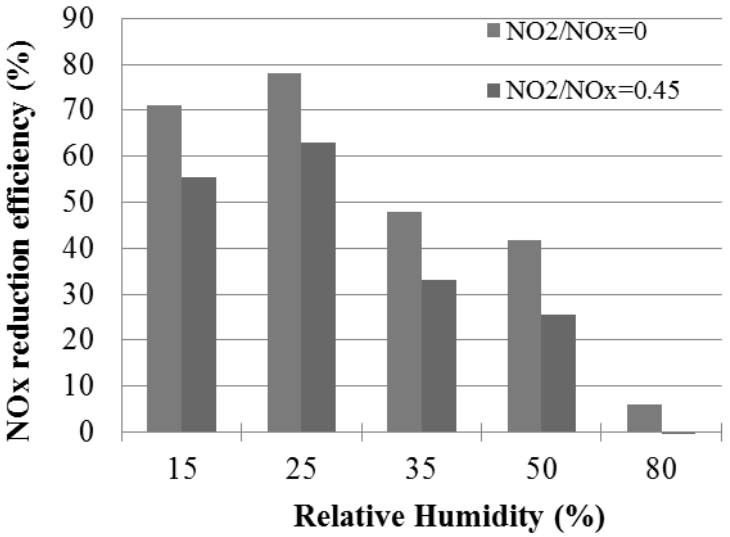



**Figure 7.** Impact of the relative humidity on the $NO_x$ reduction efficiency according to the ratio of $NO_2/NO_x$

## Conclusions

Nitrogen dioxide ($NO_2$) is a highly reactive gas that is usually not considered in the evaluation of $TiO_2$ photocatalytic activities. The unstableness of $NO_2$ is exhibited in the increasing variability in $NO_x$ reduction efficiencies as more $NO_2$ is introduced. The objective of this study was to evaluate the impact of $NO_2$ in the photocatalytic process of $NO_x$. Based on the results of the experimental program, it is determined that the $NO_2/NO_x$ ratio has a significant impact on $NO_x$ reduction efficiencies; however, the extent of this impact relies on many other factors such as flow rate.

The highest photodegradation rate was observed at 25% relative humidity, which balances the availability of hydroxyl radicals at the surface with $NO_x$ contact with the photocatalytic surface. Results of the experimental program showed that increasing the flow rate and $NO_2/NO_x$ ratio negatively affect the effectiveness of the $NO_x$ reduction efficiency. This is a result of most of the $NO_2$ being scrubbed through reaction with the concrete surface. However, within the evaluated range, the titanium content and aggregate gradation had little impact on $NO_x$ removal efficiency. Further research should be conducted to confirm the heterogeneous hydrolysis reaction occurring at the concrete surface.

## Acknowledgements

The authors would like to acknowledge the assistance of Cristal Millennium for providing this research with PC 105 ultrafine titanium dioxide used for testing and the Louisiana Transportation Research Center (LTRC) for granting access to their laboratory. The authors would also like to acknowledge the assistance of the Louisiana Department of Environmental Quality.

**References**

Beeldens, A. (2008). "Air purification by pavement blocks: final results of the research at the BRRC." *Proc., Transport Research Arena Europe*, Ljubljana.

Cassar, L., Pepe, C., Tognon, G., Guerrini, G.L., and Amadelli, R. (2003). "White cement for architectural concrete possessing photocatalytic properties." *Proc. 11th Int. Congr. On the Chemistry of Cement*. Durban, South Africa.

Dylla, H., Hassan, M. M., Mohammad, L., and Rupnow, T., (2010). "Evaluation of the environmental effectiveness of titanium dioxide photocatalyst coating for concrete pavements." *Journal of the Transportation Research Record*, National Research Council, Washington, D.C., accepted.

Environmental Protection Agency (EPA). (2010). "Final revisions to the primary national ambient air quality standard for nitrogen dioxide (NO2)." <http://www.epa.gov/airquality/nitrogenoxides/pdfs/20100124presentation.pdf> (June 21, 2010).

Hassan, M.M., Dylla, H., Mohammad L., and Rupnow T., (2010). "Evaluation of the durability of titanium dioxide photocatalyst coating for concrete pavement." *Journal of Construction and Building Material*, Elsevier, accepted.

Jimenez, J.L, Koplow, M.D., Nelson, D.D., Zahniser, M.S., and Schmidt, S.E. (1999). "Characterization of on-road vehicle NO emissions by a TILDAS remote sensor." *J. Air & Waste Manage. Assoc.,* 49, pp. 463-470.

Fujishima, A., Rao, T.N., and Tryk, D.A. (2000)."Titanium dioxide photocatalysis." *Journal of Photochemistry and Photobiology C: Photochemistry Reviews.* 1, pp. 1-21.

McConnell, R., Islam, T. , Shankardass, K., Jerrett, M., Lurmann, F., Gilliland, F., Gauderman, J., Avol, E., Kuenzli, N., Yao, L., Peters, J., and Berhane, K.. (2010). "Childhood incident asthma and traffic-related air pollution at home and school." *National Institute of Environmental Health Sciences*, pp. 1-33.

Poon, C.S., and Cheung, E. (2007). "NO removal efficiency of photocatalytic paving blocks prepared with recycled materials." *Construction and Building Materials*, 21, pp. 1746-1753.

Ramazan, K.A., Syomin, D., and Finlayson-Pitts, B.J. (2004). "The photochemical production of HONO during the heterogeneous hydrolysis of NO2." *Phys. Chem. Chem. Phys.*, 6, pp. 3836-3843.

Tao, Z., Hewings, G., and Kieran, D. (2010). "An economic analysis of Midwestern US criteria pollutant emissions trends from 1970 to 2000." *Ecological Economic*, 69, pp. 1666-1674.

Thermo Scientific. (2008). "Model 146*i* Instruction Manual." Thermo Fisher Scientific Inc., Franklin, MA.

Zhao, J., and Yang, X. (2003). "Photocatalytic oxidation of indoor air purification: a literature review. *Building and Environment*, 38, pp. 645-654.

# Energy Harvesting from Pavements using Pyroelectric Single Crystal and Nano-Composite Based Smart Materials

Sudip Bhattacharjee[1], Ashok K. Batra[2], and Jacob Cain[3]

[1]Assistant Professor, Department of Civil Engineering, Alabama A&M University, P.O. Box 367, Normal, AL, 36762, Email: sudip.bhattacharjee@aamu.edu, Ph: (256)-372-4148, Fax: (256)-372-5909
[2]Assiatant Professor, Department of Physics, Alabama A&M University, P.O. Box 1268, Normal, AL 35762, Email: ashok.batra@aamu.edu, Ph: (256)-372-8109
[3]Undergraduate Research Assistant, Department of Civil Engineering, Alabama A&M University, P.O. Box 367, Normal, AL, 36762, Email: jacobc_58@yahoo.com

## ABSTRACT

Energy harvesting (also known as power harvesting or energy scavenging) technology allows capturing unused ambient energy such as solar, wind, thermal, strain and kinetic, energy of gas and liquid flows and converting the energy into another form of usable energy. The goal of this study was to determine if there is a possibility of capturing thermal energy from pavements and storing the energy and using it as alternative power source to other devices. In this paper, pyroelectric materials have been used as the energy harvester. A single crystal based material called Lithium Tantalate ($LiTaO_3$), and a polycrystalline composite material based on ordinary Portland cement with carbon nano-fibers have been used as pyroelectric smart materials. The present research indicates that simple material such as Portland cement can be used to capture heat energy from pavements, which can be stored in capacitor for the use of power source to other sensor electronics.

## INTRODUCTION

Energy harvesting (also known as power harvesting or energy scavenging) technology allows capturing unused ambient energy such as vibration, strain, light, temperature gradients, and/or changes, energy of gas and liquid flows and converting into usable electrical energy which is stored for performing sensing and actuation and other applications. Energy harvesting is a perfect match for wireless devices and wireless networks which otherwise depend on battery power and other maintenance-free systems; thus, providing cost-effective and environment friendlier solutions for various applications. Energy can be recovered from mechanical vibration (Sodano et al 2004), light, and spatial and temporal temperature variations (Roundy et al 2003, Paradiso and Stamer 2005).

Thermal energy in the environment is a potential source of energy for low-power electronics. One of the sources of thermal energy is from pavements. The pyroelectric effect converts a temperature change into electrical current or voltage. In pyroelectric materials, the thermal energy is converted to electrical energy. Charge is produced when the material's temperature is altered as a function of time. Recent work was presented on using Lead Zirconate Titanate (PZT), PMN-PT, Polyvinylidene Fluoride (PVDF) and other pyroelectric materials for energy harvesting and storage (Bauer 2006).

If such electric energy can be trapped, it can be used for potential energy harvesting purposes, particularly from pavements. There have been several studies performed in this emerging field recently. In a recent paper Cuadrasa et al (2010) have showed the potential of using pyroelectric materials as energy harvester. The authors tested several PZT and PVDF as potential pyroelectric material under various temperature conditions. The research results indicate that both PZT and PVDF can be used as pyroelectric material.

Novel applications require diverse and specific properties in materials which cannot be met in single-phase materials. Composites contain two or more chemically different materials or phases. In these materials, it is possible to tailor electrical, piezoelectric, pyroelectric, optical, and mechanical properties catering a variety of applications in civil engineering.

Wen and Chung (2003) investigated the pyroelectric behavior of cement-based materials. They showed that the steel/carbon-nanofibers increase the dielectric properties of cement composites. However, there are limited studies on the pyroelectric and dielectric properties of cement-based nano-composites for infrared sensing and energy harvesting and other civil engineering applications. Therefore further electrical and pyroelectric investigations in cement-based composites, including nano-composites, are warranted for the purpose of fundamental understanding of their behavior.

Piezoelectric materials have high potential of being used in the field for energy harvesting, especially under high volume traffic conditions. However, limited study and applications can be found in energy harvesting from pavements. In United Kingdom, PaveGen Systems Inc has developed piezoelectric materials based energy harvesting system from human walking on pavements. The most serious attempt to generate electricity from vehicle movements has come from Innowattech in Israel, where they have installed piezoelectric devices under one kilometer length of pavement lanes. The maximum generated power is 250 kWh from 600 v/h traffic flow. Similar devices have been installed also in Tokyo Railway Station to harness 1400 kW/sec energy from human movements.

Specially, use of pyroelectric materials in energy harvesting from pavements is missing in current literature. If the pavement temperature can be converted into electrical current, it would be ideal choice for using alternative sustained power source to wireless sensors and other devices which require low but sustain power. This would ensure sustained and uninterrupted power supply to pavement management system hardware and will contribute to the energy conservation. In this paper, the pyroelectric properties of Lithium Tantalate ($LiTaO_3$) and cement composites with carbon nano-fibers and Polyvinyl Alcohol (PVA) consisting of

ordinary Portland cement as matrix are reported. The influence of carbon nano-fibers on the dielectric and pyroelectric properties of the composite has been examined for admixtures prepared under identical temperature and humidity conditions. The possibility of harvesting electrical energy from pavement temperature variation is also discussed.

## MATERIALS AND TEST METHOD

Two different pyroelectric materials have been used in this study. One such material is Lithium Tantalate ($LiTaO_3$). $LiTaO_3$ is a crystalline solid which possess both pyroelectric and piezoelectric properties. It has excellent mechanical stability making it a good candidate for piezoelectric materials as well, which can be used under heavy loading. It has Curie temperature of 601±5.5 $^oC$ and pyroelectric coefficient of $2.3\times10^{-4}$ $C/^oC/m^2$ (Crystal Technology, Inc). $LiTaO_3$ crystals were grown in the laboratory using Czochralski technique and later tested for the electrical properties. The pyroelectric property of $LiTaO_3$ was used for energy conversion and storage. $LiTaO_3$ was selected among other materials because it has better pyroelectric properties and is a common pyroelectric material used as ordinary household motion detectors.

Another pyroelectric material chosen was a cement-composite. To prepare the composite specimens, the ordinary Portland cement, PVA and carbon nano-fibers were mixed thoroughly. The PVA acts as a binder within the cement. No aggregate was added. The carbon nano-fibers (PR-19-XT-LHT) of fiber diameter 150 nm and 20-30 $m^2/gm$ surface area were used. The weighted cement mixture was mixed in water to make a paste of suitable viscosity. The paste was compressed in a steel mound to form discs of diameter 10 mm and thickness 2.0 mm approximately. Two electrical parameters were calculated which were relevant to the materials potential application of charge storage capacity: the real ($\varepsilon'$) and imaginary ($\varepsilon''$) parts of dielectric constant and dielectric loss (tan $\delta$) and they are defined as:

$$\varepsilon' = \frac{C_p d}{\varepsilon_0 A} \qquad \varepsilon'' = \varepsilon' \tan\delta \tag{1}$$

where $A$ is the electrode area (identical areas for the opposite electrodes were used in each sample), $d$ is the thickness of the sample, $\omega=2\pi f$ is the frequency of $AC$ measurement, $\varepsilon_0 = 8.854\times10^{-12}$ F/m is the permittivity of vacuum. A series of $AC$ frequency was selected from very low (10 Hz) to as high as 1 MHz to obtain a wide spectrum. The pyroelectric current $I_p$ was also measured at various temperatures and the pyroelectric coefficient ($p$) was calculated using the relationship:

$$p = \left(\frac{I_p}{A}\right) / \left(\frac{dT}{dt}\right) \tag{2}$$

where $A$ is the electrode area and $dT/dt$ is the rate of change of temperature which was kept constant throughout the measurement. The change in pyroelectric coefficient will indicate the change in dipole orientation inside the material; higher the coefficient, better the material is for converting temperature change in electrical charge. The additional charge generated via heating or cooling within a temperature change $dT$ can be calculated as:

$$dQ = dI_P dt = pA\frac{dT}{dt}dt = pAdT \quad (3)$$

where, the pyroelectric coefficient $p$ could be constant or as a function of temperature.

## ELECTRICAL MEASUREMENT RESULTS

Figure 1 (a,b) shows the frequency dependence of dielectric constant ($\varepsilon'$) and loss ($\varepsilon''$) for cement composites reinforced with carbon fibers. All of the composites show a decrease sharply in dielectric constants ($\varepsilon'$ and $\varepsilon''$) with increasing frequency in the range from 100 Hz to 10 kHz. This is mainly attributed to interfacial polarization of the composites and polarization in cement matrix. When frequency is larger than 10 kHz, the composites exhibit relatively good frequency stability. The behavior of the parameters studied can be better explained by knowing the porosity and the moisture content of the composites below percolation. However, the pyroelectric behavior of the cement composite is very similar to the ones observed in other materials (Batra et al 2009, Wen and Chung 2003).

The dependence of dielectric constant ($\varepsilon'$) and pyroelectric coefficient ($p$) of cement composites and $LiTaO_3$ on temperature are shown in Figure 2. As shown, these parameters increase with increase in temperature. The observed increase in dielectric constant in cement composites can be attributed to the increase in mobility of ions as the temperature increases, whereas in the case of $LiTaO_3$, the increase of $p$ is due to the increase in dipolar moments in the material. The figures also indicate that the inclusion of carbon nano-fibers increase the dielectric and pyroelectric coefficient of cement composite. One major conclusion from these figures is: ordinary Portland cement can act as pyroelectric material. The influence of carbon fiber content on $\varepsilon'$ and $p$ of cement composites are plotted in Figure 3a and 3b respectively at 1 kHz frequency and 40$^o$C. It can be seen that the dielectric constant values increase with the increase in carbon nano-fiber content.

## CHARGE STORAGE VIA PYROELECTRIC EFFECT

The conversion of thermal energy into electrical energy is essential in pyroelectric sensors. A large amount of heat absorbed by the pavement is radiated to the atmosphere. It was the goal of this initial study to demonstrate the feasibility of capturing the thermal energy from pavement via pyroelectric effect.

### Feasibility Study

The feasibility of using pyroelectric materials for energy conversion has been investigated using $LiTaO_3$ crystals. Figure 4a shows the circuit used to model pyroelectric sensors used for energy conversion and charge storage (Xie et al 2009). The pyroelectric sensor is modeled as a capacitor ($C$) and resistor ($R$) in parallel with a current source ($i_p$). The current is generated within the cell with the change in temperature. The circuit in the figure can be modeled with the following first order differential equation:

$$C(T)\frac{dV}{dt}+\frac{V}{R}=i_p=p(T)A\frac{dT}{dt};\ \frac{dV}{dt}=\frac{Ap(T)}{C(T)}\frac{dT}{dt}-\frac{V}{RC(T)}\ , \tag{4}$$

where $C(T)$ = capacitance as a function of temperature, $R$ = resistance, $V$ = voltage, $i_p$ = pyroelectric current, $A$ = cross sectional area of the sample, d$T$/d$t$ = temperature gradient, $p(T)$ = pyroelectric coefficient as a function of temperature. This differential equation is then numerically solved with a given temperature profile using Runge-Kutta 4$^{th}$ order numerical integration technique:

$$V_{i+1}=V_i+\frac{1}{6}(k_1+2k_2+2k_3+k_4)h, k_1=f(t_i,V_i)\ k_2=f\left(t_i+\frac{1}{2}h,V_i+\frac{1}{2}k_1h\right)$$

$$k_3=f\left(t_i+\frac{1}{2}h,V_i+\frac{1}{2}k_2h\right), k_4=f(t_i+h,V_i+k_3h)\ , \tag{5}$$

where $f(t,V)=\frac{A}{C(T)}p(T)\frac{dT}{dt}-\frac{V}{RC(T)}$, and $h = \Delta t$, the step size. The pyroelectric coefficient and the capacitance as functions of temperature are obtained from the experiment as shown in Figure 2. For the purpose of this simulation, the temperature is chosen as a linear function of time. At various time instances, the temperature is calculated and the corresponding values of $p$ and $C$ are obtained from Figure 2. The Eq (4) is then numerically solved. Once the voltage is simulated, the power ($P$) is obtained from the relation: $P = V^2/R$. Figure 4b shows the simulation result; the voltage and power as a function of time for $R$ = 40 MΩ. These initial results indicate that the $LiTaO_3$ is a viable candidate material for energy harvesting from pavements.

**Energy Harvesting Simulation**

The idea explored in this section has been proposed by Cuadras et al (2010) for pyroelectric energy harvesting. Eq. (2) indicates that the pyroelectric current is directly proportional to the rate of change of temperature. However, one problem with this behavior is that the current will flow in opposite direction when the rate changes from positive to negative or from negative to positive. In other words, a heating followed by cooling or cooling followed by heating will produce charge accumulation in different direction. However, to charge an external capacitor it is essential that the capacitor be charged continuously.

To mitigate the problem, a full bridge diode rectifier circuit, as shown in Figure 5a, can be used. The charge pyroelectric source is connected to an external capacitor ($C_E$) and an external resistor ($R_E$). There are two pairs of diodes; one pair is used for each direction of current flow. Diodes $D_1$-$D_2$ are used when current is flowing in one direction and $D_3$-$D_4$ are used when the current flows in other direction. At each time only the forward biased diodes work, the other two pairs blocks current flow under reverse biased condition. As it can be seen, in both cases, the external capacitor ($C_E$) is charged via charge flow in one direction and that causes the voltage to increase across the external storage capacitor.

The charge flow in the two-capacitor system can be modeled as follows: When the new charge is accumulated, it is distributed in both capacitors and the charge balance equation can be written as (Cuadras et al 2010):

$$\Delta Q_n = Q_{E,n} - Q_{E,n-1} + Q_{P,n} \pm Q_{P,n-1} \quad (6a)$$

$$Q_{P,n} = V_n C_P,\ Q_{P,n-1} = V_{n-1} C_P,\ Q_{E,n} = V_n C_E,\ Q_{E,n-1} = V_{n-1} C_E\ , \quad (6b)$$

where $C_p$ and $C_E$ are pyroelectric cell capacitance and external charging capacitance respectively, $V_n$ and $V_{n\text{-}1}$ are voltage at temperature data points $n$ and $n$-1 respectively, $Q_p$ and $Q_E$ are charge accumulated in pyroelectric cell and external capacitance respectively, and $\Delta Q_n$ is the additional charge generated at $n^{th}$ data point (from heating or cooling). The ± sign in front of the right hand side term indicates that charge stored in pyroelectric cell can be in opposite direction if the sign of the rate of change of temperature changes from $(n\text{-}1)^{th}$ data point to $n^{th}$ data point. Substitution of Eq. (6b) into Eq. (6a) results the following recurrence equation, from which the voltage across the external capacitance can be calculated at a given temperature data point:

$$V_n = \frac{\Delta Q}{C_P + C_E} + \left(\frac{C_E \pm C_P}{C_E + C_P}\right) V_{n-1} = \frac{pA\Delta T}{C_P + C_E} + \left(\frac{C_E \pm C_P}{C_E + C_P}\right) V_{n-1} \quad (7)$$

Once the voltage is determined, the energy stored at $n^{th}$ data point can be calculated from the following equation: $E_n = 0.5 C_E V_n^2$.

Eq. (7) has been used to simulate the voltage produced from a measured temperature profile of actual pavement temperature. Figure 5b shows the temperature profile between May-Oct at a station location in Huntsville, AL. The temperature profile was obtained from the Environmental and Climatic Database of Mechanistic Empirical Pavement Design Guide (MEPDG) (NCHRP 2004). The simulated voltage and energy stored are shown in Figure 6 for $LiTaO_3$ and cement composites. The circuit is simulated with the following values: $C_P = 3.761\times10^{-12}$ F, $C_E = 10\times10^{-6}$ F, $p = 1.5\times10^{-9}$ C/cm$^2$/$^o$C, $A = 0.907$ cm$^2$. When $C_E >> C_P$, Eq. (7) reduces to: $V_n \approx (pA\Delta T)/C_E + V_{n\text{-}1}$. This is the simplified version of Eq. (7); when storing energy from temperature fluctuations, the values of $C_P$ and $C_E$ should be optimized for highest energy storage. Furthermore, the external resistor, $R_E$, and the external capacitor, $C_E$, can be adjusted to obtain fast charging. Figure 6a shows that the accumulated voltage increases as the summer months go by. Figure 6b shows the accumulated energy (Jules). The maximum voltage accumulated at the end of October is around 0.55 V. It should be noted that this maximum voltage can be controlled by choosing suitable value for external storage capacitor, $C_E$.

It is possible to increase the voltage by placing several of the pyroelectric cells in series combination. One single device can produce 1 mV instantaneous voltage, and if 1000 of such devices are connected via a typical 1kΩ resistance sensor device, the supplied current would be 1 mA, which can be used to power wireless sensors for bridge maintenance. It is to be noted that preliminary test conducted by Innowattech in Israel produced 250 kWh energy from 1 km length of highway for 600 v/h traffic volume. Currently, the authors are testing the methodology by using several of these sensors in real pavement and monitoring the voltage generated. The present experimental results are in agreement with the predicted results. A typical configuration of the sensors in real pavement with data acquisition system is shown in Figure 7.

## CONCLUSIONS

In this study Lithium Tantalate and cement-based pyroelectric nano-composites have been fabricated and their electrical and pyroelectric properties have been measured. The material data has been used to perform simulation with real pavement temperature data obtained from climatic database of MEPDG. The preliminary experimental results obtained from composite samples can be summarized as follows: (a) Both Lithium Tantalate and cement composites with nano-fibers are viable materials which can be used for energy conversion from pavement structures. (b) Lithium Tantalate crystals show very good pyroelectric behavior which makes it excellent candidate for energy harvester. (c) Cement-Carbon nano-composites can be fabricated using normal mixing and compaction method, thus making it an attractive candidate for energy harvester. (d) The dielectric constant of cement composite increases as the carbon nano-fiber content increases. (e) The pyroelectric coefficients of the composite also increase as carbon fiber content increases. Higher content of carbon fiber in the cement nano-composites is beneficial for pyroelectric energy conversion devices. (f) The simulation result with the real time pavement temperature data indicates that the methodology is a sound one and should be explored further with real time pavement experimentation.

## ACKNOWLEDGEMENTS

The partial financial support for this work through NSF-RISE grant ID 0927644 is gratefully acknowledged. Special thanks to Drs. Mohan Aggarwal and M. A. Alim for support in our work.

## REFERENCES

Batra, A. K., Alim, M. A., Currie, J. R., and Aggarwal, M. D. (2009). "The Electrical Response of the Modified Lead Titanate Thick Films." *Physica B*, 404, 1905-1911.

Cuadras, A., Gasulla, M., and Ferrari, V. (2010). "Thermal Energy Harvesting Through Pyroelectricity." *Sensors and Actuators A*, 158, 132–139.

National Cooperative Highway Research Program (NCHRP). (2004). "Guide for Mechanistic Empirical Design of New and Rehabilitated Pavement Structures: Final Report." *NCHRP Project No. 1-37A*, Transportation Research Board, National Research Council, Washington DC, USA.

Paradiso, J. A., and Starner, T. (2005). "Energy Scavenging for Mobile and Wireless Electronics." *Prevasive Computing*, January - March, 18 – 27.

Roundy, S., Wright, P., and Rabaey, J. (2003). "A Study of Low Level Vibrations as a Power Source for Wireless Sensor Nodes." *Computer Communications,* 26(1),1131 –1144.

Sodano, H. A., Inman, D. J., and Park, G. (2004). "A Review of Power Harvesting from Vibration using Piezoelectric Materials." *The Shock and Vibration Digest,* 36, 197 –205.

Wen, S., and Chung D. D. L. (2003). "Pyroelectric Behavior of Cement-Based Materials." *Cement and Concrete Research*, 33, 1675-1679.
Xie, J., Mane, X. P., Green, C. W., Mossi, K. M., and Leang, K. K. (2009). "Performance of Thin Piezoelectric Materials for Pyroelectric Energy Harvesting." *J. Intelligent Mat. Sys and Struct*, 0-2009.
Pavegensystems. <http://www.pavegensystems.com/> (November 1, 2010).
Innowattech. <http://www.innowattech.co.il/> (November 1, 2010).
JREST. <http://www.jreast.co.jp/e/> (November 1, 2010).

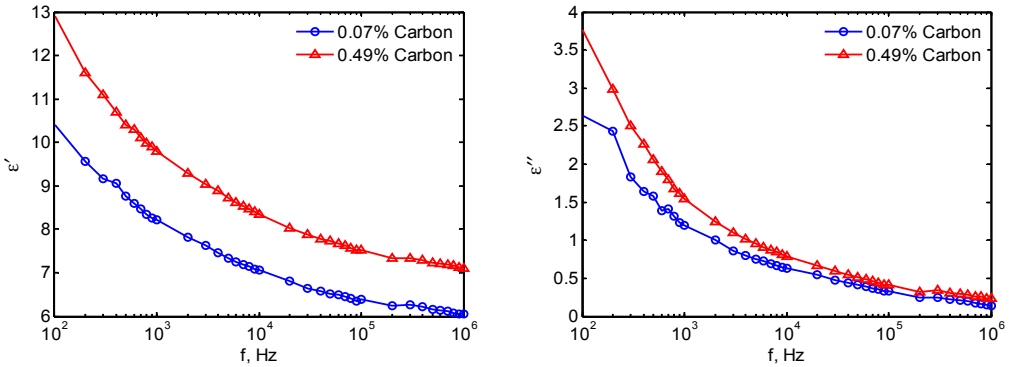



Figure 1. (a) Typical dependence of (a) dielectric constant ($\varepsilon'$) and (b) dielectric loss ($\varepsilon''$) of cement:carbon nano-composites on frequency (40 °C).

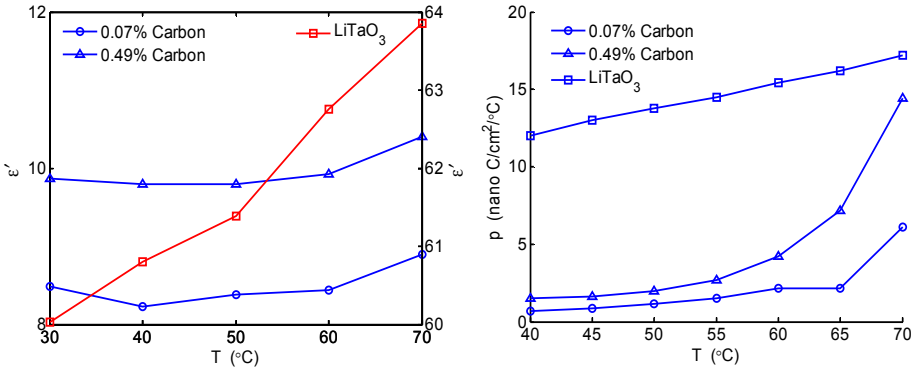
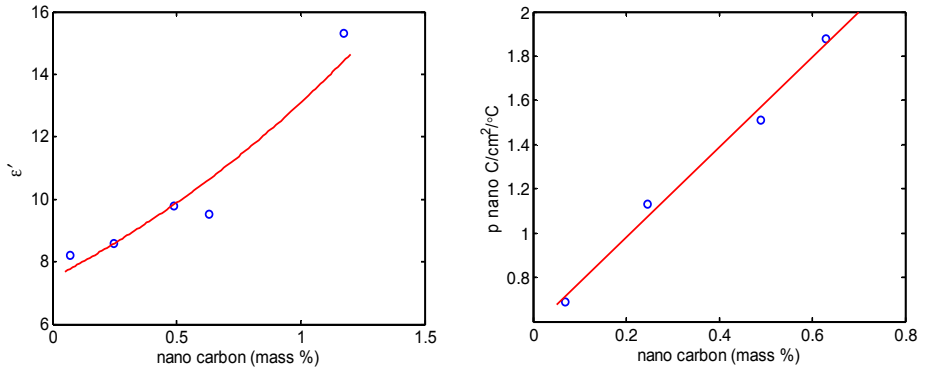
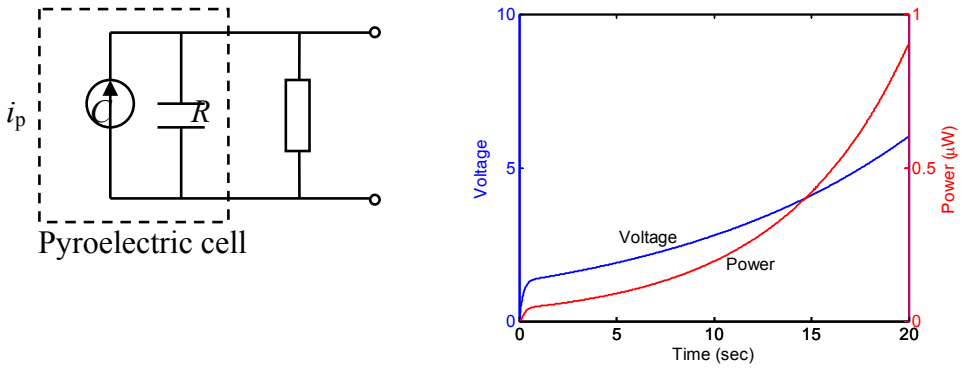



Figure 2. Typical dependence of (a) dielectric constants ($\varepsilon'$) and (b) pyroelectric coefficient of $LiTaO_3$ and cement-carbon nano-composites on temperature.



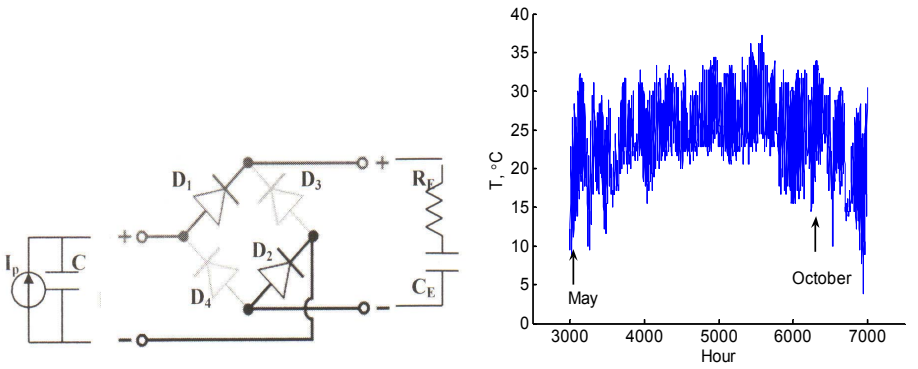

(a) (b)

Figure 3. Dependence of the (a) dielectric constants ($\varepsilon'$) and (b) pyroelectric coefficient of cement-carbon nano-composites on the carbon content (40 °C and 1 kHz frequency).



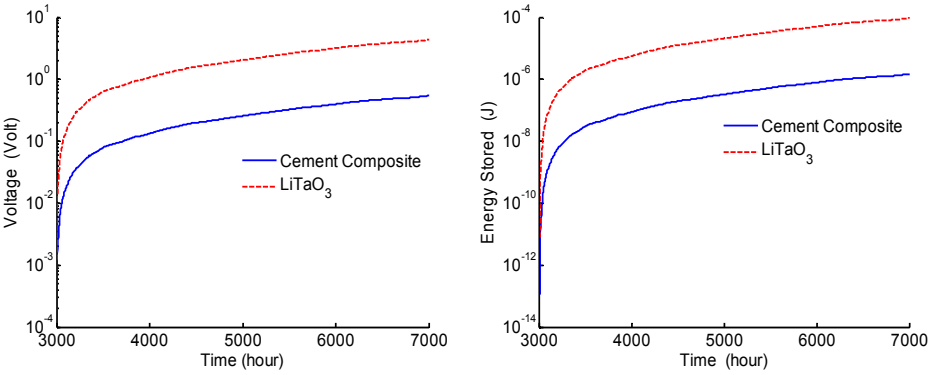

(a) (b)

Figure 4. (a) Circuit model of pyroelectric material (Xie et al 2009); (b) Voltage and Power versus time generated from $LiTaO_3$ single crystal pyroelectric sensor.




(a) (b)

Figure 5. (a) Pyroelectric cell with full bridge rectifier circuit for charge storage; $D_1$-$D_2$ are used in one direction of current flow (solid arrow), $D_3$-$D_4$ are used in other direction of current flow (dotted arrow) (Cuadras et al 2010); (b) Temperature profile at Huntsville, AL, May-Oct, Station ID = 03856.




(a) (b)

Figure 6. (a) Generated voltage and (b) Energy (in Joules) across the external storage capacitor from a single pyroelectric device.

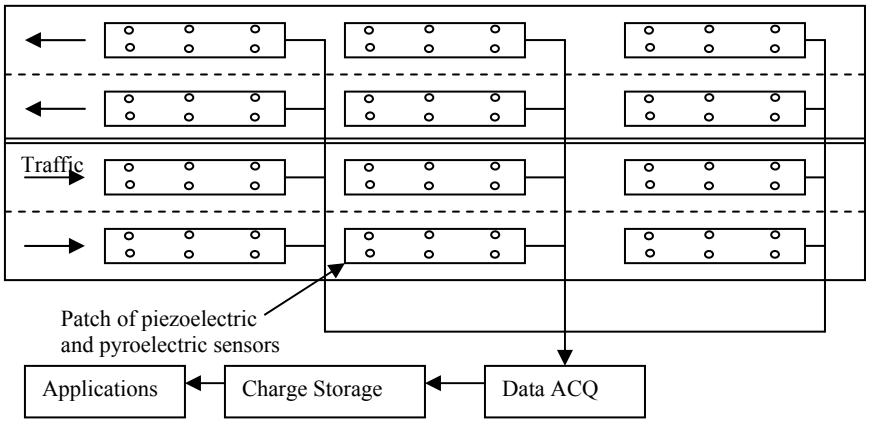



Figure 7. Combination of pyroelectric sensors under pavement layer for power harvesting from pavement.

# Laboratory Evaluation of Warm Mix Asphalt and Asphalt Mixtures with Recycled Materials[a]

Qazi Aurangzeb[1], Jeff Kern [2], Hasan Ozer [1] Imad L. Al-Qadi[3], Thomas G. Zehr [4], James S. Trepanier[5], and William J. Pine[6]

[1] Graduate Research Assistant, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, aurangz1@illinois.edu
[2] Research Engineer, Illinois Center for Transportation, Rantoul, IL, 61866, jeffkern@illinois.edu
[3] Founder Professor of Engineering, Director of Illinois Center for Transportation, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, alqadi@illinois.edu
[4] HMA Implementation Engineer, Bureau of Materials and Physical Research, IDOT, Springfield, IL, 62704, Thomas.Zehr@illinois.gov
[5] HMA Operations Engineer, Bureau of Materials and Physical Research, IDOT, Springfield, IL, 62704, James.Trepanier@illinois.gov
[6] Research Engineer, Heritage Research Group, Indianapolis, IN, 46268, Bill.Pine@heritage-enviro.com

**ABSTRACT:** Transportation infrastructure poses significant challenges to sustainable development because its construction and maintenance require large amounts of material consumption, energy input, and capital investment. A refined systematic approach to the design, construction, and management of pavement systems is necessary to meet the objectives of sustainable development. The major components of sustainable development in pavement systems include economic and social considerations in addition to environmental preservation. Asphalt mixture design and construction practices are essential to the environmental component of sustainable development. These practices include, but are not limited to, asphaltic mixture reusing and recycling as well as innovative technologies considering energy efficiency and stormwater management. Performance and durability become a critical issue with the changes introduced to the mixtures. This paper briefly summarizes the evaluation of mixture design practices and the performance of asphalt mixtures prepared with various types of recycled materials and warm mix technologies.

## INTRODUCTION

Sustainable pavement systems are defined as a network of high-quality, long-lasting

pavements whose design, construction, and management take into account economic and social development, and environmental preservation. Sustainable practices in the pavement industry target increased use of recycled materials and encourage innovations in the design and construction of asphaltic mixtures. For example, long-lasting pavement, also known perpetual or thick pavement, is a design strategy to achieve sustainable pavements that have minimum repair cost. Asphaltic mixtures with reclaimed asphalt pavements (RAP), recycled shingles (RAS), warm mix technologies, and porous asphalt mixes are examples of the most common sustainable practices.

Recycled products and RAP constitute the main ingredients of the alternative mixtures. Virgin aggregates and binder are often partially replaced by these recycled products, whose disposal in pavements is more compatible with the environment and minimizes the impact on natural resources. In another practice, warm mix asphalt (WMA) technology reduces mixing and compaction temperature to minimize energy consumption and green house gas emissions. These practices, in general, are attempts to reduce the impact on the environment by saving energy during construction and reducing the utilization of natural resources. However, sustainable development can only be achieved when the pavements constructed with these materials and methods can properly withstand traffic and thermal loads. Therefore, the major challenge is that the prepared mixtures achieve the desired performance and design volumetrics while addressing environmental concerns. Eventually, short- and long-term performance and field volumetrics are inextricably linked; insufficiently compacted mixes in the field are more prone to develop segregation, cracking, and other distresses. The objective of this study is to evaluate the volumetric properties and the laboratory performance of the asphaltic mixtures designed with alternative materials and technologies.

The first section of this paper describes the ongoing efforts to design asphaltic mixtures with a high percentage of RAP, presents an aggregate packing theory utilized in the design phase, and includes results from a moisture damage test on the RAP mixes. The next section presents highlights from the performance tests conducted on the alternative mixes.

## ACHIEVING VOLUMETRICS IN THE HIGH RAP MIXES

Reclaimed asphalt pavement (RAP) has been used in asphaltic pavements for almost four decades. Apart from hot-in-place recycling (HIPR) or Cold-in-Place Recycling (CIPR) which can utilize 80% to 100% RAP, the percentage of RAP in conventional asphaltic mixture design seldom increases above 20-25%. One of the main causes of RAP's limited amount of use is the variability in aggregate gradation introduced with RAP. In addition, the high percentages of fines in RAP and the need for overheating virgin aggregates pose challenges to mixture design and production. The current economic crisis coupled with environmental concerns has forced the departments of transportation (DOTs) in the United States to increase the amount of RAP in flexible pavements. At asphalt plants, measures like splitting the RAP into two stockpiles, one below and one above 9.5 mm sieve, have been taken to increase the RAP percentages. Other plants conduct further fractionation of RAP to achieve

better-controlled gradation.

In line with the nationwide attempts to increase the RAP usage in new flexible pavements, the Illinois Department of Transportation (IDOT) is exploring the use of higher percentages of RAP in asphaltic mixtures. The Illinois Center for Transportation (ICT) has been working on a high-RAP project to provide guidelines for selecting the amount of RAP that can be safely used in asphaltic mixtures without sacrificing quality compared to standard mixes.

Voids in mineral aggregates (VMA) play an important role in the performance of flexible pavements. It is defined as the volume of intergranular void space between the aggregate particles of compacted paving mixture that includes the air voids and the effective asphalt content (Asphalt Institute, 2001). The thickness of asphalt film around aggregate particles is approximated by the aggregate gradation and volume of asphalt binder. Durability and stability are the two asphaltic mixture problems caused by low and high VMA, respectively. AASHTO M323-04 includes a warning that the mixtures prepared with VMA values more than 2.0% above the specified minimums might be prone to rutting and flushing.

The aggregate packing is influenced by gradation, shape, surface texture, and strength of aggregate, and type and amount of compactive effort (Vavrik et al., 2002). The major challenge with increased use of RAP in asphaltic mixtures is the potential for an increase or decrease in the VMA due to the changes in aggregate gradations as well as, to a higher extent, aggregate characteristics, i.e. shape, texture, and strength of the RAP aggregates.

Al-Qadi et al. (2009) investigated the effects of the amount of RAP on the volumetric and mechanical properties of asphaltic mixtures using six different job mix formulae (JMFs) designed with two materials. It was observed that the optimum asphalt content for the mix designs with different percentages of RAP was not significantly changed. Voids in mineral aggregate at optimum asphalt content could have opposite trends for two materials. While for one material, VMA decreased with increased RAP percentage; for the other material, VMA increased with increased RAP percentage. In another study by West et al. (2009), VMA showed a decreasing trend with increasing RAP percentage. The optimum asphalt contents of the mixtures were also decreased by 1% with an increase in RAP from 0 to 45%. Kim et al. (2009) also demonstrated similar results, i.e. decrease in optimum asphalt content and VMA with an increase in RAP percentage. The study by Mogawer et al. (2009) showed the same trend as well.

Daniel and Lachance (2005) observed some contrary results from their study on RAP. They observed that the VMA and VFA of the mixtures with RAP increased at 25% and 40% levels. According to the authors, the difference between the VMA values was due to the extent of blending of the RAP material with the virgin materials. They suggested an optimum heating time for the RAP material to allow for the greatest extent of blending between the virgin and RAP materials. Hajj et al. (2008) also observed similar increasing trends in VMA and VFA with increased RAP percentages.

This study used an aggregate packing theory to better understand and control the asphalt mixture volumetrics with RAP. In addition, very stringent protocols have been developed to ensure quality control of the laboratory mixes prepared with RAP.

As stated earlier, gradation is one of the key factors influencing aggregate packing. To gain some control on variability caused by the RAP gradation, each RAP (plus 9.5 mm and minus 9.5 mm) material was fractionated over various sieve sizes. Agglomerates (RAP chunks retained on 19 mm sieve) were heated at about 80 $^{o}$C for 20-25 min and then broken down by hand. The broken material was again fractionated using all considered sieves. The apparent gradation, obtained by fractionating the RAP, was used to blend the RAP back to make the extraction and Rice Gravity ($G_{mm}$) samples. Extraction of RAP was conducted at IDOT's facility in Springfield, Illinois. The extracted/recovered aggregate gradation was used in the asphalt mixtures. The $G_{mm}$ determined was used to find the effective specific gravity ($G_{se}$). In this study, bulk specific gravity ($G_{sb}$) of the RAP aggregates was then calculated using the following empirical relationship. IDOT, on the other hand, uses 0.1 as the reduction factor to determine bulk specific gravity of RAP aggregates.

$$G_{sb}(RAP) = G_{se}(RAP) - 0.075 \quad (1)$$

The Bailey method (Vavrik et al., 2002) was used to develop 0% (control), 30%, 40%, and 50% RAP mix designs. The Bailey method is a very efficient and automated approach to design an asphalt mix. The approach is based on the aggregate packing theory. It also provides useful insight into the aggregate packing effect on asphalt mixture volumetrics. The Bailey method was primarily developed for virgin aggregate mix design. The unit weight of aggregates, measured as part of the Bailey method, takes into account the effects of aggregate gradation, texture, shape and size, and compaction effort. The unit weight test is not performed on RAP and mineral filler.

After calculating the mix characteristic using Bailey method for a mix with new aggregates, different percentages of RAP were added and the new aggregate percentages were altered such that the new combined blend containing RAP has the same percent passing the primary control sieve (PCS) as the control mix. Primary control sieve is defined as the closest sieve size to the result of 0.22 x NMAS (Nominal Maximum Aggregate Size). For example, for a 19.0 mm NMAS mixture, PCS is No.4 sieve. To maintain the desired split of coarse and fine aggregate, the percent passing the PCS were kept approximately the same with the inclusion of RAP.

Table 1 shows the design mix formula (DMF) and the extracted aggregate gradation of $G_{mm}$ samples with 30%, 40%, and 50% RAP. Due to using stringent specimen preparation protocols and the Bailey method, a very good control on gradation variability was achieved during this study.

As described earlier, researchers have faced considerable problems in achieving the required VMA values with RAP mixes. Following a better aggregate and RAP processing procedure and with the help of the Bailey estimation process, similar values of VMAs were achieved for all of the mixes with RAP. Table 2 shows the VMA values for the control and RAP mix designs. This approach ensures that the difference in performance of the four mixtures will be independent of VMA variability.

Table 1 Comparison between Target and Achieved Aggregate Gradation.

| Sieve Size/No | 30% RAP | | 40% RAP | | 50% RAP | |
|---|---|---|---|---|---|---|
| | DMF[1] | Extracted | DMF | Extracted | DMF | Extracted |
| 25mm | 100.0 | 99.5 | 100.0 | 100.0 | 100.0 | 100.0 |
| 19mm | 96.1 | 95.9 | 96.4 | 96.2 | 96.6 | 97.4 |
| 12.5mm | 75.9 | 76.3 | 77.8 | 77.9 | 79.1 | 79.7 |
| 9.5mm | 63.7 | 64.8 | 65.6 | 65.8 | 66.6 | 67.4 |
| No.4 | 38.0 | 38.4 | 37.9 | 38.4 | 37.3 | 37.8 |
| No.8 | 23.2 | 23.4 | 22.5 | 22.7 | 21.7 | 22.0 |
| No.16 | 16.2 | 16.3 | 16.3 | 16.5 | 16.2 | 16.3 |
| No.30 | 12.4 | 12.6 | 12.8 | 13.1 | 13.1 | 13.4 |
| No.50 | 9.4 | 9.6 | 9.9 | 10.1 | 10.1 | 10.5 |
| No.100 | 6.8 | 7.0 | 7.1 | 7.5 | 7.2 | 7.6 |
| No.200 | 5.4 | 5.7 | 5.7 | 5.7 | 5.8 | 6.2 |

[1]DMF: Design Mix Formula

Table 2 VMA and TSRs for Control and Mixes with RAP Designs.

| Mix | Control | 30% RAP | 40% RAP | 50% RAP |
|---|---|---|---|---|
| VMA (%) | 13.7 | 13.6 | 13.7 | 13.7 |
| TSR | 90.2 | 93.4 | 89.7 | 99.9 |

IDOT's moisture damage test was conducted as part of the mix design. Tensile strength ratios (TSR) from the moisture damage test are reported in Table 2 which shows that TSR increases as RAP percentage increases. This increase in TSR may be due to increased stiffness of mixes because of the RAP hard aged binder.

Achieving volumetrics in the mixtures prepared with the use of recycled and alternative materials is essential for an acceptable performance. Performance characteristics can be measured using laboratory tests and field evaluation. The next section discusses the performance evaluation of the mixes prepared using warm mix technologies, various alternative friction aggregate, and recycled materials.

## WARM MIX ASPHALT STUDY

In conjunction with increasing, the asphalt mixture industry has also taken some steps towards creating more sustainable pavements. Reuse of flexible pavements, mixes with various recycled and alternative materials, and some asphaltic technologies are just a few examples of the sustainability practices. However, it is vitally important to achieve standard performance characteristics with the use of

recycled and alternative materials as well as innovative technologies.

Some performance characteristics of the new and alternative asphalt mixtures are briefly summarized in this section; the details can be found elsewhere (Al-Qadi et al., 2010). The new mixtures include warm mix asphalt (WMA), stone-matrix asphalt (SMA) with recycled asphalt shingles (RAS) and asphalt mixtures containing alternate friction aggregates (diabase, quartzite, and granite). Mixtures were collected by IDOT and delivered to the Advanced Transporation Research and Engineering Laboratory (ATREL) in Rantoul, Illinois for testing. Hamburg Wheel Tracker and dynamic modulus tests were conducted to evaluate rutting potential and stiffness of these mixtures.

A brief summary of the mixtures evaluated is as follows:

*Warm Mix Asphalt (WMA)*

Two warm mix technologies were evaluated. These are WMA containing 0.5% of the Evotherm additive (FRWARMT) and WMA with foamed asphalt. These mixtures were produced and sampled from a plant.

*Alternate Friction Aggregates (SMAs)*

These mixtures were considered to evaluate alternate friction aggregates in SMAs. Each mixture contains a different alternate friction aggregates (diabase, granite, and quartzite). These mixtures were all produced in a laboratory.

*Alternate Friction Aggregates (N90F Mixes)*

These mixtures were considered to evaluate alternate friction aggregates in standard N90 (N90F) mixtures. N90 is defined as a measure of compactive effort in the Superpave gyratory compactor. The mixtures were prepared with steel slag and quartzite. These mixtures were all produced in a laboratory.

The results obtained from the two performance tests are presented next.

**Hamburg Wheel Test Results**

Hamburg Wheel testing was performed on all of the mixtures evaluated in this study. Gyratory samples were all compacted to 7% ± 0.5% voids. The 124.5 to 132.1 mm tall gyratory samples were cut in half to obtain two specimens that were 60.96 ± 2 mm height. All mixtures were conditioned and tested in a water bath at 50 °C. The testing was performed for 20,000 passes or a rut depth of 12.5 mm and terminated when one of these criteria was reached. Figure 1 summarizes the results to compare rut depths from each test.

Some highlights from the Hamburg Wheel tests are as follows:

1. The addition of recycled asphalt shingles (RAS) shows an increase in displacement for these mixtures (Block 1 in Figure 1).
2. WMA does show a greater displacement than the other HMA of the same classification (Block 2 in Figure 1). The second WMA prepared with foamed asphalt yielded the highest rut depth among all of the tested mixtures. However, it is important to note this is a N50 mix, which might have contributed to the relatively poor performance. Since N50, here, is used as a measure of compactive effort, this is a mix that requires lower compactive effort when compared to a N90 mix.
3. The Hamburg Wheel data for the three SMA surface mixtures with alternate friction aggregates are also shown in Figure 1. There is little difference

between the average maximum displacements for these mixes.

4. The effect of alternate friction aggregates on N90F mixes are shown in Block 4 of Figure 1. The control mixture with steel slag has the largest average maximum displacement of 6.86 mm while the quartzite mixes range from 3.56 to 4.57 mm.
5. Average rut depth varies from approximately 3.05 to 7.87 mm (without the 10.2 mm rut depth observed in a N50 mix). The results indicate that alternate materials and technologies have a potential to provide similar performance characteristics as the standard mixes.

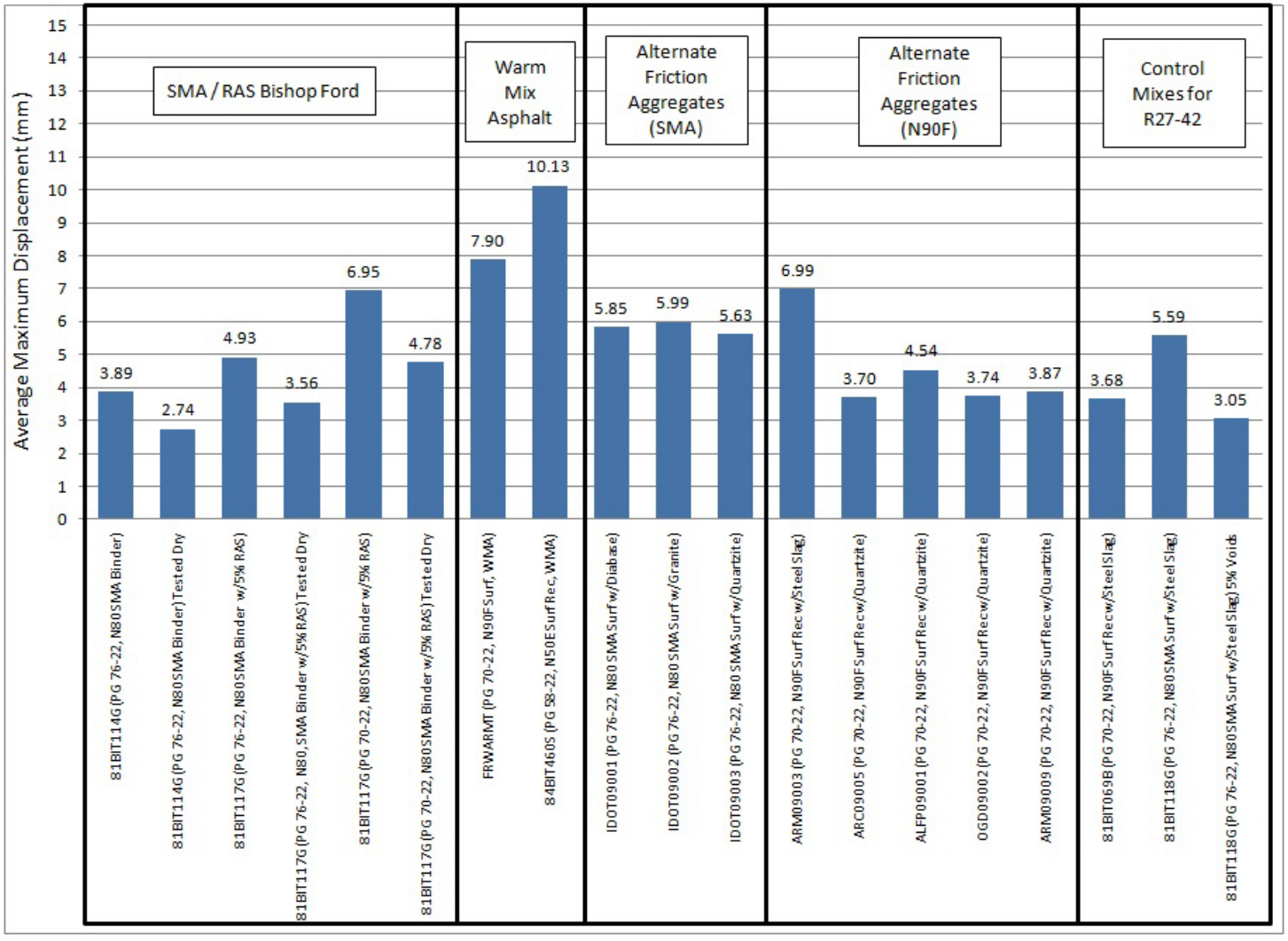



Figure 1. Summary of Hamburg Wheel testing data at 7% +/- 0.5% voids (after Al-Qadi et al., 2010).

### Complex Modulus Test Results

The dynamic modulus test provides a structural characterization of asphalt mixtures and is used as an input for the Mechanistic Empirical Pavement Design Guide (MEPDG) (Carpenter 2007). For each specimen, dynamic modulus tests were conducted at -10, 4, 21, 38, and 54 °C and frequencies of 25, 10, 5, 1, 0.5 and 0.1 Hz. Dynamic modulus testing was only performed on the RAS mixtures for the SMA / RAS Bishop Ford group.

Master curves were constructed using time-temperature superposition with a reference temperature of 21 °C. The results are shown in Figure 2 with master curves fitted to the experimental data. The results indicate no significant difference between the mixtures prepared with 5% shingles and the control mixtures.

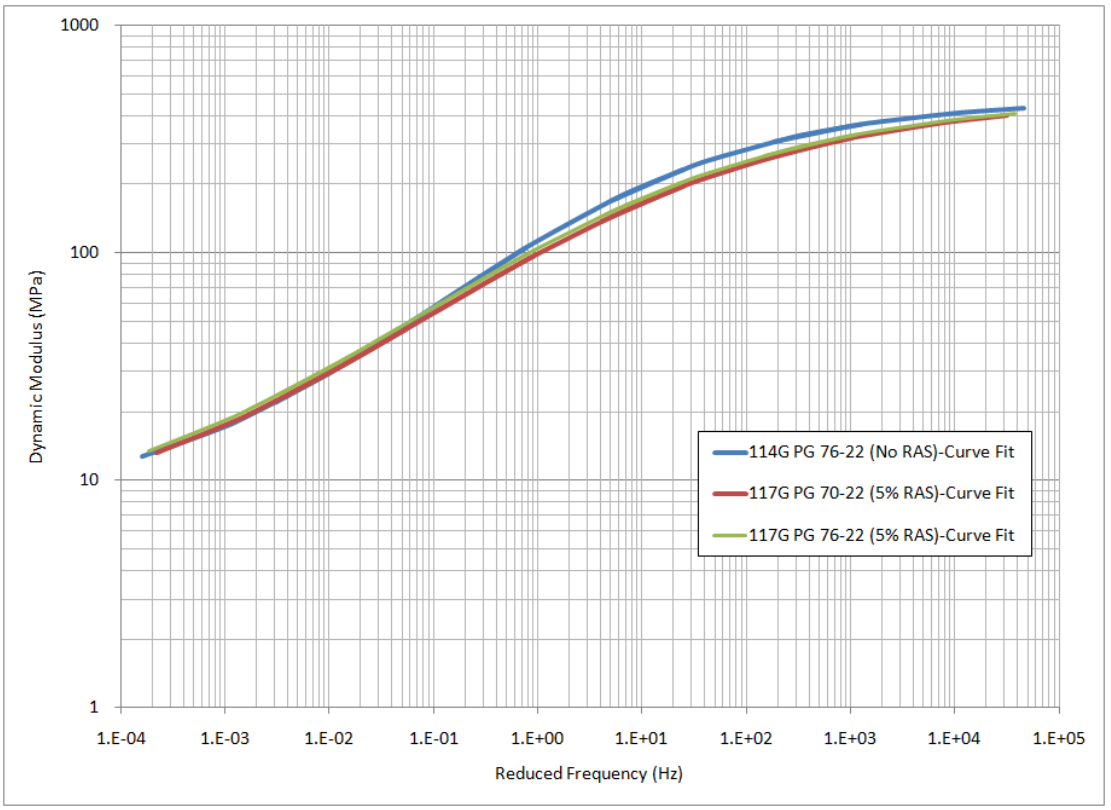



Figure 2. Master curve for the RAS mixtures (after Al-Qadi et al., 2010).

## SUMMARY

Using recycled and other alternative materials is a step forward towards sustainable pavement system. The asphalt mixtures prepared using these materials were evaluated in order to ensure better performance and durability of the mixtures. Achieving consistent volumetrics, particularly VMA, have long been a challenging task with increasing amounts of RAP in asphaltic mixes. Stringent specimen preparation protocols in conjunction with an aggregate packing theory, also known as the Bailey method, ensured the quality of the product with design volumetric

properties. Four mix designs prepared in accordance with these protocols are presented with their TSR values. TSR values tend to increase as RAP percentages increase. These results are only valid for the type of materials and mixes used in the study.

Warm mix asphalt (WMA), and mixes with recycled shingles and alternative friction aggregates were tested using the Hamburg Wheel and complex modulus. These tests were used to evaluate rutting potential and stiffness of these asphalt mixes. The results suggest that these mixes have the potential to provide similar performance characteristics as the standard mixes. Additional testing, including moisture susceptibility and fracture at low temperature, as well other sustainable mixtures are being considered.

## ACKNOWLEDGMENTS

This publication is based on the results of ICT-R27-SP17, Warm Mix Asphalt Study and ICT-R27-37, High-Rap Study. ICT-R27-SP17 and ICT-R27-37 were conducted in cooperation with the Illinois Center for Transportation; the Illinois Department of Transportation, Division of Highways; and the U.S. Department of Transportation, Federal Highway Administration.

The authors would like to acknowledge the assistance of the members of the Technical Review Panel for ICT-R27-37. The contents of this paper reflect the view of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Illinois Center for Transportation, the Illinois Department of Transportation, or the Federal Highway Administration. This paper does not constitute a standard, specification, or regulation.

## REFERENCES

Al-Qadi, I.L., Kern, J., and Meister, J. (2010). "Warm Mix Asphalt Study"-*Illinois Center for Transportation Report (ICT)*, R27-SP17, Rantoul, Illinois (in preparation).

Al-Qadi, I. L., Carpenter, S. H., Roberts, G. L., Ozer, H., and Aurangzeb, Q. (2009). "Investigation of Working Binder in Hot-Mix Asphalt Containing Recycled Asphalt Pavements," Paper Number 09-1262*, presented at 88th Annual Meeting of the Transportation Research Board*, Washington, DC.

Daniel, J. S. and Lachance, A. (2005). "Mechanistic and Volumetric Properties of Asphalt Mixtures with RAP," *Journal of the Transportation Research Board*, Vol. 1929, pp. 28-36.

Kim S., Sholar G.A., Byron T., and Kim J. (2009). "Performance of Polymer Modified Asphalt Mixture with Reclaimed Asphalt Pavement", *presented at 88th Annual Meeting of the Transportation Research Board,* Washington, DC.

Hajj, E. Y., Sebaaly, P. E. , Kandiah, P. (2008). "Use of Reclaimed Asphalt

Pavements (RAP) in Airfields HMA Pavements," *Airfield Asphalt Pavement Technology Program*, AAPTP Project No. 05-06, Final Report.

Mogawer, W. S., Austerman, A. J., Engstrom, B., and Bonaquist R. (2009). "Incorporating High Percentages of Recycled Asphalt Pavement (RAP) and Warm Mix Asphalt (WMA) Technology into Thin Hot Mix Asphalt Overlays to be Utilized as a Pavement Preservation Strategy," *presented at 88th Annual Meeting of the Transportation Research Board*, Washington, DC.

Superpave Mix Design (2001). Asphalt Institute Superpave Series No. 2(SP2), 3rd edition.

Vavrik, W. R., Huber, G., Pine W. J., Carpenter, S. H. and Bailey, R. (2002). "Bailey Method for Gradation Selection in HMA Mixture Design," *Transportation Research E-Circular,* E-C 044.

West, R., Kvasnak, A., Tran, N., Powell, B., and Turner, P. (2009). "Laboratory and Accelerated Field Performance Testing of Moderate and High RAP Content Mixes at the NCAT Test Track", *presented at 88th Annual Meeting of the Transportation Research Board*, Washington, DC.

# Moisture Damage and Fatigue Cracking of Foamed Warm Mix Asphalt Using a Simple Laboratory Setup

Shu Wei Goh[1] and Zhanping You[2]

[1]Research Assistant, Department of Civil and Environmental Engineering, Michigan Technological University, Houghton, Michigan 49931, USA. Tel: (906)487-2528, Fax: (906)487-1620, Email: sgoh@mtu.edu
[2]Associate Professor, Department of Civil and Environmental Engineering, Michigan Technological University, Houghton, Michigan 49931, USA. Tel: (906)487-1059, Fax: (906)487-1620, Email: zyou@mtu.edu

## ABSTRACT

Warm mix asphalt (WMA) appears to allow a reduction in the temperature at which asphalt mixes are produced. This technology is rapidly gaining prominence due to its uniqueness compared with hot mix asphalt (HMA). In this study, the WMA was produced using the foaming method under a simple laboratory setting, and evaluated by the indirect tensile strength and four-point beam fatigue testing. A control mixture, WMA with 1.0%, 1.5% and 2.0% water based of asphalt binder weight were used in the test. These WMAs were produced at production temperatures of 100°C, 115°C and 130°C. The details of the method used to produce the foamed WMA were discussed in this study. Tensile strength and four-point beam fatigue tests were carried out to evaluate the moisture damage and fatigue characteristics of WMA. Generally, when lower production temperature was used, the WMA had higher fatigue potential based on four-point beam fatigue testing. The findings from this study were summarized in the paper.

Keywords: Warm Mix Asphalt, Foaming, Tensile Strength, Moisture Damage, Asphalt Mixes, Four Point Beam Fatigue Cracking

## INTRODUCTION

Traditionally, asphalt mixtures were produced at high temperatures, between 150°C to 180°C, and thus often referred to as Hot Mix Asphalt (HMA). Recently, a new technology named Warm Mix Asphalt (WMA) was developed in Europe that allows HMA to be produced at a lower temperature. The goals for Warm Mix Asphalt (WMA) was to allow the use of existing HMA plants, standards and specifications; noting that it has similar qualities to that of HMA (Newcomb 2006). Over years of research efforts, quite a few technologies were used to produce WMA including the foaming method using Aspha-min®, Advera®, and WAM-FOAM®; organic

additives such as Sasobit® and Asphaltan B; chemical packages such as Evotherm®; and surfactants such as Cecabase RT® (FHWA 2007; Von Devivere et al. 2003).

Some benefits were found when lower temperatures are used to produce asphalt mixtures, especially when it comes to environmental and energy savings. Past research indicates that both emissions and energy usage (fuel) were reduced significantly when the WMA concept was used (Button et al. 2007; Gaudefroy et al. 2009; Goh and You 2009; Hassan 2009; Kristjansdottir 2006; Ventura et al. 2009). Some other potential benefits included cold weather paving, reduced thermal segregation of materials, extended paving window, improved workability, earlier traffic opening after construction, reduced worker exposure to asphalt fumes, and slowed binder aging potential (Gaudefroy et al. 2009; Hassan 2009; Ventura et al. 2009).

Even though WMA has shown promising results in energy savings and emission reduction, however, only limited studies and laboratory tests have been conducted to date. Recently, WMA using the foaming technique has slowly gained the interest of the researcher because it involves lower cost (no additional additive) to produce the WMA mixtures. This paper presents the foamed WMA that has been produced under laboratory setting, and its performance evaluated under the tensile strength and four-point beam fatigue tests.

## LITERATURE REVIEW

This literature review part collects information from laboratory experiences and previous research done on WMA using the foaming method. In this section, the methods for creating the foaming effect and laboratory test results of WMA using foaming method were reviewed.

The production of WMA using the foaming method has been gaining popularity recently due to its relatively lower cost. Additionally, it was the first technique used to produce WMA in 1956 by Professor Csanyi at Iowa State University (Button et al. 2007). The history of WMA could be traced back to 1956 when Dr. Csanyi found that foamed asphalt could potentially be used as soil binder. This invention was then modified by adding cold water instead of steam in asphalt, and it was patented by Mobil Oil Australia in 1968 (Button et al. 2007). This invention was later licensed to Conoco Inc. to promote foamed asphalt in United States and to further develop the product as a base stabilizer for both laboratory and field evaluation (Kristjansdottir 2006; Little et al. 1983). Since the 1970s, researchers have been trying to investigate a new method to reduce asphalt's mixture production temperature (Zettler 2006). This method was later termed as Warm Mix Asphalt (WMA).

Currently, several kinds of WMA technologies have been developed and are in use in European countries and the USA. They are mainly categorized as foaming techniques, organic additives and chemical packages. In this study, different kinds of foaming techniques and laboratory test results of WMA using this technique will be discussed.

There were a few methods used in this foaming technique, including the use of the foaming nozzle to introduce small amount of water into asphalt binder; using a

hydrophilic material (i.e. Aspha-min® or Advera®); and using wet aggregate to create a foaming effect during the mixing process (D'Angelo et al. 2008; You and Goh 2008). The concept behind this foaming technique was that the water would create a volume expansion of the asphalt binder that results in asphalt foam, and allows increased workability and aggregate coating at lower temperatures (Von Devivere et al. 2003; You and Goh 2008).

Laboratory studies and field experiences have been conducted in the past few years on foamed WMA. Their result mainly indicated that a foamed WMA would aid compaction, increase moisture damage, and provide comparable or higher rutting potential compared to the Hot Mix Asphalt (HMA) mixture (Goh and You 2008; Hurley et al. 2006; Wasiuddin et al. 2007; Wasiuddin et al. 2008). Recently, it was found that foamed WMA could be used for asphalt mixture containing high recycled asphalt pavement (RAP) content. Since RAP mixtures usually required a higher production temperature, using the WMA technique would aid the production of RAP mixtures and this is becoming an interesting development in the asphalt industry; however, there were no solid conclusions drawn for each test section on the performance of foamed WMA (Losa et al. 2009; Reyes et al. 2009; Robjent and Dosh 2009).

Although it shows a significant promise in energy saving and expanded construction season, the use of WMA is relatively new to the asphalt industry. The performance of the WMA is still not well understood. Therefore, many transportation agencies and contractors are not confident about the application of WMA.

## OBJECTIVE

The objectives of this study are: 1) to produce WMA using foaming method under a simple laboratory setting and; 2) to compare its performance with HMA mixture based on the tensile strength and four-point beam fatigue testing.

## SAMPLE PREPARATION

In this study, HMA mixtures (control) and WMA mixtures, that were produced using the foaming method, were evaluated and compared. All the mixtures gradations were designed based on specifications for a local asphalt mixture used in Michigan, USA. The nominal maximum aggregate size is 9.5mm and the designed traffic level is less than 3 million equivalent single axles loads (ESALs) based on the current Superpave™ asphalt mixture design procedure. A performance grade of PG 58-34 asphalt binder was used in this study. Tap water at the rate of 1%, 1.5% and 2% (based on binder weight) was injected into the asphalt binder using a syringe. It should be noted that a certain pressure should be applied to the syringe to allow water injected into the asphalt a short period of time (less than a second). Additionally, it was noteworthy that the asphalt binder was heated up to mixing and compacting temperature, which were 100°C, 115°C and 130°C, before the water was introduced. When the water came into contact with the asphalt, the molecules of the water became very volatile due to the high temperature of asphalt which was close to or above its boiling point. The water then vaporized and turned into steam. Immediately after water was injected to the bottom part of the asphalt binder, a spatula was used to

rapidly mix the asphalt and the water in order to allow the steam to disperse completely in the asphalt binder. Figure 1 shows the procedure for producing the foamed asphalt binder.

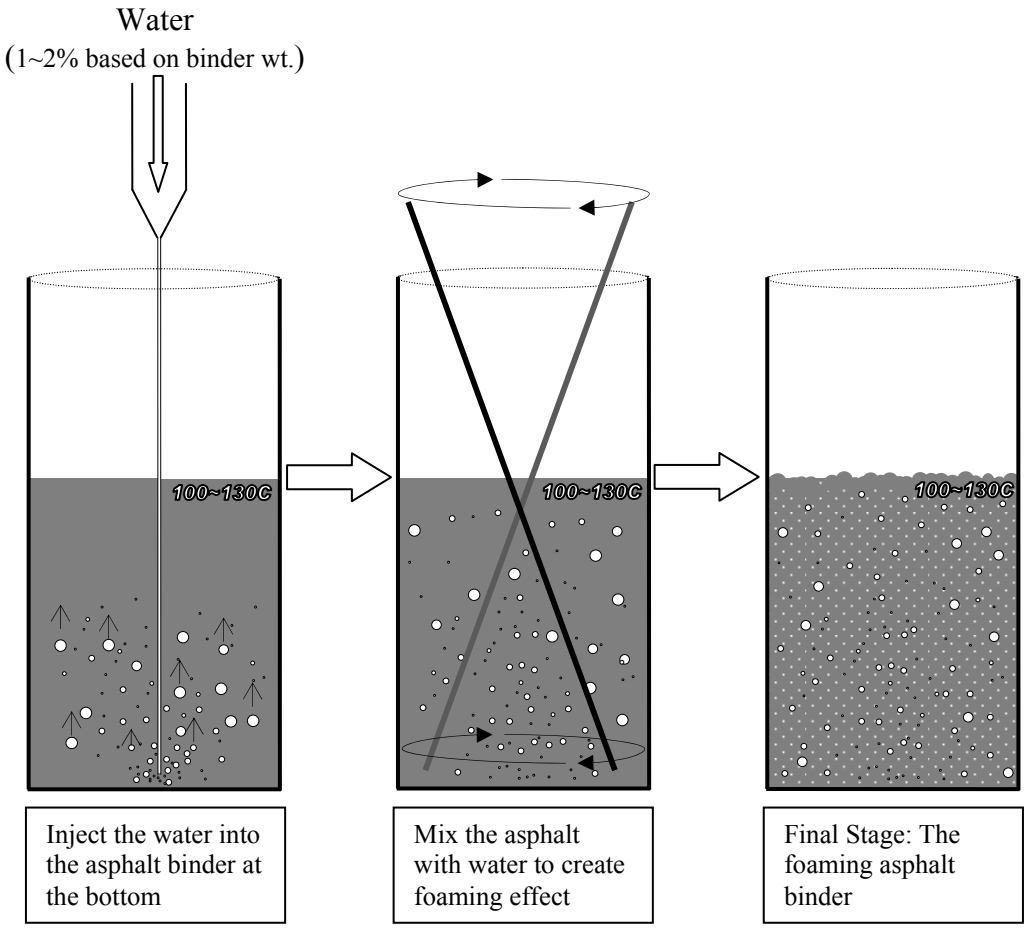



**Figure 1. Procedure to produce foamed asphalt binder**

When foam formed throughout the asphalt binder, the asphalt binder was then immediately mixed with the aggregate at the same temperature (100°C, 115°C, and 130°C, respectively). The foamed asphalt mixtures, also referred to as foamed WMA, were compacted at the temperature similar to its mixing temperature (100°C, 115°C, and 130°C, respectively). A gyration number of 86 was applied during the compaction process using the Superpave™ gyratory compactor. Figure 2 shows the procedure of mixing and compaction of foamed WMA in this study; and Figure 3 shows the final product of WMA using this foaming method. The control mixtures were mixed at 165°C and compacted at 153°C. A similar gyration number of 86 was used for the control HMA mixture, and the Superpave™ specification was followed in the mix preparation. The volumetric properties of samples were evaluated as well after the compaction. It was found that the average air void level for control samples are 6.1%; and for WMA samples are ranged from 5.5% to 7.9%.

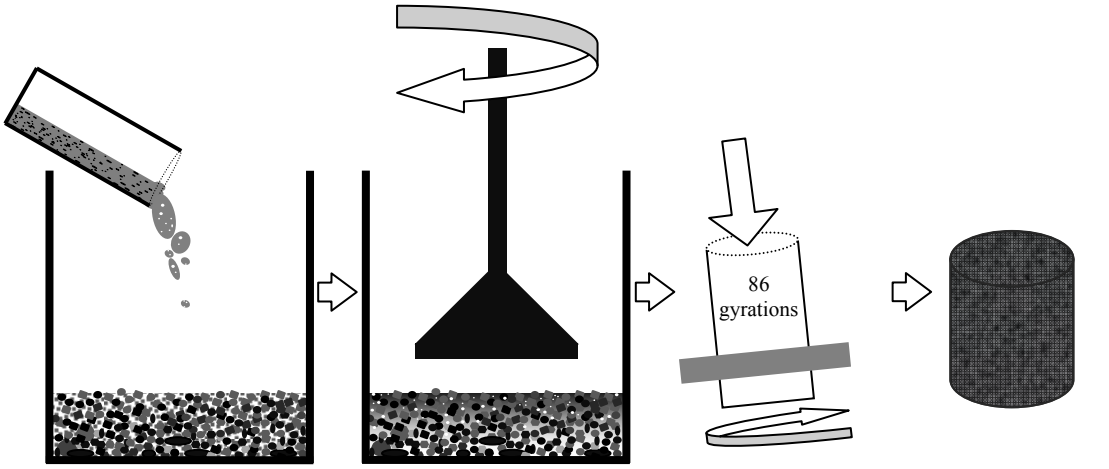



**Figure 2. Mixing and compacting the foam asphalt with aggregate**

**Figure 3. Warm asphalt mixture produced using the water foaming method**

## ASPHALT MIXTURE PERFORMANCE TESTING

In this study, two performance tests were conducted – the indirect tensile strength and the four-point beam fatigue tests. For the indirect tensile strength tests, the control samples of foamed WMA mixtures containing 1%, 1.5% and 2% water were produced at temperatures of 100°C, 115°C and 130°C using the Superpave™ gyratory compactor. For the four-point beam fatigue testing, only control mixtures and foamed WMAs using 1% water produced at 115°C and 130°C using the linear

kneading compactor were used due to limited material available. Three replicated samples were used for each tensile strength and four-point beam fatigue testing. It is noteworthy that the descriptions used in the graphs for each asphalt mixture are shown in Table 1.

**TABLE 1 Description Asphalt Mixture used in the Graphs**

| Descriptor | Description |
|---|---|
| CTRL | Control HMA Mixture |
| 1% Water 100C | WMA using 1% water compacted at 100°C |
| 1% Water 115C | WMA using 1% water compacted at 115°C |
| 1% Water 130C | WMA using 1% water compacted at 130°C |
| 1.5% Water 100C | WMA using 1.5% water compacted at 100°C |
| 1.5% Water 115C | WMA using 1.5% water compacted at 115°C |
| 1.5% Water 130C | WMA using 1.5% water compacted at 130°C |
| 2.0% Water 100C | WMA using 2.0% water compacted at 100°C |
| 2.0% Water 115C | WMA using 2.0% water compacted at 115°C |
| 2.0% Water 130C | WMA using 2.0% water compacted at 130°C |

**Tensile strength Testing**

The purpose of the tensile strength ratio testing is to evaluate the asphalt mixture's fatigue potential and moisture susceptibility. Previous research has indicated that the tensile strength of hot-mix asphalt is related to its fatigue cracking (Chapuis and Gatien 1995). A higher tensile strength indicated that the asphalt pavement can tolerate higher strains before failing (i.e. cracking). Additionally, the moisture susceptibility of the asphalt mixture can be evaluated by comparing the tensile strength of asphalt mixtures exposed to wet and dry conditions. In this study, all samples were tested based on AASHTO T283 (AASHTO:T283-03 2003). A loading rate of 0.83 mm/s and a testing temperature of 25°C were used during the testing. As mentioned previously, control mixtures and foamed WMA mixtures using the 1%, 1.5% and 2% water produced at 100°C, 115°C and 130°C were evaluated.

Figure 4 shows the results for the tensile strength testing. Based on the results, it is observed that the foamed WMA have lower tensile strength in general compared to the control mixtures. One interesting finding is that during testing, the tensile strength for all the foamed WMA at production temperatures at around 115°C was the highest among all the foamed WMA mixtures tested. Additionally, the production temperature at around 115°C could be the effective temperature for WMA because it shows the highest tensile strength compared to WMA produced at 100°C and 130°C. The main reason behind this was likely due to the effect of binder aging and aggregate coating. Aged binder from higher production temperature (stiffer binder) could result in lower tensile strength value. On the other hand, using lower mixing temperature could result in another problem that the aggregate may not be fully coated.

In this study, the tensile strength ratio (TSR), the tensile strength between dry and conditioned (mixtures went through one freeze-thaw cycle) were compared and the results are shown in Figure 4 as well. Typically, the final result for TSR testing

would have a value of less than 1.00 because it is expected that the conditioned samples would suffer moisture damage and exhibit lower tensile strength; this phenomenon was observed in the control sample. However, it was found that some of the foamed WMA mixtures exhibited TSR values greater than 1.00. This indicated that the sample after conditioning has higher tensile strength. The best mixture in this case was the foamed WMA mixture using 1% water compacted at 130°C. Additionally, it was observed that when the WMA production temperature increased, the TSR increased this held true in all cases.

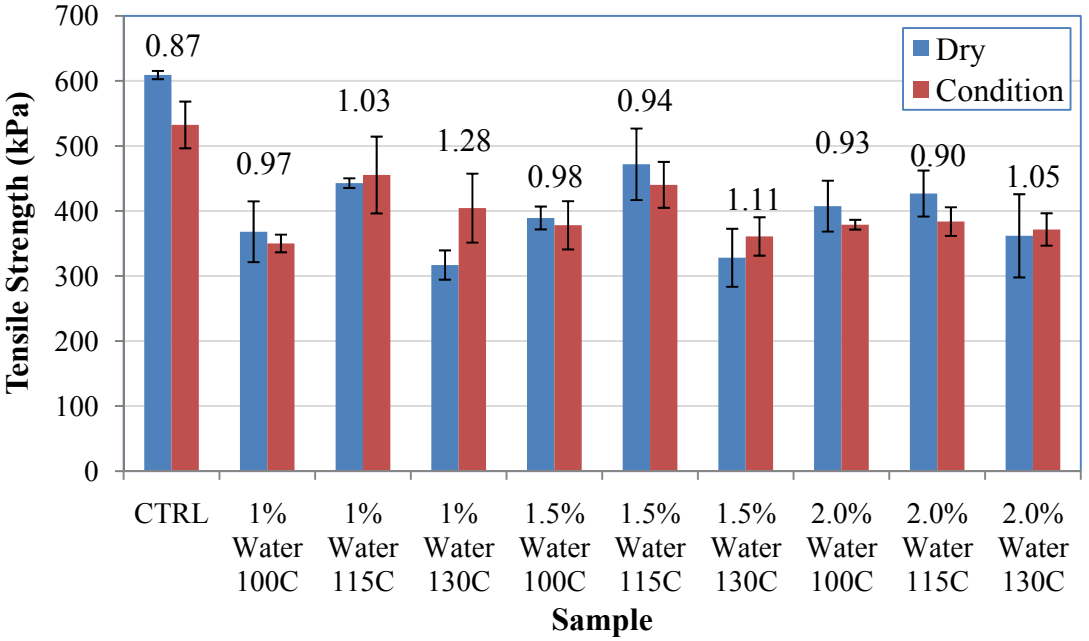



**Figure 4. Comparison of Indirect Tensile Strength and TSR for the Control Mixture, and WMA using 1%, 1.5% and 2% Water at 100°C, 115°C and 130°C**

**Four-point beam fatigue testing**

The results from the four-point beam fatigue tests are presented in this section. The purpose of this test is to determine the fatigue life of the asphalt mixture subjected to the repeated bending until failure where the fatigue failure was defined as 50% reduction of initial stiffness (AASHTO:T321-07 2007). In this test, a frequency of 10 Hz and 400 micro-strain (constant strain) were used for all the samples tested. As mentioned previously, only control mixtures and foamed WMAs using 1% water produced at 115°C and 130°C were tested due to the limited material available. The results of the four-point beam fatigue testing are presented in Figure 5.

From Figure 5, it can be found that all the foamed WMA fatigue life was higher than the control HMA. It is also noticed that when the foamed WMA's production temperature increased, the fatigue life increased as well. There are several factors that would affect the fatigue life associated with production temperatures and WMA additives when comparing HMA and WMA, including: 1) absorption – lower mixing temperature (WMA) may result in less binder absorption into the aggregate,

which will reduce the adhesion and thus affect the asphalt mixture fatigue life (Button et al. 2007) and; 2) aging of the asphalt binder – lower mixing temperature of WMA will reduce binder's aging and thus improve workability of asphalt mixture;

The evaluation of factors described above is underway and more tests on other types of WMA were ongoing to evaluate the fatigue life of WMA.

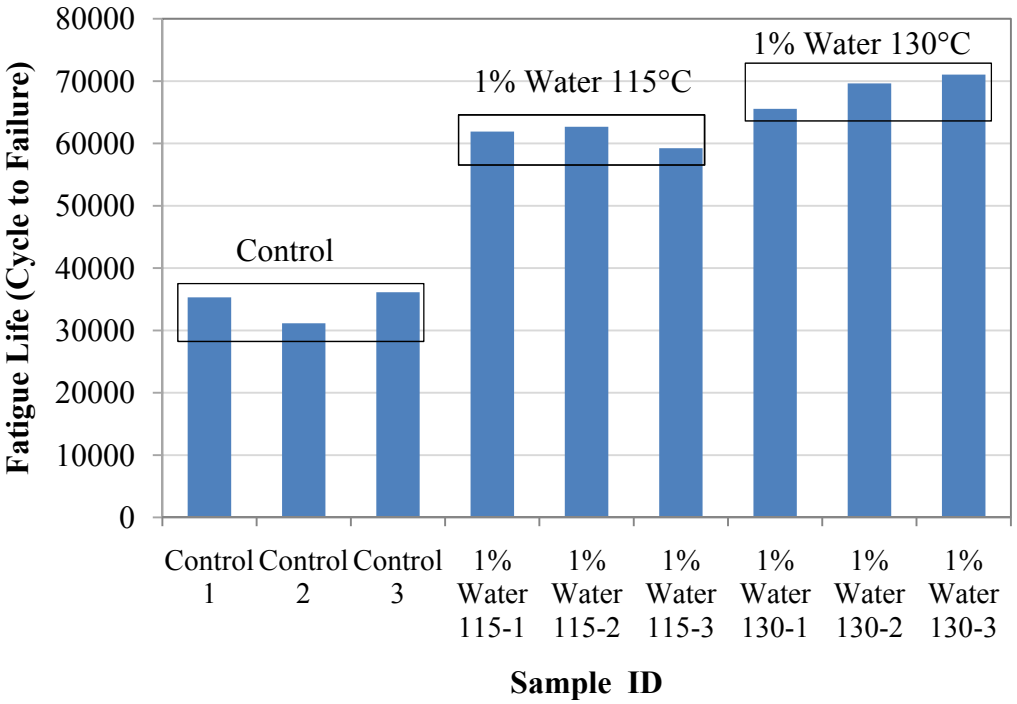



**Figure 5. Comparing the Fatigue Life of HMA (control mix) and WMA**

## SUMMARY

In this paper, WMAs using the foaming method under laboratory settings were produced. The average air void level for control samples (HMA) are 6.1%, and for WMA samples are ranged from 5.5% to 7.9%. The performance was compared based on the indirect tensile strength and four point beam fatigue test results. Tensile strength using the AASHTO T283 specifications and the four-point beam fatigue tests using AASHTO T321 specifications were performed on both the control and foamed WMA mixtures. For tensile strength testing, it was found that the foamed WMA generally have lower tensile strength when compared to the control mixture. The tensile strength for all the foamed WMA at a production temperature of around 115°C was the highest among all the foamed WMA mixtures tested, and it was suspected that this could be an effective temperature for WMA because it shows the highest tensile strength compared to WMAs produced at 100°C and 130°C.

The TSR for foamed WMA were higher than the control mixture. In addition, it was found that some of the foamed WMA mixtures exhibited TSR values greater than 1.00 and TSR increases when the production temperature of WMA increases.

For the four-point beam fatigue tests, all the foamed WMA's fatigue life was higher than the control HMA. Factors such as absorption and aging could be the main factors that affect WMA performances and the evaluations of these factors are currently under underway.

## ACKNOWLEDGEMENT

The authors wish to express their gratitude to James Vivian, Moua Lee, David Porter and Travis Aho at Michigan Technological University for their help in running some preliminary tests. The research work was partially sponsored by the Federal Highway Administration through Michigan Department of Transportation (MDOT). The authors also appreciate the guidance and involvement of John Barak of MDOT as the Project Manager.

## REFERENCES

AASHTO:T283-03. (2003). "Standard Method of Test for Resistance of Compacted Asphalt Mixture to Moisture Induced Damage." American Association of State Highway and Transportation Officials.

AASHTO:T321-07. (2007). "Determining the Fatiuge Life of Compacted Hot-Mix Asphalt (HMA) Subjected to Repeated Flexural Bending." American Association of State Highway and Transportation Officials.

Button, J. W., Estakhri, C., and Wimsatt, A. (2007). "A Synthesis of Warm-Mix Asphalt." *Report 0-5597-1*, Texas Transportation Institute, Austin, Texas.

Chapuis, R. P., and Gatien, A. (1995). "Temperature dependent tensile strength of asphalt mixtures in relation to field cracking data." ASTM, Philadelphia, PA, USA, 180-193.

D'Angelo, J., Harm, E., John Bartoszek, Baumgardner, G., Matthew Corrigan, Cowsert, J., Harman, T., Mostafa Jamshidi, Jones, W., Newcomb, D., Brian Prowell, Sines, R., and Yeaton, B. (2008). "Warm Mix Asphalt: European Practice." Office of International Programs, Office of Policy, Federal Highway Administration, U.S. Department of Transportation, American Association of State Highway and Transportation Officials, National Cooperative Highway Research Program.

FHWA. (2007). "Warm Mix Asphalt Technologies and Research." U.S. Department Of Transportation Federal Highway Administration.

Gaudefroy, V., Olard, F., Ridane, M., Beduneau, E., and de La Roche, C. (2009). *Laboratory Environmental Assessment of Half-Warm Mix Asphalts by Factorial Experiment Design Approach*, Transportation Research Board.

Goh, S. W., and You, Z. (2008). "Resilient Modulus and Dynamic Modulus of Warm Mix Asphalt." GeoCongress 2008, Geosustainability and Geohazard Mitigation, Geotechnical Special Publication (GSP 178), American Society of Civil Engineers, 1000-1007.

Goh, S. W., and You, Z. (2009). "Preliminary Study of Evaluating Asphalt Pavement Rutting Performance Using The Mechanistic-Empirical Pavement Design Guide." Proceedings of the 14th Conference on Cold Regions Engineering, American Society of Civil Engineers, Duluth, Minnesota, 366-373.

Hassan, M. M. (2009). *Life-Cycle Assessment of Warm-Mix Asphalt: Environmental and Economic Perspectives*, Transportation Research Board.

Hurley, G. C., Prowell, B. D., Reinke, G., Joskowicz, P., Davis, R., Scherocman, J., Brown, S., Hongbin, X., and Bonte, D. (2006). "Evaluation of Potential Processes For Use In Warm Mix Asphalt." Association of Asphalt Paving Technologist, White Bear Lake, MN 55110, United States, Savannah, GA, United States, 41-90.

Kristjansdottir, O. (2006). "Warm Mix Asphalt for Cold Weather Paving," University of Washington, Seattle.

Little, D. N., Button, J. W., and Epps, J. A. (1983). "Structural Properties of Laboratory Mixtures Containing Foamed Asphalt and Marginal Aggregates." *Transportation Research Record*, 911.

Losa, M., Bacci, R., and Leandri, P. (2009). "Experimental Characterization of High Performance WAM-Foam Trial Sections." Sixth International Conference on Maintenance and Rehabilitation of Pavements and Technological Control (MAIREPAV6), Turin , Italy.

Newcomb, D. (2006). "An Introduction to Warm Mix Asphalt." National Asphalt Pavement Association, Lanham, Maryland.

Reyes, M., Forfylow, R. W., and Middleton, B. (2009). "Initial Performance of Foam WMA Mixes in Western Canada." 54th Annual CTAA Conference, Moncton, NB.

Robjent, L., and Dosh, W. (2009). "Warm-Mix Asphalt for Rural County Roads." ASCE, Duluth, MN, 43-43.

Ventura, A., Moneron, P., Jullien, A., Tamagny, P., Olard, F., and Zavan, D. (2009). *Environmental Comparison at Industrial Scale of Hot and Half-Warm Mix Asphalt Manufacturing Processes*, Transportation Research Board.

Von Devivere, M., Barthel, W., and Marchand, J. P. (2003). *Warm Asphalt Mixers by adding a synthetic zeolite*, World Road Association - PIARC.

Wasiuddin, N. M., Selvamohan, S., Zaman, M. M., and Guegan, M. L. T. A. (2007). "A Comparative Laboratory Study of Sasobit® and Aspha-min® in Warm-Mix Asphalt." Transportation Research Board 86th Annual Meeting, Washington DC, United States.

Wasiuddin, N. M., Zaman, M., and Rear, E. A. (2008). *Effect of Sasobit and Aspha-min on Wettability and Adhesion Between Asphalt Binders and Aggregates*, Transportation Research Board.

You, Z., and Goh, S. W. (2008). "Laboratory Evaluation of Warm Mix Asphalt: A Preliminary Study." *International Journal of Pavement Research and Technology*, 1(1), 34-40.

Zettler, R. (2006). "Road manager: Warm mix stands up to its trials." *Better Roads*, 76(2), 16-21.

# A Field Investigation of the Effect of Pavement Type on Fuel Consumption

Imen Zaabar, PhD[1] and Karim Chatti, PhD[2]

[1] Visiting Research Associate, Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824, USA, Email: zaabarim@egr.msu.edu

[2] Professor, Department of Civil and Environmental Engineering, Michigan State University, 3546 Engineering Building, East Lansing, MI-48824, USA.
Email: chatti@egr.msu.edu

## ABSTRACT

Fuel consumption costs are influenced by vehicle technology, pavement condition, roadway geometrics, environment, speed, and other factors. The goal of this paper is to investigate the effect of pavement type on fuel consumption. The study has entailed the use of five instrumented vehicles to make fuel consumption measurements over different concrete (PCC) and asphalt (AC) pavement sections. The sections selected have similar characteristics (grade, roughness and texture) and differ only in the type of pavement. The data was collected at three different speeds. The results showed that the difference in fuel consumption between asphalt and concrete pavements is statistically significant at 95 percent confidence level for (loaded) light and heavy trucks at low speed (56 km/h) and summer conditions. Under these conditions, trucks driven over AC pavements will consume about 4% more than if they were driven over PCC pavements. Fuel consumption data for heavy truck in winter was not available. The analysis also showed that the mean differences of fuel consumption between asphalt and concrete pavements for passenger car, van and SUV are statistically not significant for both winter and summer conditions.

# INTRODUCTION

Understanding the costs of highway construction, highway maintenance and vehicle operation is essential to sound planning and management of highway investments, especially under increasing infrastructure demands and limited budget resources. While the infrastructure costs conceived by road agencies are substantial, the cost borne by road users are even greater. In 2009, the American Automobile Association (AAA, 2009) reported an average vehicle operating cost of 54.9¢ per vehicle mile based on 2008 prices. For conventional vehicles, these costs are related to fuel and oil consumption, tire wear, repair and maintenance, and depreciation. These costs depend on the vehicle class and are influenced by vehicle technology, pavement-surface type, pavement condition, roadway geometrics, environment, speed of operation, and other factors. Therefore, vehicle operating costs are part of the costs that highway agencies must consider when evaluating pavement-investment strategies.

Reduction in vehicle fuel consumption is one of the main benefits considered in technical and economic evaluations of road improvements considering its significance. According to the Bureau of Transportation Statistics (BTS), the 255 million vehicles in the United States consume about 200 billion gallons of motor fuel annually. With today's gas prices, this will translate to 400 to 600 billion dollars. Many researchers have investigated the effect of pavement conditions on fuel consumption. Zaniewski et al. (1982) reported that pavement conditions have no effect on fuel consumption. Their conclusions were based on field trials that were performed in the 1980s with vehicles that vary substantially from the ones used currently in the US. Their vehicles were first driven for long distances then the amount of fuel consumed was measured. Since the variability in the pavement conditions will be high for long distances, the effect of pavement conditions could be cancelled out. Sandberg (1990) reported that the texture of the pavement has significant effect on rolling resistance and therefore on fuel consumption. Zaabar and Chatti (2010) reported that roughness has a significant effect on rolling resistance forces and consequently on fuel consumption.

Zaniewski also looked at the effect of pavement type on fuel consumption (Zaniewski, 1989). He conducted a more detailed analysis of the data collected in the 1980's and had pointed out that fuel consumption of a truck when travelling on PCC pavements was lower than when travelling on AC pavements (20% difference).

Netherlands Pavement Consultants (NPC, 2002) made a theoretical calculation of the maximum energy dissipation when driving on asphalt versus concrete pavement. They reported that the possible difference is likely to be much lower than 1%.

A Canadian study (Taylor et al, 2000; Taylor and Patten, 2006) investigated whether vehicles driven over different types of pavement would display fuel savings based on pavement type. The study used a multivariate linear regression analysis including different variable categories: vehicle, weather and pavement related variables. They then used the model to account for the effects of pavement type, roughness, temperature, speed, etc. to predict the fuel consumption differential. One issue with this approach is that it can lead to large errors in interpretation especially when using models with a large number of independent variables and dealing with small main effects. Nonetheless, they reported that concrete pavements generally provided fuel savings over asphalt pavements, ranging from 0.8 per cent to 6.9 per cent for heavy

trucks (semi-trailers). The same study found that for passenger vehicles, concrete showed statistically significant fuel savings over asphalt only in winter, although it was reported that due to the limited sample size and the number of temperature ranges, the results from the passenger car testing were less conclusive than for the heavy trucks.

Sumitsawan et al (2006) also investigated the effect of pavement type on fuel consumption. An instrumented van was driven over two new pavement sections: one rigid and one flexible. All tests were conducted under dry pavement conditions. They reported that the differences in fuel consumption rates were statistically significant at 10% level of significance (vehicle driven over rigid pavement consumes less than when it is driven over asphalt pavement).

The Swedish National Road and Transport Research Institute (VTI) investigated the difference in fuel consumption between asphalt and concrete pavements (Jonsson, and Hultqvist, 2008). The measurements showed that there was 1.1 per cent less fuel consumption on the concrete pavement compared to the asphalt pavement. The difference was mainly attributed to the macro texture differences (the asphalt pavement had rougher macro texture than the concrete pavement).

In summary, several studies have investigated the effect of pavement type on fuel consumption. Even though all the reported results show savings from operating on concrete pavements, the reported savings seem to vary significantly. Therefore, there is a need to further investigate this question. In this paper, an attempt was made to extract the effect of pavement type directly from the data by controlling all other variables except for pavement type. This was done by selecting only flat and smooth pavement sections.

## EXPERIMENTAL DESIGN AND DATA COLLECTION

Five different locations (near Lansing, Michigan) were selected based on the variability level of their pavement conditions (i.e. roughness, gradient, texture and pavement type). The tests were conducted during wet (winter) and dry (summer) conditions. The weather conditions (temperature and wind speed) were recorded using a portable weather station. A maximum change in ambient temperature of 3°C was imposed for the test to be acceptable. Tables 1 and 2 summarize the testing and pavement surface conditions for each section. The test equipment used during the field tests is an instrument that could access and log data instantaneously from the vehicle's Engine Control Unit (ECU) via On Board Diagnostic (OBD) connector. The pavement condition data (raw profile and texture depth) were collected by Michigan Department of Transportation (MDOT) using a Rapid Travel Profilometer and a Pavement Friction Tester equipped with a laser-based measurement system to measure texture depth. The slope data surveys were collected by a third party using a high precision GPS (average error of about twice that of the total station). The sampling rate was every 1 second at highway speed (every 30 m).

Table 1 Test Matrix for Field Trials

| Section ID | Pavement Type | | IRI range (m/Km) | Length (Km) | Speed limit (Km/h) | Test Speed (Km/h) | | Replicates |
|---|---|---|---|---|---|---|---|---|
| | AC | PCC | | | | | | |
| Creyts Rd | X | | 1.3 - 8.5 | 1.44 | 72 | 56 | 72 | 2 |
| Creyts Rd | X | | 1.7 - 7 | 1.6 | 72 | 56 | 72 | 2 |
| Waverly Rd | X | | 3.5 - 6 | 0.48 | 72 | 56 | 72 | 2 |
| Waverly Rd | X | | 3.3 - 6 | 0.64 | 72 | 56 | 72 | 2 |
| I69 | | X | 1.1 - 2.5 | 4.8 | 112 | 88 | 104 | 2 |
| M99 S | | X | 1.5 - 2.6 | 6.4 | 80 | 56 | 72 | 2 |
| M99 S | | X | 1.5 - 2.6 | 0.64 | 80 | 72 | 88 | 2 |
| M99 N | X | | 0.8 - 4.6 | 1.6 | 80 | 56 | 72 | 2 |
| M99 N | | X | | 0.48 | 80 | 56 | 72 | 2 |
| M99 N | | X | | 1.28 | 72 | 56 | 72 | 2 |

Table 2 Weather Conditions

| Weather conditions | Winter | | Summer | | |
|---|---|---|---|---|---|
| | Day 1 | Day 2 | Day 1 | Day 2 | Day 3 |
| Ambient temperature (°C) | 0-2 | 1-3 | 28.9-29.2 | 27.2-28.3 | 22.5-25.2 |
| Wind speed (m/s) | 1.7-2.4 | 0.4-1 | 2.1 – 2.9 | 1.4-2.4 | 1.7-2.4 |

Five different vehicles were used: (1) Medium car, (2) SUV, (3) Van, (4) Light truck and (5) Heavy (articulated) truck. Table 3 summarizes the characteristics of these vehicles. The vehicles were driven over the pre-selected sections. The data acquisition system was connected to the vehicle. The vehicles were driven at 35, 45 and 55 mph (56, 72 and 88 km/h) with and without cruise control. Multiple and repeated runs were performed. Tests for trucks were conducted under two different loading conditions: loaded and unloaded. The cargo loads for the light and heavy articulated trucks were 2.82 tonnes and 21.32 tonnes. These loading conditions are typical for the US. More details of the test set up and equipments are included in (Zaabar and Chatti, 2010).

The variables recorded during the tests included the following:

- Weather conditions: ambient air temperature, humidity and wind speed/direction.
- Vehicle conditions: tire pressure, vehicle weight, vehicle speed, engine speed, temperature of fuel flowing in the engine.
- Status of auxiliary devices (A/C, Radio, Headlights).
- Pavement conditions: IRI, Mean Profile Depth (MPD) and Slope data.

Table 3 Characteristics of the Vehicles Used in the Field Trials

| | Vehicle class | | | | |
|---|---|---|---|---|---|
| Characteristics | Medium car | SUV | VAN | Light truck | Heavy truck |
| Make | Mitsubishi | Nissan | Ford | GMC | International |
| Model | Galant | Pathfinder | E350 | W4500 | 9200 6x4 |
| Year | 2008 | 2009 | 2008 | 2006 | 2005 |
| Drag coefficient | 0.4 | 0.5 | 0.5 | 0.6 | 0.8 |
| Frontal area ($m^2$) | 1.9 | 2.9 | 2.9 | 4.2 | 9 |
| Tare Weight (t) | 1.46 | 2.5 | 2.9 | 3.7 | 13.6 |
| Maximum allowable Load (t) | - | - | - | 2.9 | 22.7 |
| GVW (t) | - | - | - | 6.6 | 36.3 |
| Weight of the load (t) | - | - | - | 2.8 | 21.3 |
| Gas type | Gas | Gas | Gas | Gas | Diesel |
| Tire diameter (m) | 0.38 | 0.4 | 0.4 | 0.4 | 0.57 |
| Tire pressure (psi) | 35 | 39 | 43 | 75 | 110 |
| Tire type | radial | radial | radial | radial | bias |
| Cargo length (m) | - | - | - | 4.88 | 15.85 |
| Other | - | 4WD | 15 seats | - | Flat bed |

## EFFECT OF PAVEMENT TYPE ON FUEL CONSUMPTION

A detailed analysis to find the effect of pavement type on fuel consumption was conducted. The methodology is as follows:

1. Divide each of the roads mentioned in Table 1 into 30 m-subsections (100ft).
2. Consider only smooth sections (1 m/km ± 10%), based on the sensitivity of fuel consumption to IRI (Zaabar and Chatti, 2010).
3. For the sections extracted in step 2, consider only sections that have a texture depth (Mean Profile Depth) of 0.5 ± 20%, based on the sensitivity of fuel consumption to texture (Zaabar and Chatti, 2010).
4. For the sections extracted in step 3, consider only the sections that are flat (based on the sensitivity of fuel consumption to grade, sections with slope in the range of ± 0.1% are considered to be flat).
5. Divide the sections extracted in step 4 into two groups based on their pavement type (i.e., concrete and asphalt).
6. Conduct univariate analysis having IRI as a covariate and pavement type as fixed factor.

7. Repeat for all vehicles at 56, 72 and 88 km/h.

The above methodology led to a total of 138 sections for each vehicle and speed. Tables 4 and 5 present summary statistics for articulated and light trucks, respectively. Figures 1 and 2 show the mean and standard deviation of fuel consumption for articulated and light trucks, respectively, driven over different pavement types at 56, 72 and 88 km/h. Tables 6 summarizes the results of the main effect analysis using SPSS. It can be seen that, for both truck types and for summer conditions, the mean difference of fuel consumption between asphalt and concrete pavements is statistically significant at 56 km/h; whereas, it is statistically not significant at higher speeds (i.e. 72 and 88 km/h). For winter conditions, the mean difference of fuel consumption between asphalt and concrete pavements is statistically not significant. However, it should be noted that heavy (articulated) truck was only tested in summer. These observations could be explained by the viscoelastic behavior of asphalt pavement. Since in summer AC pavements are more viscous, trucks will cause more deflection on flexible pavements than on rigid pavements. Therefore, trucks need more power to overcome the traction caused by the deflected pavement; thus, more fuel is required to drive on flexible pavements. However, in winter, AC pavements become more rigid (less deflection will occur). Therefore, in winter, the power required to overcome the deflected AC pavement will be similar to rigid pavement.

For passenger car, van and SUV, the analysis showed that the mean differences of fuel consumption between asphalt and concrete pavements are statistically not significant (Table 6).

Table 4 Estimated Marginal Means – Articulated Truck

| Pavt type | Speed (km/h) | Mean (mL/km) | Std. Deviation | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| PCC | 56 | 201.391[a] | 1.077 | 198.610 | 204.171 |
| | 72 | 222.938[a] | 1.053 | 220.220 | 225.655 |
| | 88 | 248.401[a] | 1.190 | 245.328 | 251.473 |
| AC | 56 | 209.422[a] | 1.008 | 206.821 | 212.024 |
| | 72 | 225.213[a] | 1.034 | 222.543 | 227.883 |
| | 88 | 247.588[a] | 1.183 | 244.535 | 250.641 |

a. Covariates appearing in the model are evaluated at the following values: IRI = 1.2298.

Table 5 Estimated Marginal Means – Light Truck

Dependent Variable: FC_mLKm

| Pavt type | Speed (km/h) | Mean (mL/km) | Std. Deviation | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| PCC | 56 | 151.130[a] | 1.080 | 148.344 | 153.916 |
| | 72 | 187.778[a] | 1.071 | 185.015 | 190.541 |
| | 88 | 225.188[a] | 1.294 | 221.850 | 228.527 |
| AC | 56 | 156.739[a] | 1.061 | 154.002 | 159.475 |
| | 72 | 188.398[a] | 1.052 | 185.685 | 191.112 |
| | 88 | 219.973[a] | 1.225 | 216.811 | 223.134 |

a. Covariates appearing in the model are evaluated at the following values: IRI = 1.2298.

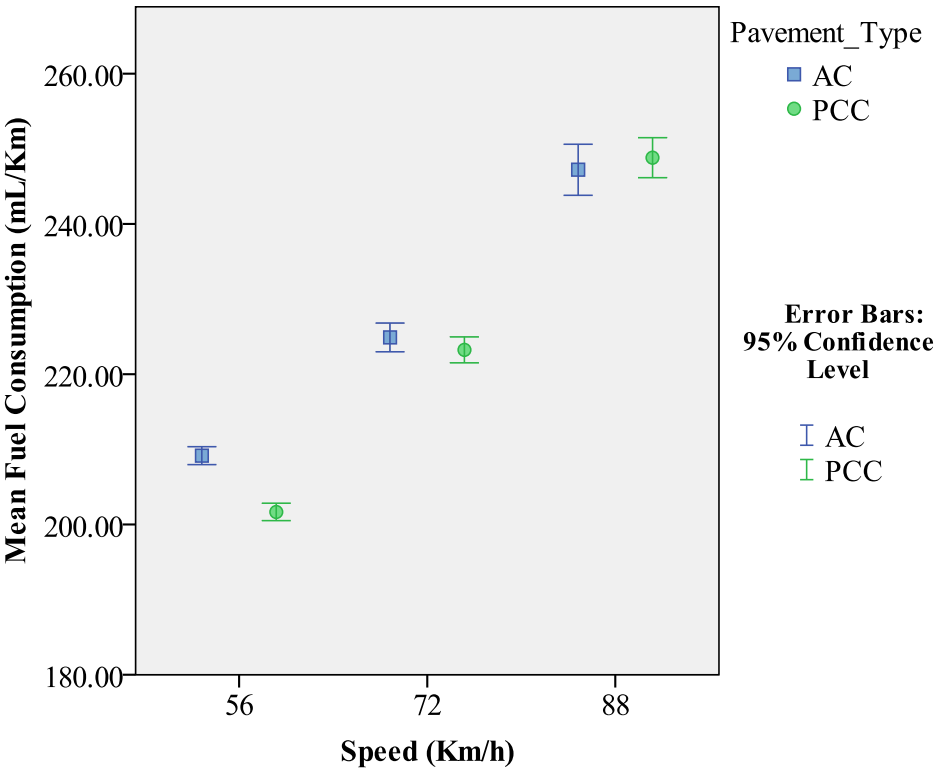



Figure 1 Mean and Standard Deviation of Fuel Consumption for Different Pavement Type and Speed– Articulated Truck

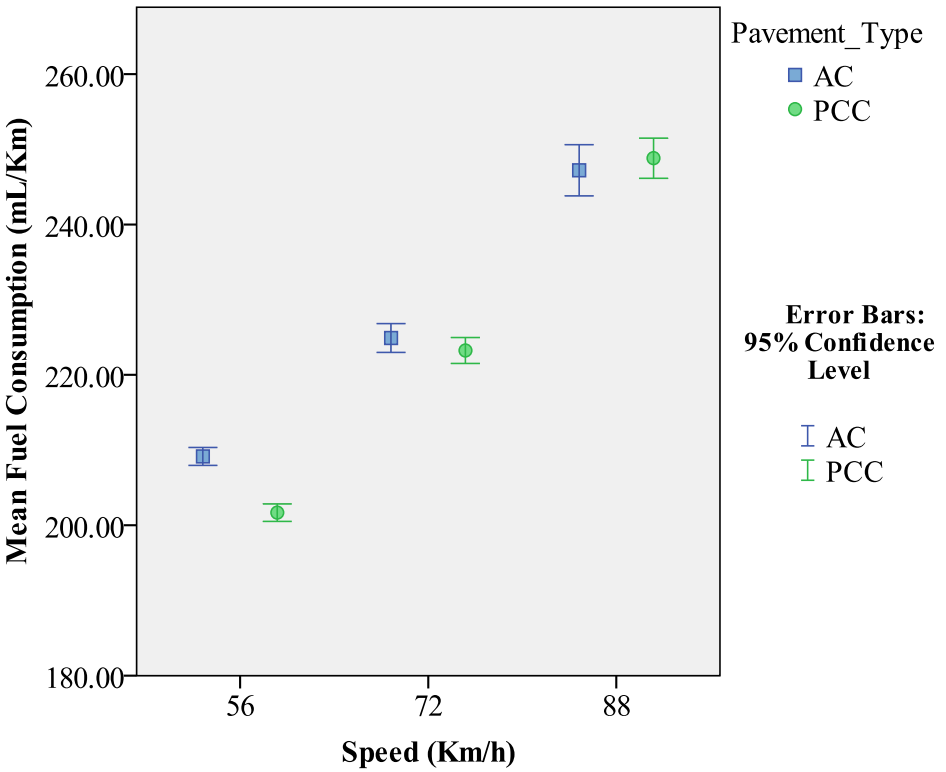



Figure 2 Mean and Standard Deviation of Fuel Consumption for Different Pavement Type and Speed – Light Truck

## CONCLUSION

The good quality of the data allowed for the investigation of the effect of pavement type on fuel consumption. A pairwise comparison was conducted between sections that have similar characteristics (grade, roughness and texture) but different pavement type. The results showed that the difference in fuel consumption between asphalt and concrete pavements is statistically significant at 95 percent confidence level for (loaded) light and heavy trucks at low speed (56 km/h) and summer conditions. Under these conditions, trucks driven over AC pavements will consume about 4% more than if they were driven over PCC pavements. These observations could be explained by the viscoelastic behavior of asphalt concrete. It should be noted that fuel consumption data for heavy truck in winter was not available. The analysis also showed that the mean differences of fuel consumption between asphalt and concrete pavements for passenger car, van and SUV are statistically not significant for both winter and summer conditions.

## ACKNOWLEDGMENTS

This research was conducted as part of the NCHRP project 1-45. The authors would like to acknowledge the financial support of the NCHRP of the National Academies, the input of the technical panel and the TRB senior program manager, Dr. Amir

Hanna. The authors also would like to thank the technical support from MDOT for conducting pavement condition testing, coordinated by Mr. Tom Hynes.

## REFERENCES

1. American Automobile Association (AAA), "*Your driving costs*", AAA's annual guide, 2009.

2. Jonsson, P. and Hultqvist, B., "Measurement of Fuel Consumption on Asphalt and Concrete Pavements", North of Uppsala, Swedish National Road and Transport Research Institute, Linköping, Sweden, 2008.

3. OECD, "Full-scale Pavement Test", Report IIRD No.837329, Organization for Economic Cooperation and Development, Paris, France, 1991.

4. Sandberg, Ulf S. I. Road Macro- and Megatexture Influence on Fuel Consumption. ASTM STP 1031 page 460-479, USA 1990.

5. Sumitsawan, P., Romanoschi, S., and Ardekani, S. A., "Effect of Pavement Type on Fuel Consumption and Emissions", Proceedings of the 2009 Mid-Continent Transportation Research Symposium, Ames, Iowa, August 2009.

6. Taylor, G., Marsh, P., and Oxelgren, E., "Effect of Pavement Structure on Truck Fuel consumption - Phase I and II", National Research Council's Centre for Surface Transportation Technology (CSTT), Report # CSTT- HVW-CTR-041, August 2000.

7. Taylor, G. and Patten, J., "Effects of Pavement Structure on Vehicle Fuel Consumption - Phase III", National Research Council's Centre for Surface Transportation Technology (CSTT), Report # CSTT-HVC-TR-068, January 2006.

8. VEROAD® Calculations. "Maximum energy dissipation when driving on asphalt pavement versus driving on rigid cement concrete", NPC, Netherlands Pavement Consultants, Netherlands 2002.

9. Zaabar, I. and Chatti, K., "Calibration of HDM4 Models for Estimating the Effect of Pavement Roughness on Fuel Consumption for US Conditions", In Press, Transportation Research Record: Journal of the Transportation Research Board Journal of the Transportation Research Record, Volume 2010.

10. Zaniewski, J. P., Butler, B. C., Cunningham, G., Elkins, G. E., Paggi, M. S., and Machemehl, R., "Vehicle Operating Costs, Fuel Consumption, and Pavement Type and Condition Factors." FHWA-PL-82-001, Texas Research and Development Foundation, Austin, 1982.

11. Zaniewski, J. P., "Effect of Pavement Surface Type on Fuel Consumption", Report # SR289.01P, Portland Cement Association, Skokie, Illinois, 1989.

Table 6 Summary of Pairwise Comparison for All Vehicles

| Vehicle class | Speed (km/h) | Mean fuel consumption (mL/km) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Summer | | | | | Winter | | | | |
| | | AC | PCC | Mean Difference | Number of data points | Significance (p-value)* | AC | PCC | Mean Difference | Number of data points | Significance (p-value)* |
| Medium car | 56 | 52.9 | 53.1 | -0.2 | 138 | 0.71 | 56.2 | 56.1 | 0.2 | 138 | 0.8 |
| | 72 | 63.9 | 64.6 | -0.7 | 138 | 0.22 | 67.2 | 68.2 | -1.0 | 138 | 0.1 |
| | 88 | 78.9 | 79.1 | -0.2 | 138 | 0.9 | 82.6 | 82.4 | 0.2 | 138 | 0.8 |
| Van | 56 | 81.8 | 81.1 | 0.7 | 138 | 0.35 | 85.7 | 85.6 | 0.2 | 138 | 0.8 |
| | 72 | 96.9 | 97.6 | -0.7 | 138 | 0.38 | 102.0 | 103.0 | -1.1 | 138 | 0.2 |
| | 88 | 113.7 | 115.3 | -1.6 | 138 | 0.29 | 119.1 | 121.4 | -2.3 | 138 | 0.2 |
| SUV | 56 | 101.7 | 100.2 | 1.5 | 138 | 0.2 | 106.6 | 106.1 | 0.5 | 138 | 0.6 |
| | 72 | 119.9 | 119.1 | 0.8 | 138 | 0.4 | 125.8 | 125.5 | 0.3 | 138 | 0.8 |
| | 88 | 159.6 | 160.5 | -0.9 | 138 | 0.7 | 164.8 | 162.6 | 2.2 | 138 | 0.3 |
| Light truck | 56 | 156.7 | 151.1 | 5.6 | 138 | 0 | 159.6 | 159.3 | 0.2 | 138 | 0.9 |
| | 72 | 188.4 | 187.8 | 0.6 | 138 | 0.6 | 198.2 | 198.3 | 0.0 | 138 | 1.0 |
| | 88 | 219.9 | 225.2 | -5.3 | 138 | 0.1 | 227.2 | 228.1 | -1.0 | 138 | 0.7 |
| Articulated truck | 56 | 209.4 | 201.4 | 8 | 138 | 0 | N/A | N/A | N/A | N/A | N/A |
| | 72 | 225.2 | 222.9 | 2.3 | 138 | 0.2 | N/A | N/A | N/A | N/A | N/A |
| | 88 | 247.6 | 248.4 | -0.8 | 138 | 0.9 | N/A | N/A | N/A | N/A | N/A |

*If more than 5%, the mean difference is considered statistically significant.

# EVALUATION OF PAVEMENT TEMPERATURES OF VARIOUS PAVEMENT SECTIONS

Adriana Vargas-Nordcbeck[1] and David H. Timm[2], P.E., A.M. ASCE

[1]Graduate Research Assistant, Dept. of Civil Engineering, Auburn University, 238 Harbert Engineering, Auburn, AL 36849; PH (334)-844-1596; email: vargaad@auburn.edu
[2]Associate Professor, Dept. of Civil Engineering, Auburn University, 238 Harbert Engineering, Auburn, AL 36849; PH (334)-844-6282; email: timmdav@auburn.edu

## ABSTRACT

Pavement temperature is an important factor of the urban heat island (UHI) effect and it also plays an important role in pavement design. The objective of this study was to compare the pavement temperatures and heat absorbed for nine different pavement sections placed during construction of Phase IV at the National Center for Asphalt Technology (NCAT) Test Track. Temperature probes were installed in each section to measure temperatures at the top, middle and bottom of the asphalt concrete (AC) and three inches into the underlying aggregate base layer. Temperatures were recorded on an hourly basis over a one month period from August to September 2009. Results indicated that an increase in density may also increase pavement temperatures and stored energy. Additionally, changing components or the structure of the asphalt mixture may change its thermophysical properties as well as the amount of energy the pavement is able to store.

## INTRODUCTION

The urban heat island (UHI) effect occurs when temperatures in a metropolitan area run significantly warmer than its surrounding rural areas and is caused mainly by the modification of land surface by urban development and the use of materials that effectively retain heat. Studies have shown that such an effect will result in higher energy and water consumption, as well as the acceleration of smog production (Gui et al., 2007). The use of cool pavements has been shown to mitigate this effect by increasing solar reflectivity to reduce surface heating and also by promoting cooling through increased air filtration and evaporation (Drainscape, 2010).

Research conducted on the thermophysical properties of asphalt pavements suggests that thermal conductivity, thermal diffusivity and heat capacity have an effect on pavement temperature, but changes on these properties alone do not produce a significant impact on the UHI effect (Gui et al., 2007). Highter and Wall analyzed the thermal properties of different asphaltic concrete mixes and determined that for limestone mixes thermal conductivity depended on asphalt content and aggregate gradation, but for lightweight aggregate mixes the conductivity showed little variation with asphalt content. The specific heat and thermal diffusivity of the mixes varied

with mineralogy and gradation but changed little with asphalt content (Highter and Wall, 1984). Additionally, no correlation has been found between pavement density and these parameters (Luca and Mrawira, 2005).

Pavement temperature is also an important variable for pavement design. It is used to specify performance grades for asphalt binders and modeling parameters for pavement performance and design. AC stiffness and strains are directly related to mid-depth pavement temperature and can be used to predict pavement performance (Priest and Timm, 2006). The objective of this study was to compare the pavement temperatures and heat absorbed for different pavement sections placed at the National Center for Asphalt Technology (NCAT) Test Track.

## SITE DESCRIPTION

The National Center for Asphalt Technology (NCAT) Test Track located in Opelika, Alabama is a state-of-the-art, full-scale, closed loop accelerated testing facility. The Test Track was designed to test 46 pavement sections (200 feet in length) with 10 million equivalent single axle loads (ESALs) of traffic over two years. The test sections allow pavement engineers and researchers to study pavement responses and distresses to make pavement design more economical and efficient (Willis et al., 2009).

Construction of the Phase IV research cycle was carried out during the summer of 2009 and the Test Track was opened to traffic on August 28, 2009. As part of this test cycle, new structural sections were placed that included non-conventional mixtures such as warm mix asphalt (WMA), mixes with high reclaimed asphalt pavement (RAP) content, and mixtures containing modified and alternative binders. A control section and a control section with porous friction course were also constructed. Table 1 shows a description of the test sections evaluated in this study. The thicknesses shown are an average of twelve measurements taken along each test section and the total density of the asphalt layer was taken as a weighted average of the densities of each asphalt lift. Though all sections met their respective density requirements, it is important to note that section S12 achieved the highest total layer density.

**Table 1. Section Description**

| Section Number | Description | As-built thickness, cm | Total Layer Density, g/cm$^3$ |
|---|---|---|---|
| N6 | Shell Thiopave® – Sulfur Modified WMA | 17.5 | 2.37 |
| N7 | Kraton – 7.5% SBS Modified Asphalt | 14.7 | 2.34 |
| N10 | 50% RAP HMA | 18.0 | 2.39 |
| N11 | 50% RAP WMA | 18.0 | 2.38 |
| S8 | Control with Permeable Surface | 17.8 | 2.35 |
| S9 | Control | 17.8 | 2.29 |
| S10 | Foamed WMA | 17.8 | 2.36 |
| S11 | Additized WMA | 17.5 | 2.35 |
| S12 | 25% asphalt replacement with Trinidad Lake Asphalt Pellet | 17.5 | 2.41 |

The Campbell-Scientific model 108 thermistor temperature probes used in the previous three test cycles at the Test Track were again selected for use in 2009. Individual probes were bundled together and custom built for each section based on as-built thicknesses to measure temperature at the top, middle and bottom of the AC and three inches (7.6 cm) into the underlying aggregate base layer. After paving was complete, a vertical hole was drilled just outside the edge paint stripe, approximately two feet from the edge of the finished pavement and the bundled array of thermistors was coated with roofing cement and inserted into the hole (Timm, 2009).

Temperature readings were recorded on an hourly basis and the minimum, maximum and average temperatures for each depth were stored in a database. Additionally, environmental conditions such as air temperature, wind speed and relative humidity were also recorded at an onsite weather station.

## METHODOLOGY

This study was based on temperature readings taken between August 28 and September 28, 2009 for the test sections mentioned above. Mid-depth temperatures were used as an estimate of the average temperature of the asphalt layer. First, the hourly temperature differences between S9 (control) and all other sections were calculated and distribution curves were plotted and used for comparison. The daily ΔT, hourly ΔT and cooling and heating rates (change in temperature per hour) were also calculated for all sections and compared at 95% confidence level using t-tests and the Tukey method. Figure 1 illustrates the concepts used in this study.

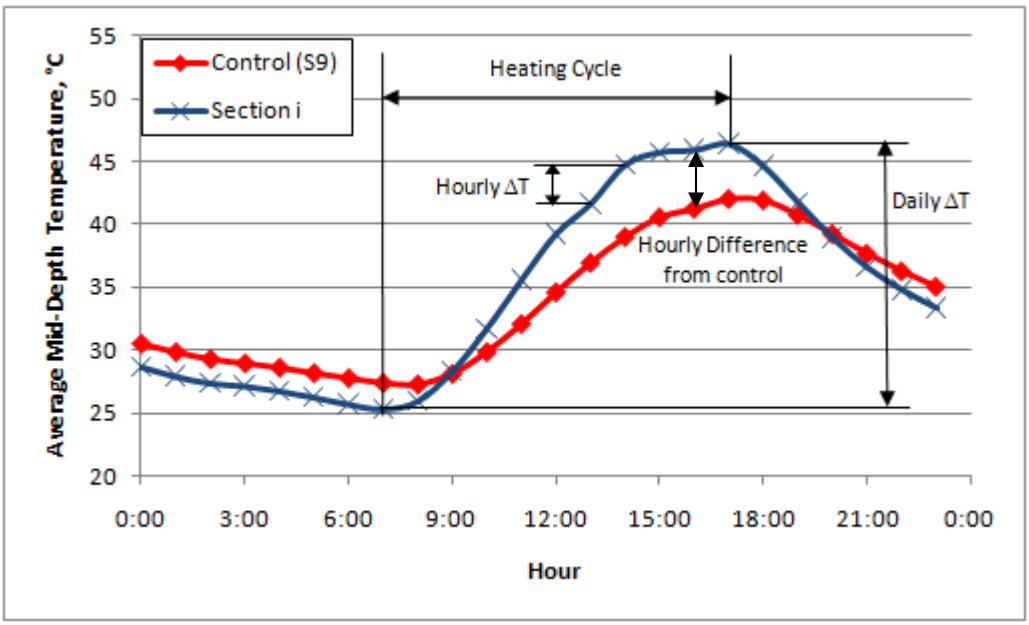



**Figure 1. Example of Variables Used in Analysis**

The pavement warms as it absorbs energy from solar radiation and later radiates heat when the cooling cycle begins. The energy released by the pavement contributes to the UHI effect by causing an increase in energy demand for air conditioning (Kevern et al.). The amount of energy absorbed by the pavement was calculated using the following equation:

$$\Delta E = C_v \cdot \Delta T \cdot \rho \cdot h \quad (1)$$

Where:
ΔE = amount of energy stored during the daily heating cycle per unit area of the asphalt layer, J/cm$^2$.
$C_v$ = volumetric heat capacity of the asphalt layer, J/g°C.
ΔT = change in mid-depth temperature of the asphalt layer, °C.
ρ = density of the asphalt layer, g/cm$^3$.
h = as-built thickness of the asphalt layer, cm.

Measuring the heat capacity of the asphalt was beyond the scope of this study. Instead, $C_v$ was assumed to be equal for all sections and ΔE of each section was expressed as a percentage of the control section, as shown in Equation 2. The subscript *i* indicates the section being compared to the control. All statistical evaluations of ΔE were based on these percentages.

$$\%\Delta E = \frac{C_{v_i} \cdot \Delta T_i \cdot \rho_i \cdot h_i}{C_{v_{S9}} \cdot \Delta T_{S9} \cdot \rho_{S9} \cdot h_{S9}} = \frac{\Delta T_i \cdot \rho_i \cdot h_i}{\Delta T_{S9} \cdot \rho_{S9} \cdot h_{S9}} \quad (2)$$

## RESULTS

Figure 2 shows the cumulative distribution of hourly temperature differences of each section compared to S9. Positive values indicate the section is warmer than S9 while negative values indicate cooler. In general, all sections tend to have warmer temperatures than the control section, but the differences do not exceed 2.5°C. The table inset in Figure 2 lists the percentage of readings that exceed 0°C. For example, 80% of the S8 measurements were warmer than the corresponding S9 measurements. However, when the temperatures were cooler than the control the differences were as much as 7.3°C. It can also be observed that section S11 has the smallest differences from the control section, and sections S10 and S12 have considerably larger temperature differences toward the cooler end of the spectrum. The increased difference between the "warmer" and "cooler" temperatures shown in these sections suggests that while the pavement warms up more (absorbs more energy) than the control section, it also cools down (releases energy) even more that the control section.

The amount of energy stored is directly proportional to the temperature differential in the asphalt layer during the daily heating cycle. Figure 3 shows the average daily temperature change for all sections. It was found that the temperature differential was

significantly higher for sections S10 and S12 relative to the control section. This could be caused by the higher pavement densities, or differences in heat capacity, compared to the control section.

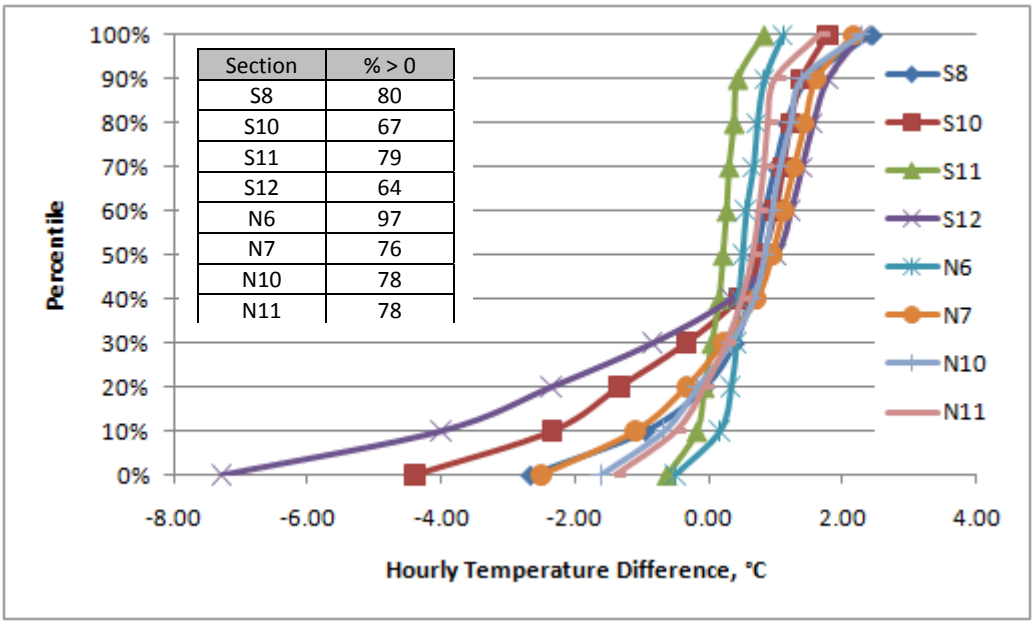



**Figure 2. Percent Distribution for Hourly Temperature Differences from Control Section**

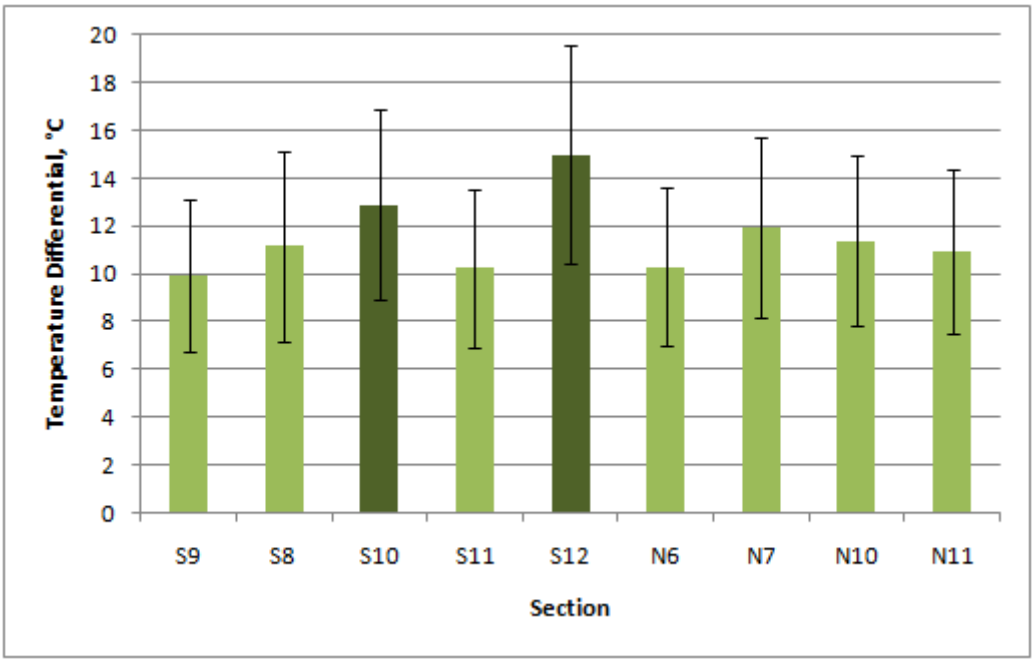



**Figure 3. Average Daily Temperature Differential During the Pavement Heating Cycle (darker bars denote statistically significant differences)**

Figure 4 shows an example of the daily heating cycle of the pavement sections that are significantly different than the control for September 25, 2009, which registered the highest air temperature for the time period of this study. It can be observed that the mid-depth pavement temperatures are always warmer than the air temperature. Early in the day, the temperatures are very similar among pavement sections, but at the hottest period the differences become significant, with sections S10 and S12 approximately 3 to 5 degrees warmer than S9, respectively. Once the cooling cycle begins, all pavement sections reach similar temperatures again.

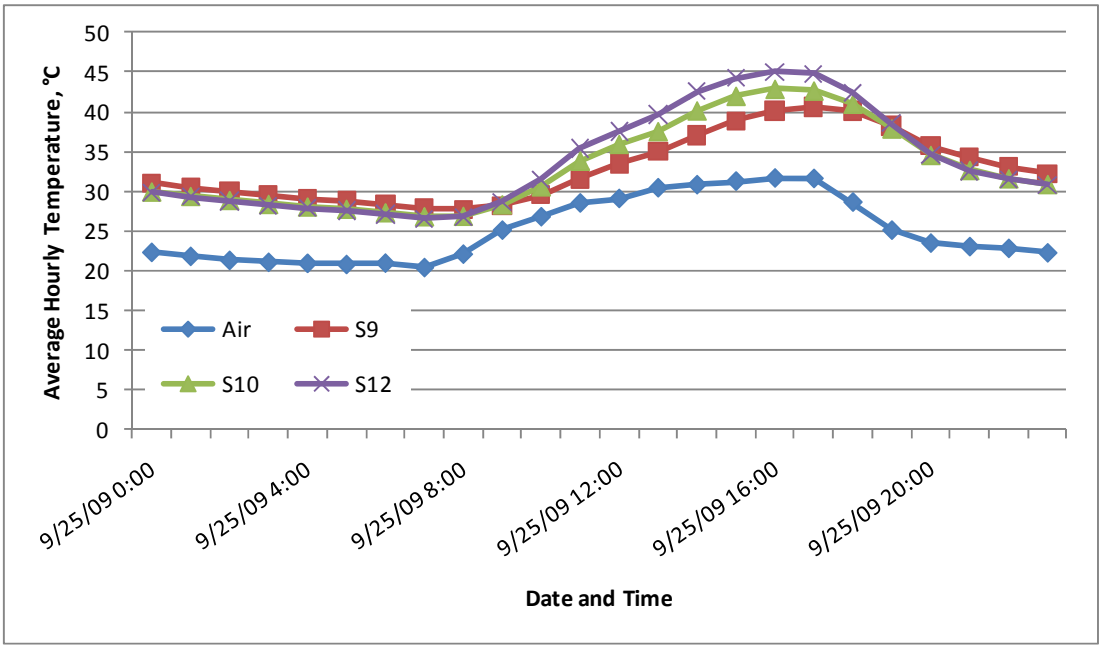



**Figure 4. Pavement Temperatures for 9/25/2009.**

The temperature differential from one hour to the next was calculated for all sections. As expected from previous results, sections S10 and S12 have significantly larger rates of heating and cooling, as shown in Figure 5. Although the differences are statistically significant, they do not exceed 0.6°C/hr, which from a practical point of view may not require any changes in mixture or pavement design. For example, 0.6°C/hour over a nine-hour heating cycle is still within 6°C which is the Superpave performance graded binder specification interval.

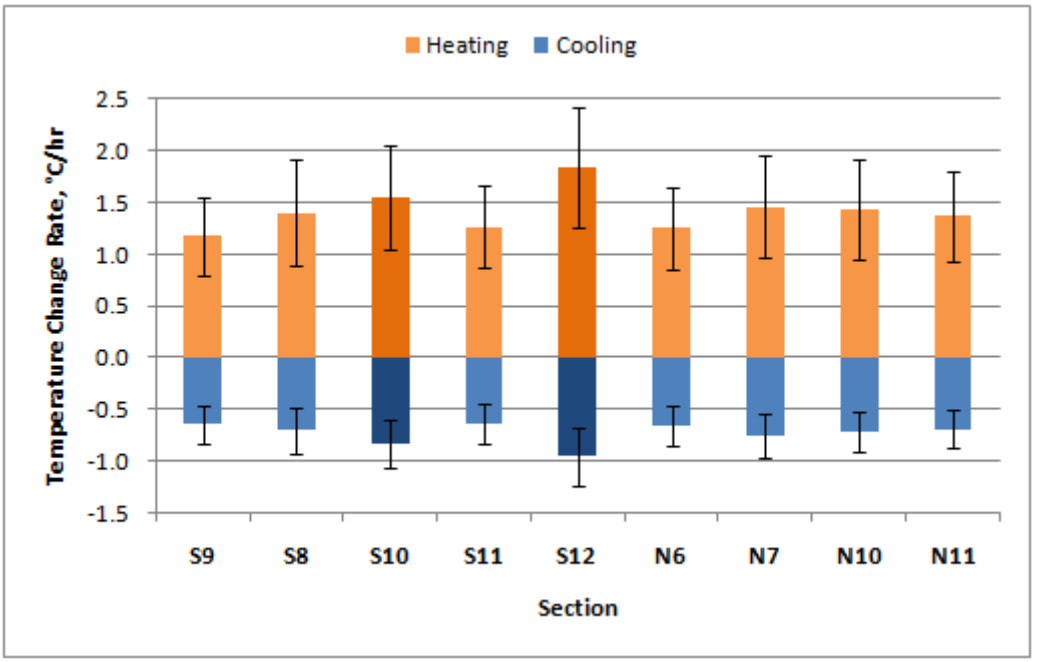



**Figure 5. Rate of Temperature Change (darker bars denote statistically significant differences)**

An example of the cumulative hourly ΔT is shown in Figure 6 for the period of 9/25/2009 to 9/28/2009. The slope of the curve is an indicator of how quickly temperature changes hourly for a given section. The differences in the peaks provide a comparison between the different sections. As stated above, sections S10 and S12 have the largest deviation from the control section, which indicates that they warm up faster. Section S8 exhibits the largest difference during the cooling cycle, meaning that the section with the porous surface cools down faster.

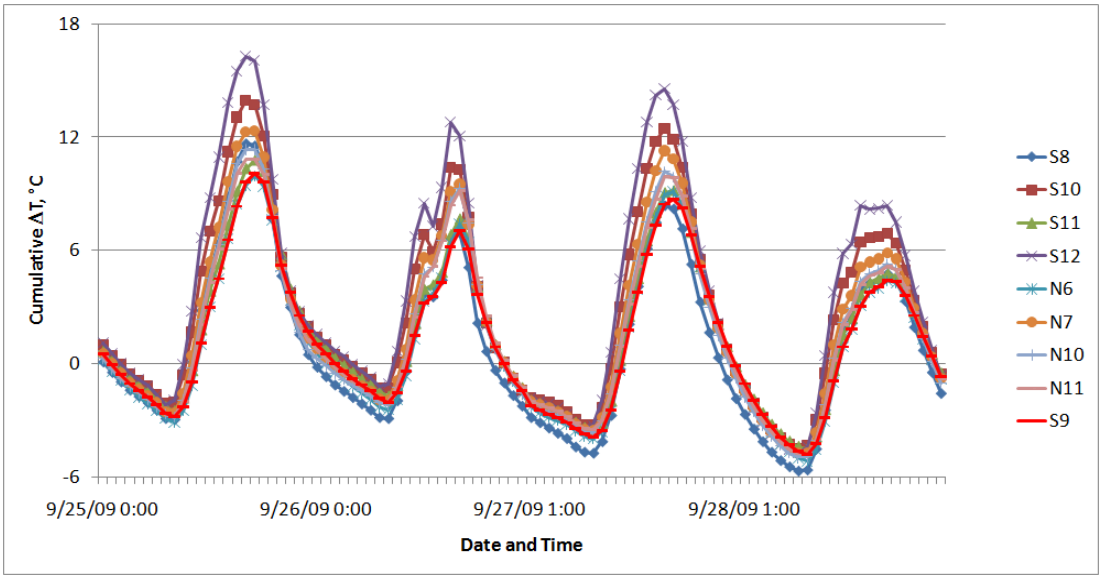



**Figure 6. Cumulative Temperature Differential of Pavement Sections.**

The average daily amount of absorbed energy, as calculated by equation 2 for August 28 through September 28, as a percentage of S9 is shown in Figure 7. Sections S8, S10, S12, N10 and N11 store a significantly higher amount of energy than the control section. This was somewhat expected for sections S10 and S12 because of the higher temperature differentials mentioned above. Since the pavement thicknesses are similar for all sections and density has already been accounted for, it is possible that the assumption that the specific heat of the asphalt is the same as the control is not accurate, especially in the case of sections S8 (porous surface), N10 and N11 (high RAP sections), which do not have significantly higher ΔTs but do show a significant increase in ΔE. The amount of absorbed energy appears to be affected by a combination of pavement density and specific heat.

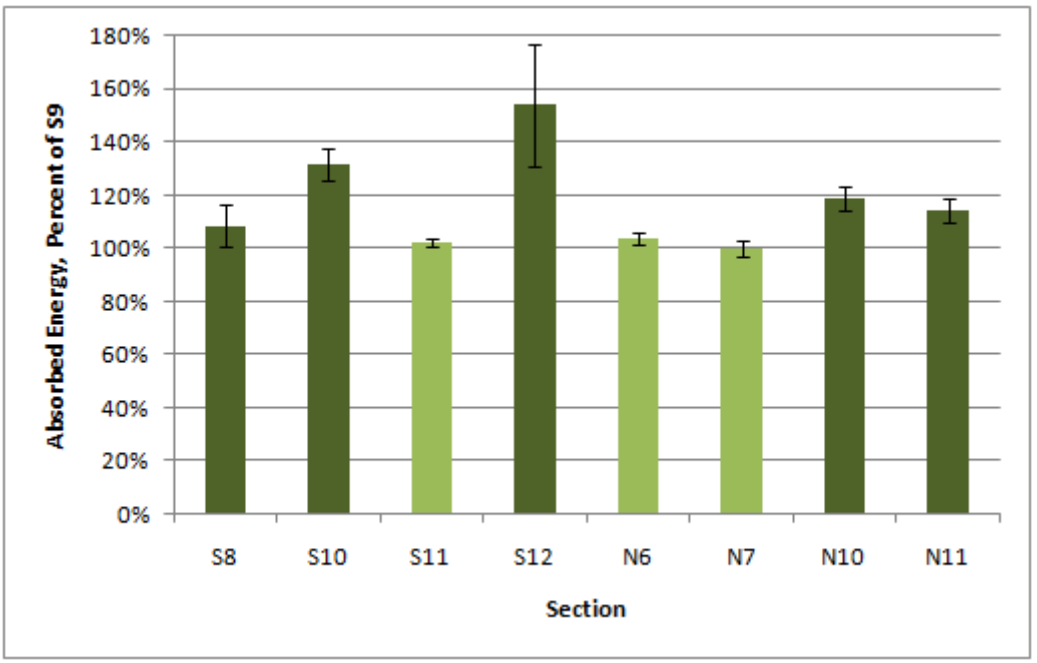



**Figure 7. Absorbed Energy as a Percent of the Control Section (darker bars denote statistically significant differences)**

## CONCLUSIONS

The following conclusions were reached from this study:

- The use of different types of asphalt mixtures has an effect in mid-depth pavement temperature, but the differences do not appear large enough to require changes in the pavement and mixture design process.
- An increase in pavement density can lead to higher mid-depth temperatures and larger rates of heating and cooling, as well as an increased amount of stored energy.

- Use of certain materials such as RAP or aggregate structure such a porous friction courses may change the thermophysical properties of the asphalt layer and consequently the amount of stored energy. It is recommended that the specific heat of the different mixtures be measured to confirm the effect of these changes.

## ACKNOWLEDGEMENTS

The authors wish to thank the following state departments of transportation for their continued cooperation and support of this research: Alabama, Florida, North Carolina, Oklahoma and Tennessee. The authors are also appreciative of Kraton Polymers LLC, Lake Asphalt of Trinidad and Tobago, Ltd. and Shell Sulphur Solutions for their support and cooperation. Finally, the Federal Highway Administration also deserves special recognition for their support and cooperation.

## REFERENCES

Drainscape. "Brochure on the Use of Cool Pavements to Reduce the Urban Heat Island Effect, prepared by the Town of Gilbert", http://www.drainscape.com/custdocs/Gilbert-Heat-Island-Cool-Pavements-Brochure.pdf. Accessed June, 2010.

Gui, J., P. E. Phelan, K. E. Kaloush and J. S. Golden. "Impact of Pavement Thermophysical Properties on Surface Temperatures", Journal of Materials in Civil Engineering, Vol. 19, No. 8, August 1, 2007.

Highter, W. H, and D. J. Wall. "Thermal Properties of Some Asphaltic Concrete Mixes", Journal of the Transportation Research Board, No. 968, National Academies, Washington, D.C., 1984, pp. 38-45.

Kevern, J. T., L. Haselbach, V. R. Schaefer. "Hot Weather Comparative Heat Balances in Pervious Concrete and Impervious Concrete Pavement Systems", 2nd International Conference on Countermeasures to Urban Heat Island, Berkeley, CA, 2009.

Luca, J. and D. Mrawira. "New Measurement of Thermal Properties of Superpave Asphalt Concrete", Journal of Materials in Civil Engineering, Vol. 17, No. 1, February 1, 2005.

Priest, A.L. and D.H. Timm. "Methodology and Calibration of Fatigue Transfer Functions for Mechanistic-Empirical Flexible Pavement Design", Report No. 06-03, National Center for Asphalt Technology, Auburn University, 2006.

Timm, D.H. "Design, Construction and Instrumentation of the 2006 Test Track Structural Study", NCAT 09-01, National Center for Asphalt Technology, Auburn University, 2009.

Willis, R., D. Timm, R. West, B. Powell, M. Robbins, A. Taylor, A. Smit, N. Tran, M. Heitzman, A. Bianchini. “Phase III NCAT Test Track Findings”, NCAT 09-08, National Center for Asphalt Technology, Auburn University, 2009.

# Performance Evaluation of Lime and Cement Treated Soil Layers under Laboratory and Full Scale Accelerated Pavement Testing

Zhong Wu[1], Gavin Gautreau[2] and Zhongjie Zhang[3]

[1]Louisiana Transportation Research Center, Baton Rouge, LA 70808, (225)767-9163, Zhong.Wu2@LA.GOV
[2]Louisiana Transportation Research Center, Baton Rouge, LA 70808, (225)767-9110, Gavin.Gautreau@LA.GOV
[3]Louisiana Transportation Research Center, Baton Rouge, LA 70808, (225)767-9162, Doc.Zhang@LA.GOV

## ABSTRACT

This paper documents a research effort conducted regarding chemical stabilization of the naturally wet and problematic clayey soils typically found as subgrade in south Louisiana, and provides detailed information on experiment design, instrumentation, and conducted field and laboratory tests. The objectives of the study include the exploration and development of a methodology to build reliable and conservatively achievable subgrade layers, stabilized with cementitious agents at various field moisture contents so that a treated subgrade layer would not only provide a working table for pavement construction, but can also function as a pavement subbase layer that contributes to the overall pavement structural capacity. The laboratory and field results confirmed that among subbase treatments evaluated, cement-treated soil provided the best-performance followed by lime-treated soil.

## INTRODUCTION

Many Louisiana pavements were built in areas of naturally low shear strength and minimum bearing capacity. The wet climate (over fifty inches of rain per year) combined with the soft alluvial fine-grained soils exacerbate the potential of moisture sensitivity and bearing capacity problems for a subgrade layer. The Louisiana Department of Transportation and Development (LADOTD) historically provided a lime treatment item to be used at the discretion of project engineers when subgrade-pumping conditions were encountered. The common use of this item, regardless of soil type, is to "dry-out" wet areas and produce a temporary working platform (or "working table") for pavement construction. In the pavement design, no structural values (e.g. layer coefficient) are allowed to be assigned to such a working table layer. Another type of soil improvement is called "soil stabilization", where a higher

level of stabilizer is chosen to achieve a desired soil strength properties and durability [Prusinski et al. 1999].

A study conducted at the University of New Orleans indicated that fine-grained soils containing silt percentages of 50% or more, and a Plastic Index (PI) less than or equal to 10 should be considered to have a high pumping potential [McManis 2003]. To improve such soils different chemical additives–lime, lime-fly ash, Portland cement, and slag cement–were studied. The results indicated that cement or slag cement treated silty subgrade would provide more enhanced strength characteristic necessary for both short- and long-term benefits, as well as greater performance benefits than a lime treated subgrade in wet conditions [McManis 2003].

The concept of using cement to treat silty subgrade in wet conditions has been successfully implemented on LADOTD projects consistent with the "working table" concept to address the problem discussed. Both laboratory and field results indicated that cement treated silty soils have significantly higher modulus and strength values than lime treated silty soils. In addition, the experience from LADOTD and many other state agencies indicates that sandy soils perform well when treated with cement, while lime is more effective in treating heavy clays.

However, since most times a treated subgrade layer is called for at locations where a wet subgrade is encountered (i.e. the in situ moisture content of a soil is higher than its optimum moisture content), field compaction to reaching to the maximum dry density of a subgrade soil is not realistic or even achievable in some cases. In other words, the effectiveness of a treatment and the performance of a treated subgrade layer could be largely varied, depending on not only soil types but also in situ moisture contents.

## OBJECTIVES

The main objective of this study was to develop a methodology to build reliable and conservatively achievable subgrade layers stabilized with cementitious agents at various field moisture contents so that a treated subgrade layer would not only provide a working table for pavement construction, but can also function as a pavement subbase layer that contributes to the overall pavement structural capacity.

## EXPERIMENT DESIGN

This project included laboratory and field two testing programs. The laboratory experiment was to find out how to build a treated subgrade layer with reliable, consistent strength and performance. The objective of the field experiment was to verify, validate, and quantify the benefits of using a cement-treated subgrade in lieu of a lime-treated working table with respect to the overall pavement performance. Brief description of each program is given below.

**Laboratory testing.** Three soil types, representing the wet and weak subgrade soils commonly found in many field cases in south Louisiana, were selected for the laboratory testing program of this study. The three soils, designated as Soil I, II, and III, were all fine-grained in size and had low, medium, and high PIs, respectively.

Table 1 presents the physical properties of each soil. The corresponding particle size distribution curves are shown in Figure 1. As seen in Table 1, the maximum dry densities for the three soils were 108.0 pcf, 119.0 pcf, and 85.0 pcf, respectively, at the corresponding optimum moisture content ($w_{opt}$) of 17.5%, 13.5%, and 33.1%. According to the USCS and AASHTO classifications, the three soils are classified as CL/A-6, CL/A-7, and CH/A-7, respectively.

**Table 1. Physical Properties of Tested Soils.**

| Soil Type | Silt (%) | Clay (%) | LL (%) | PI (%) | $w_{opt}$ (%) | Maximum dry density, pcf | Sp. Grav | Classification USCS/AASHTO |
|---|---|---|---|---|---|---|---|---|
| Soil I | 64.5 | 27.5 | 34 | 12 | 17.5 | 108.0 | 2.65 | CL/A-6 |
| Soil II | 30.6 | 27.9 | 37 | 22 | 13.5 | 119.0 | 2.72 | CL/A-7 |
| Soil III | 13.7 | 81.9 | 83 | 49 | 33.1 | 85.0 | 2.73 | CH/A-7 |

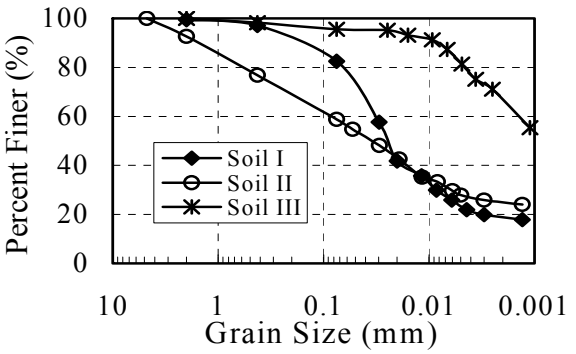



**Figure 1. Particle Size Distribution of Tested Soils**

To stabilize those soils, type I Portland cement and high-calcium hydrate lime were used as stabilization chemicals. Table 2 presents the stabilization factorial design of this study. As shown in the table, due to the consideration of stabilization effectiveness, cement was used to treat Soil I and Soil II, and lime was applied to only Soil III (the high PI soil). It also can be noted from Table 2 that various molding moisture contents were used to investigate the possible relationship between sample moisture contents and the corresponding sample strengths.

**Table 2 Factorial Design of Stabilization Schemes.**

| Soil type | Cement % by weight | Lime % by weight | Molding moistures, $w_n$ % | Curing, days |
|---|---|---|---|---|
| Low PI (Soil I) | 0, 4, 8, 12 | -- | From at least 6% below the optimum to too wet to mold | 7, 28 |
| Medium PI (Soil II) | 0, 4, 8, 12 | -- | | 7, 14, 28 |
| High PI (Soil III) | -- | 0, 5, 18, 26 | | 7, 28 |

The effectiveness of cement stabilization was evaluated through three

laboratory tests: unconfined compressive strength (UCS), tube suction (TS), durability. For the lime stabilization schemes, only UCS tests were performed. Briefly, UCS tests were used to determine strengths of treated/untreated soils according to ASTM D1633 (Standard Test Methods for Compressive Strength of Molded Soil Cement Cylinders). The TS tests were used to evaluate moisture susceptibility of soil with or without stabilization using a modified version of the Texas Transportation Institute method. The durability tests were performed according to ASTM D559 to evaluate a material's ability to withstand numerous wet and dry cycles.

**Field testing.** As outlined in Figure 2, two 13ft wide by 107.5ft long accelerated pavement testing (APT) sections were designed for this research study. The APT sections were constructed and tested at the Louisiana pavement research facility site at Port Allen of Louisiana. Each section had a 2.0 in. hot mix asphalt (HMA) top layer, an 8.5 in. crushed stone base layer, and a 12 in. treated soil layer. The constructed HMA layer used a 19-mm Level 2 Superpave mixture designed at 100 gyrations using the Superpave Gyratory Compactor. The crushed stone base used is a typical Class II base course specified by the LADOTD specification [LADOTD 2000].

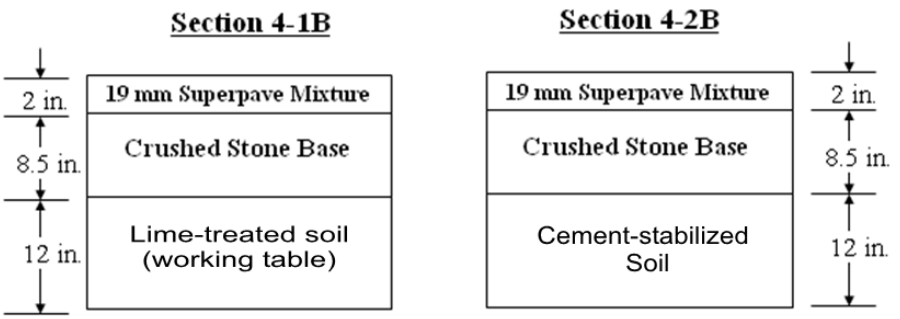



**Figure 2. Pavement Structures of APT Test Sections**

The only difference of two sections was the treated soil layers. The lime-treated soil layer of section 4-1B was an in-place 10 percent by volume lime-treated "working table" layer, whereas, the cement-stabilized soil (CSS) used in section 4-2B was an in-place 8 percent by volume cement-stabilized soil. A silty-clay soil was used in both treated soil layers as well as the beneath embankment subgrade. Basic properties of the silty-clay soil are presented in Table 3. It consists of 60.3 percent silt and 23.5 percent clay. The liquid limit and the plastic index (PI) are 31- and 11-percent, respectively.

**Table 3. Soil Properties.**

| Passing # 200 (%) | Clay (%) | Silt (%) | LL (%) | PI | $W_{opt}$ (%) | $\gamma$ (kN/ $m^3$) | Classification | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | USCS | AASHTO |
| 91 | 23.5 | 60.3 | 31 | 11 | 18.5 | 17.1 | CL-ML | A-6 |

It should be noted that the lime treated working table layer used in section 4-1B represents a current construction practice of LADOTD. However, an eight percent cement treatment of the subgrade soil used section 4-2B was determined in the laboratory based on a design criterion of a minimum unconfined compressive strength of 150 psi under an in-situ moisture condition.

In this study, the APT test sections were accelerated loaded using the Louisiana Accelerated Loading Facility (ALF). The ALF device models one-half of a single axle load with a dual-tire wheel assembly (using two MICHELIN radial 11R22.5 tires with target cold tire inflation pressure of 105 psi). The magnitude of an ALF wheel load per load application can be adjusted from 9,750 lb to 18,950 lb. In this experiment, both sections had a rutting failure (i.e. the average rut depth of 0.5 inch) at the end of ALF loading. Specifically, Section 4-1B was failed within 150,000 ALF passes of 9,750 lb per load application, whereas, Section 4-2B had a loading history of 200,000 passes of 9,750-lb ALF load in the beginning, and plus subsequent 25,000 passes of 12,050-lb ALF load, and finally plus another 75,000 passes of 14,350-lb ALF load.

During the APT testing, vertical compressive stresses and in-depth deflection profiles under the ALF wheel loading were monitored through two Geokon 3500 pressure cells and one multi-depth deflectometer (MDD) installed on each of the two sections. More details about the instrumentation of this study may be referred to elsewhere [Wu et al. 2006]. In addition, the falling weight deflectometer (FWD) test and surface rutting profiles were collected at an internal of every 25,000 ALF passes.

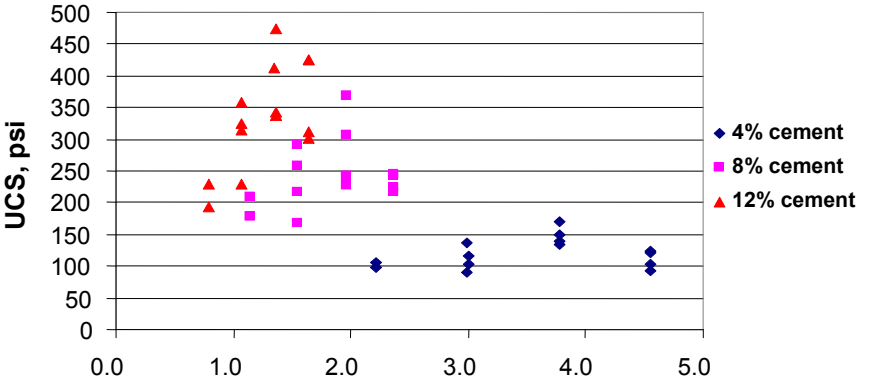
## RESULTS OF LABORATORY EVALUATION

**Summary of cement stabilization results.** As expected, laboratory test results confirmed that cement content, molding moisture content, and curing period all affect the UCS strengths of cement stabilized soils. A higher cement content and longer curing time generally can result in a higher strength for stabilized soils. The researchers' experience with cement indicates that the water-cement ratio is main controlling factor for the strength of cement-stabilized materials. It was found that, however, cement-stabilized soils performed differently on the wet and dry sides of the compaction curves (Note: the wet or dry sides represent a moisture content either higher or lower than the optimum moisture content in a compaction curve).

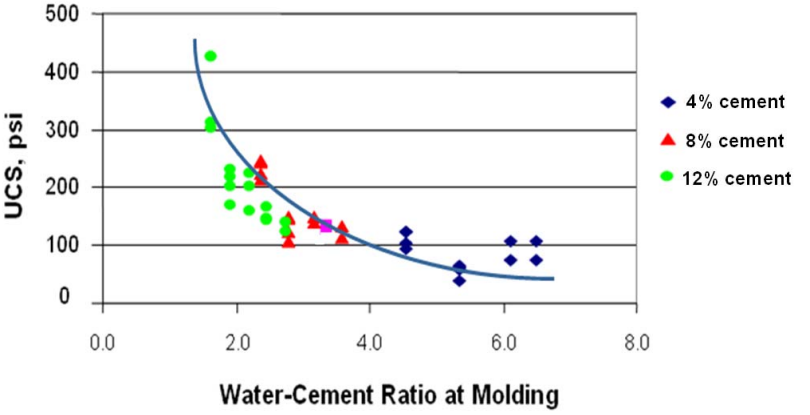
Figure 3 presents the UCS results for cement stabilization of Soil I. It can be seen in Figure 3 that the UCS strength has a good correlation with the water-cement ratio at molding on the wet side of compaction curves, but no strong correlation with the water-cement ratio on the dry side of compaction curves. Similar results were obtained for the cement stabilization of Soil II. More detailed results and discussion may be referred to elsewhere [Gavin et al 2010].

It is concluded that on the wet side of the compaction curve the water-cement ratio is a good indicator for the strength of cement-mixed soils. The strength will increase with the decrease of water-cement ratio. The strong correlation of water-cement ratio with UCS values on the wet side of compaction curves provides a quick means of approximating strength. As a result of this study, a preliminary guideline for

cement stabilization under a wet subgrade condition (e.g. in situ soil moisture is on the wet side of a compaction curve) may be set as follows: cement-soil mixtures with a water-cement ratio of 2 shall have a minimum 28-day UCS of 150 psi while cement-treated soils with a water-cement ratio of 3 will have a minimum 28-day UCS of 100 psi. With this guideline, it is expected that a reliable cement stabilized subgrade layer with consistent strength will be able to design and construct under an in situ moisture condition.




**(a) Samples on Dry Side of Compaction Curves**




**(b) Samples on Wet Side of Compaction Curves**

**Figure 3 UCS Results of Cement Stabilized Soil I**

The TS and durability tests on cement stabilized soils of this study generally indicate that there are good correlations among the soil-cement loss (from durability tests), the maximum dielectric values (from TS tests), and the 7-day UCS. These good relationships provide some support to the equivalency of wetting-drying durability, tube suction, and 7-day UCS tests. More detailed results and discussions may be referred to elsewhere [Gavin et al. 2010].

**Summary of lime treated soils.** The UCS results of different percentages (5%, 18%, and 26%) lime-treatment of Soil III at 28-day 100% room curing are presented in Figure 4. As expected, the UCS strengths of treated soils increase as the increase of

lime percentage. However, the difference between the higher additive percentages is less apparent. Also, lime treatment tends to shift optimum moisture contents to the wet side, and reduces maximum dry densities from the untreated soil values.

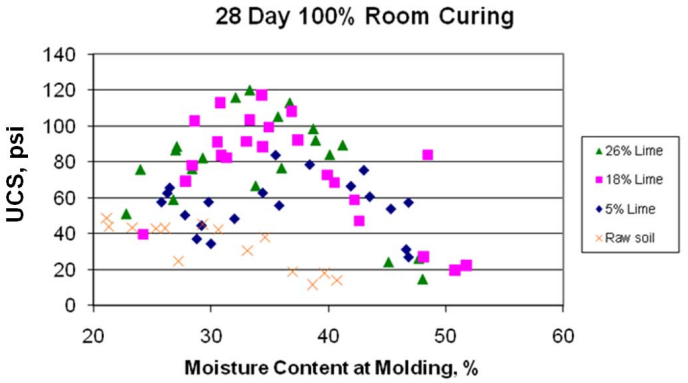



**Figure 4. UCS of Lime-treated Soil III verse Molding Moisture**

Figure 5 presents the relationships between the UCS and dry unit weights of lime-treated soils. As seen in the figure, increasing of dry unit weights indicates an increasing in UCS strengths, with the strengths peaking at roughly 80 pcf. In addition, it was found that lime treated soil samples prepared near maximum dry densities generally produced the highest UCS test results.

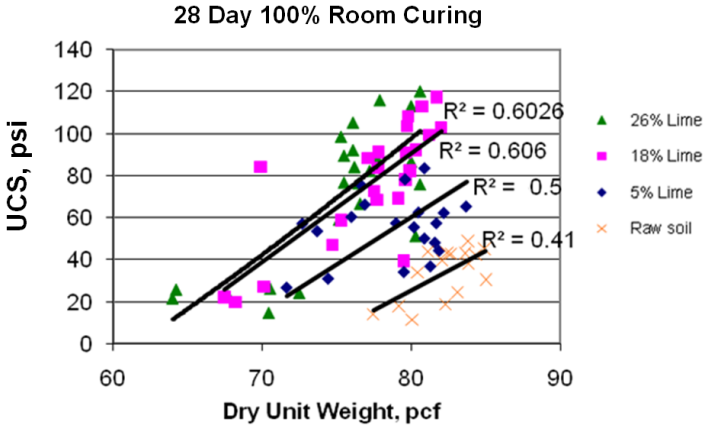



**Figure 5. UCS of Lime-treated Soil III verse Dry Unit Weight**

In summary, lime treatment was proved to have reduced clay PIs and lead to greater workability and reduction in volume change potential. Heavy clay soil treated with 18% lime and 26 % lime (by weight) generally produced similar UCS results, reflecting saturation and the limits of necessary and effective lime. Although a laboratory UCS strength of 150 psi may be attainable, more research on how to determine the right combination of lime, water, and curing techniques to achieve a reliable strength result is still warrant.

## APT TEST RESULTS

It should be mentioned that the UCS results obtained from those field mixed and

laboratory compacted samples indicated that the average 28-day UCS strengths for the lime-treated soil of section 4-1B and cement-stabilized soil of section 4-2B were 34.6 psi and 139.4 psi, respectively. While the 10% by volume lime treated working table layer did not or only slightly enhance the subgrade strength, the 8% by volume cement stabilized subgrade layer did improve the strength of a soil layer significantly. By considering in situ moisture content of the soil was about 20.5 percent, the water-cement ratio would be close to 2.5. Thus, the average UCS strength of 139.4 psi for the cement stabilized soil layer of section 4-2B could match well with the criteria set by this study: minimum UCS of 100psi for water-cement ratio of 3 and minimum UCS of 150 psi for water-cement ratio of 2.

**Pavement rutting life of test sections.** As mentioned earlier, the APT results generally indicated that both test sections had a rutting failure. Figure 6 presents variation of measured average rut depths along the number Equivalent Single Axle Loads (ESALs) for the two sections tested. As shown in Figure 6, section 4-2B with a cement stabilized subbase layer performed significantly better that section 4-1B with a lime-treated working table layer. The rutting lives (i.e. the number of load repetitions for reaching to a rutting limit of 0.5-in) of sections 4-1B and 4-2B were found to be 121,000-, and 786,000- ESALs, respectively.

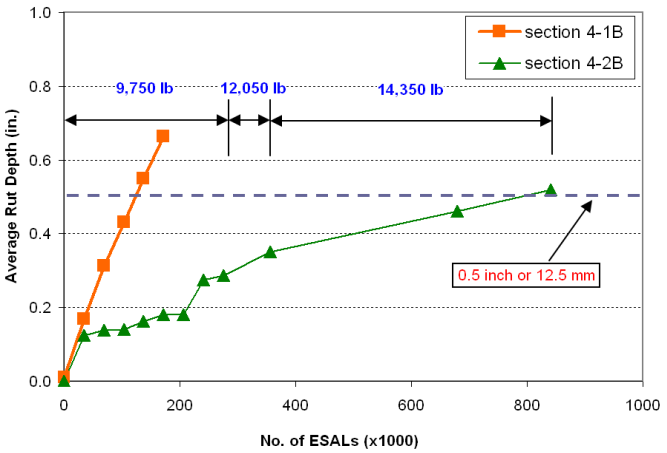



**Figure 6. Rut Depth Development on Test Sections**

In a previous ALF research, a term of "ESAL advantage" was introduced to quantify different performance of two pavement materials under an accelerated loading [King et al. 2004]. "ESAL advantage" is defined as a ratio of pavement lives for two similar pavement structures (e.g. sections 4-1B and 4-2B) under an APT experiment. The "ESAL advantage" implies that one material will perform how many times better than another material under an APT testing within a certain pavement structure. Based the APT results obtained in this study, section 4-2B with a 12-in cement stabilized subbase would have 6.5:1 ESAL advantage over section 4-1B with a 12-in lime treated working table layer. In addition, a structural layer coefficient for the cement-stabilized soil used in the APT experiment was predicted to be 0.06; while for lime-treated soil, no structural contribution should be allowed [Gavin et al. 2010].

**Instrumentation results.** Under the ALF self-load of 9,750-lb, the average vertical compressive stresses at bottom of the base layers were 32.4-, and 18.6- psi for sections 4-1B and 4-2B, respectively, whereas, the corresponding values on top of subgrades were 0.6-, and 0.7- psi, respectively. Such results indicate that both test sections received a comparable compressive stress on top of subgrades. However, the significantly smaller stress value developed at the bottom of the stone base layer of section 4-2B. Such results seemed to indicate that a stronger support underneath the stone layer of section 4-2B would have provided a stress relief for this layer. Another explanation could be the stress-hardening property of a crushed stone base. The difference in vertical stresses also indirectly explains why section 4-2B had a longer pavement life than section 4-1B.

The MDD permanent deformation results indicated that the majority (about 85%) of surface rut depths were directly come from the crushed stone layer on section 4-2B and negligible deformation was observed for the cement stabilized soil layer. On the other hand, for Section 4-1B, the stone base, working table and subgrade each contributes 60%, 20% and 20% of the total MDD measured permanent deformation, respectively. This observation further confirms that the cement stabilized soil layer of section 4-2B had a higher strength and greater load carrying capability than the working table in this study.

## COST-BENEFIT ANALYSIS

The aforementioned APT results generally revealed that using a cement stabilized soil subbase layer in a flexible pavement structure (e.g. section 4-2B) can extend the pavement service life by several times when compared to a pavement with only a lime-treated working table layer (e.g. section 4-1B), as illustrated by an ESAL advantage of 6.5:1. However, the decision on whether to implement the design of a cement stabilized subbase layer should be also dependent on its economic aspects or benefits, as the initial construction cost of using a cement stabilized soil subbase layer in a flexible pavement could be higher.

To demonstrate the potential cost benefits of using a cement stabilized soil subbase in lieu of a lime-treated working table layer in a flexible pavement design, a life-cycle cost analysis (LCCA) was performed on a typical Louisiana flexible pavement structure of low volume roads. As shown in Figure 7a, two alternative structure designs are to be used in the LCCA: *Alternative A*-pavement structure contains a 12 in. lime-treated working table layer, and *Alternative B*-pavement structure uses a 12 in. cement stabilized soil subbase layer. According to Louisiana's Alternate Design / Alternative Bid programs, a 30-year analysis period was used in the LCCA analysis. The LCCA analysis assumed that future maintenance requirements would include the following: full reconstruction (including both base and HMA layers) at the end of pavement life for *Alternative A* (pavement with 12 in. lime-treated working table), and a one-time 2 in. milling and 3.5 in. HMA overlay at the end of year 15 for *Alternative B* (pavement with 12 in. cement stabilized soil subbase). Additional assumption included a 5 percent inflation rate and zero salvage value at the end of year 30.

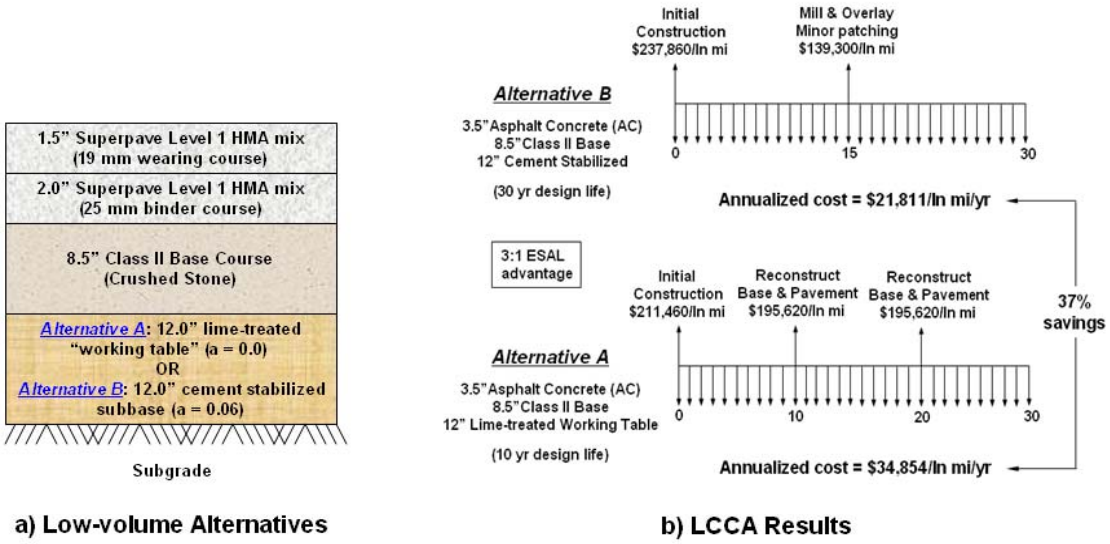



**Figure 7. LCCA on Low Volume Roads**

As shown in Figure 7b, the LCCA analysis for low volume roads reveals that using a 12 in. cement-stabilized soil subbase (*Alternative B*) in lieu of a 12 in. lime-treated working table layer (*Alternative A*) creates a 37 percent annual cost savings during a 30-year analysis period. It should be noted that an ESAL advantages of only 3:1 (as compared to 6.5:1 in the APT results) was used in the LCCA for conservative cost estimation.

## CONCLUSIONS

The following observations and conclusions were drawn from this study:

- Both laboratory and field results confirmed that a stabilized soil layer possesses a higher load-induced structural capacity than a "working table" layer.
- A structural layer coefficient for the cement-stabilized soil used in the APT experiment was predicted to be 0.06; while for lime-treated soil, no structural contribution should be allowed.
- The life cycle cost analysis indicated that using a 12 in. cement stabilized soil subbase in lieu of a 12 in. lime-treated working table layer would creates the annual cost savings of 37 percent for a typical Louisiana low volume pavement.
- It is recommended to stabilize all subgrades susceptible to moisture intrusion rather than optional working table or subgrade treatment. Stabilization of clays with lime and silts with cement will create stronger foundations for pavement structure. For a cement stabilization, a preliminary guideline may be set as: Under a wet subgrade condition (e.g. in situ soil moisture is on the wet side of a compaction curve), cement-soil mixture with a water-cement ratio of 2 at molding will have a minimum 28-day UCS of 150 psi while cement-treated soils with a water-cement ratio of 3 at molding will have a minimum 28-day UCS of 100 psi. For the lime stabilization to achieve a reliable strength, more research is warrant.

## REFERENCES

Gavin, G., Zhang, Z., Wu, Z., Mohammad, M. and King. W. (2010). "Laboratory and Accelerated Loading Evaluation of Subbase Layers in Pavement Performance", Louisiana Transportation Research Center (LTRC) Report No. 468, Baton Rouge, LA.

King, W. and Rasoulian, M. (2004). "Experimental and operational progress with a Benefit/Cost Analysis for Louisiana's Pavement Research Facility". Proceedings of the 2nd International Conference on Accelerated Pavement Testing, Minneapolis.

LADOTD. (2000). Edition. *Louisiana Standard Specification for Roads and Bridges*. Baton Rouge, LA: Louisiana Department of Transportation and Development (LADOTD), 2000 Edition.

McManis, K. (2003). "Identification and Stabilization Methods for Problematic Silt Soils: A Laboratory Evaluation of Modification and Stabilization Additives". LTRC Report No. 371. 2003.

Prusinski, J. R. and Bhattacharja, S. (1999). "Effectiveness of Portland Cement and Lime in Stabilizing Clay Soils." Sugar Land, TX : Portland Cement Association.

Wu, Z., Zhang, Z., King, W., Raghavendra, A. and Martinez, M. (2006). "Instrumentation and Accelerated Testing on Louisiana Flexible Pavements" *Airfield and Highway Pavements - Meeting Today's Challenges with Emerging Technologies.* ASCE Special Publication.

# Measurement and Modeling of Hydraulic Characteristics of Permeable Friction Course (PFC)

J. Brandon Klenzendorf[1], Randall Charbeneau[2], Brad Eck[3], and Michael Barrett[4]

## ABSTRACT

Permeable Friction Course (PFC) is a layer of porous asphalt approximately 50 mm thick placed as an overlay on conventional impervious roadway surfaces. The use of PFC improves driver safety due to reduced splash/spray, improved visibility, and decrease in hydroplaning during rainfall events. In addition, PFC improves the stormwater runoff quality from the roadway and has recently been approved as a new stormwater best management practice (BMP) in the State of Texas. However, over time the pore space in the PFC can become clogged with trapped sediment, suggesting a decrease in driver safety and water quality benefits. Therefore, to ensure proper utilization of these benefits, accurate measurement of the drainage properties of PFC is required. In particular, measurement of the hydraulic conductivity is used to determine whether the drainage benefits of PFC are expected to persist. This paper discusses methods for measuring the porosity of core specimens and hydraulic conductivity under non-Darcy flow behavior characterized by the Forchheimer equation. Laboratory and field experimental results of four years of data (2007 to 2010) for three roadways around Austin, TX are presented. The accurate measurement of in-situ hydraulic conductivity using a new nondestructive field test is essential to maintain proper drainage of the PFC layer and to determine when the pore space has become clogged with sediment. In addition, an overview of numerical modeling results of nonlinear two-dimensional flow through PFC is outlined and results are presented in order to properly determine the hydraulic conductivity. Statistical analysis of porosity data show a decrease in porosity with time suggesting clogging of the pore space. Porosity values range from 12% to 23%. However, there has been no observed statistical decrease in hydraulic conductivity with time, suggesting the driver safety and water quality benefits of the PFC layer have persisted after six years of operation. Hydraulic conductivity values on PFC core specimens range from 0.02 cm/s to nearly 3 cm/s. Furthermore, no maintenance of the PFC layer has been conducted during that time.

[1] Senior Staff Engineer, Geosyntec Consultants, 3600 Bee Caves Rd, Suite 101, Austin, TX 78746; Ph: (512) 451-4003; Fax: (512) 306-8042; email: bklenzendorf@geosyntec.com
[2] Professor, Department of Civil, Architectural and Environmental Engineering, 1 University Station C1700, University of Texas, Austin, TX 78712; Ph: (512) 471-0070; Fax: (512) 471-0072; email: charbeneau@mail.utexas.edu
[3] Post-Doctoral Researcher, Center for Research in Water Resources, PRC #119 (R8000), University of Texas, Austin, TX 78712; Ph: (512) 471-3131; Fax: (512) 471-0072; email: brad.eck@mail.utexas.edu
[4] Research Associate Professor, Center for Research in Water Resources, PRC #119 (R8000), University of Texas, Austin, TX 78712; Ph: (512) 471-0935; Fax: (512) 471-0072; email: mbarrett@mail.utexas.edu

## INTRODUCTION

Permeable Friction Course (PFC) is an innovative roadway pavement system which consists of a conventional impervious roadway surface overlain with a 50 mm thick porous asphalt surface course. The porous surface course allows rainfall and runoff to enter the pore space and be removed from the surface. Water within the PFC layer flows vertically until it hits the underlying impervious conventional roadway surface and then flows by gravity to the roadway shoulder where the runoff enters a drainage ditch or swale. This reduction of surface runoff within the driving lanes improves driver safety by drastically reducing splash/spray behind vehicles, increasing visibility, and reducing the chance of hydroplaning. In addition, PFC has a greater surface roughness when compared to conventional asphalt resulting in improved skid resistance (NCHRP, 2009). The structural integrity of a PFC system is greater when compared to completely porous pavement systems. In completely porous pavement systems, problems arise with respect to structural integrity because the entire pavement system is porous. This concern is avoided with PFC systems since the PFC layer is placed as an overlay on the conventional pavement surface.

In addition to the driver safety benefits of PFC, research has also shown improved water quality benefits (Stotz and Krauth, 1994; Barrett et al., 2006). PFC helps to reduce total suspended solids (TSS) and particulate related pollutants through two mechanisms: reduction of the source of pollutants by minimizing wash-off of pollutants from vehicles and entrapment of pollutants associated with particles by retaining sediment within the pore space of the PFC layer. The latter removal mechanism suggests that the PFC pore space may become clogged over time due to trapped sediment. This clogging is expected to result in a decrease in porosity and drainage capacity of the PFC with time. In general, PFC is expected to have a design life of roughly 10 years. However, if the PFC becomes clogged with sediment at a faster rate, then the PFC may require maintenance or replacement at an interval shorter than the 10 year life span. This paper briefly discusses stormwater monitoring conducted on a PFC layer together with porosity and hydraulic conductivity measurements conducted on PFC core specimens extracted from the roadway. In addition, a field test is described for measurement of in-situ hydraulic conductivity as an assessment method to determine whether the PFC is performing as desired.

## STORMWATER QUALITY BENEFITS

Loop 360 is a PFC roadway in Austin, Texas which has been monitored for stormwater quality since February 2004. At the onset of monitoring, the roadway surface was conventional impervious asphalt. In October 2004, the roadway was overlain with a PFC surface. Monitoring results prior to the installation of the PFC surface showed levels of TSS and other pollutants typically found in highway runoff. After installation of the PFC surface, pollutant levels decreased significantly for TSS and other particulate pollutants. Barrett et al. (2006) and Stanard et al. (2008) provide the specific details of this monitoring system as well as the results, shown in Figure 1

for TSS. Over the past six years, the improved water quality from the PFC surface has persisted.

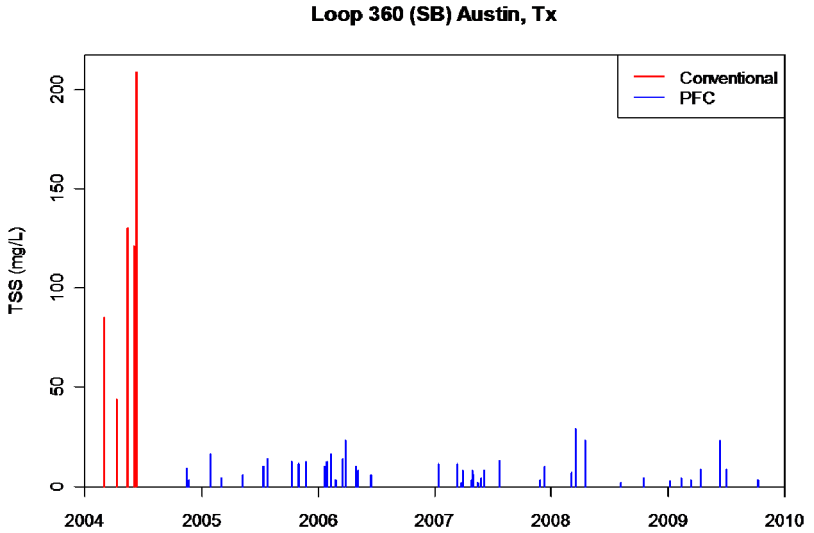



**Figure 1 – Stormwater quality reduction in total suspended solids for continuous sampler at Loop 360**

Although the above monitoring system shows the longevity of water quality improvements that PFC can provide, it does not compare both a PFC surface and conventional asphalt surface for the same rainfall event and traffic conditions. Therefore, paired samples were collected starting in February 2007 at a nearby location on Loop 360. These samples also show a significant decrease in pollutant levels from the PFC runoff, particularly for pollutants associated with sediment as shown in Figure 2 for TSS. Details of this monitoring system are also provided by Stanard et al. (2008).

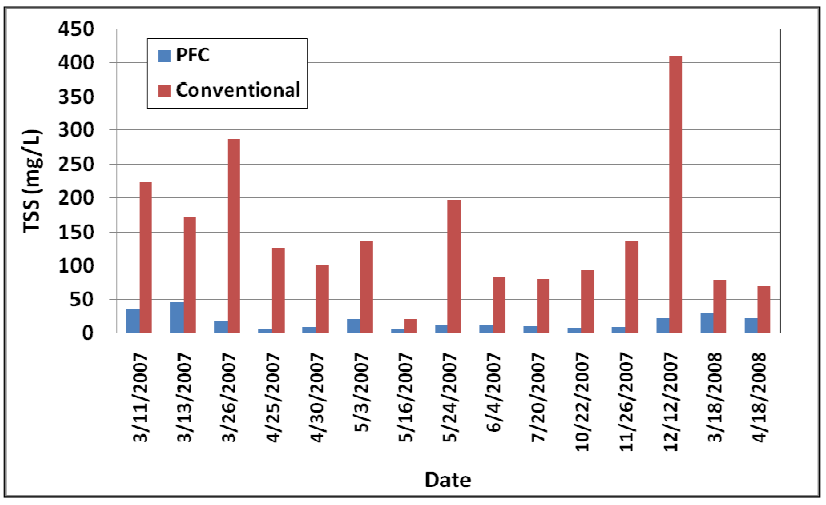



**Figure 2 – Stormwater quality reduction in total suspended solids for paired samples at Loop 360**

The results of these stormwater monitoring efforts have provided the scientific evidence for the Texas Commission on Environmental Quality (TCEQ) to approve the use of PFC as a new stormwater Best Management Practice (BMP) in the State of Texas on uncurbed roadways with a design speed of greater than 50 miles per hour. The roadway surface itself can now be used for water quality and treatment of stormwater runoff in Texas. However, the question now arises as to how to assess the performance of PFC as a BMP. It is expected that the pore space of PFC becomes clogged with sediment over time. Therefore, measurement of the hydraulic properties of a PFC layer is necessary in order to ensure proper drainage and determine whether the benefits of PFC are expected to persist in the future. Measurement of the hydraulic properties is also useful for design purposes to ensure that the PFC layer does not become saturated and result in surface runoff. A numerical model coupling the surface and subsurface flows in the PFC system is described by Eck et al. (2010).

## POROSITY MEASUREMENTS

PFC core specimens can be extracted from the roadway surface by saw-cutting. A typical core specimen is cylindrical in shape and recovers both the PFC layer on the surface together with the underlying conventional asphalt. Core specimens of varying diameter were collected at three roadways around Austin, TX for four years: Loop 360, FM 1431, and RR 620. Klenzendorf et al. (2010) provide details of the coring locations. 15.2 cm diameter cores were extracted in March 2007, 20.3 cm diameter cores were extracted in February 2008 and February 2009, and 15.2 cm diameter cores were extracted in February 2010. Both porosity and hydraulic conductivity measurements can be conducted on the core specimens in order to assess the drainage properties of the PFC overlay.

Porosity measurements were conducted either with a destructive image analysis method or a nondestructive submerged unit weight method. Candaele et al. (2008) provides an overview of both methods. The submerged unit weight method is the preferred porosity measurement method because the core specimen remains intact for subsequent hydraulic conductivity measurements. Figure 3 shows the average porosity with plus/minus one standard deviation for each of the three roadways over the past four years. In general, only three core specimens were extracted from each roadway in any given year. However, for Loop 360, six core specimens were extracted each year: three from the travel lane and three from the shoulder.

Nonparametric statistical tests were conducted on the porosity data which show a statistical decrease in porosity with time for each of the three roadways (Klenzendorf et al., 2010). This suggests that the pore space in the PFC layer is becoming clogged with sediment and suggests a possible decrease in drainage capacity. In order to determine whether the drainage capacity of the PFC is decreasing, hydraulic conductivity measurements must be conducted.

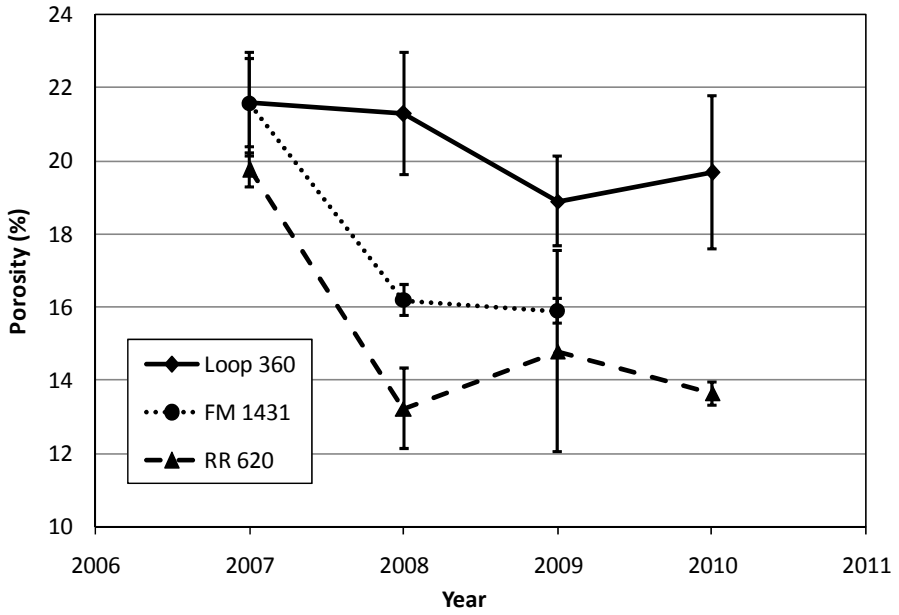



**Figure 3 – Average laboratory porosity plus/minus one standard deviation of PFC core specimens for three roadways over four years**

## HYDRAULIC CONDUCTIVITY MEASUREMENTS

Hydraulic conductivity measurements can be conducted in the laboratory on PFC core specimens using a series of constant head permeability tests. The core specimen is submerged in water and compressed between two metal plates. Rubber gaskets are inserted between the core specimen and the metal plates so that during the compression, the rubber will enter the surface voids of the PFC surface and create a no flow boundary. Water is pumped at a constant rate though a standpipe centered on the core specimen. The water enters the specimen vertically at the top surface and turns to exit the core in the radial direction along the radial periphery of the core specimen. Figure 4 shows a schematic of the test setup and details are provided by Klenzendorf et al. (2010).

During a constant head test, the volumetric flow rate $Q$ produced by the pump and the change in head through the core specimen (head in the standpipe) $h_s$ are measured. The flow rate is measured with laboratory glassware and a stopwatch and the standpipe head is measured with a bubble flow meter. In addition, the core geometry consists of the standpipe radius $R_s$, core radius $R_c$, and core thickness $b_c$. A typical graph of flow rate versus standpipe head is provided in Figure 5.

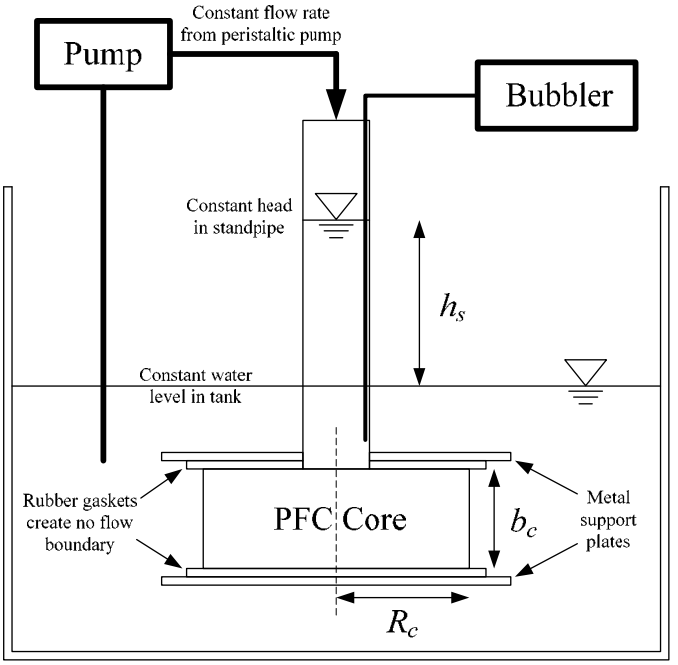



**Figure 4 – Schematic of laboratory setup for constant head hydraulic conductivity measurements under two-dimensional flow pattern**

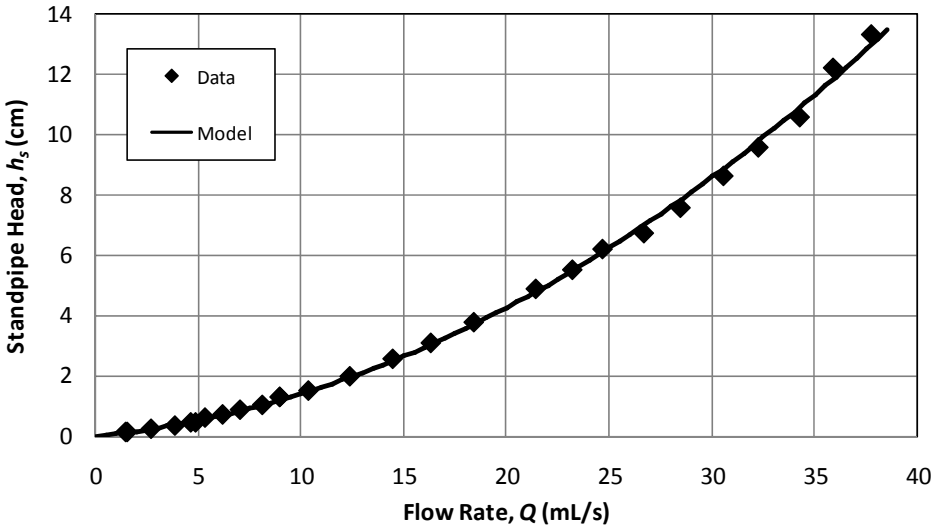



**Figure 5 – Typical constant head results showing nonlinear flow relationship**

Figure 5 shows a nonlinear flow relationship which cannot be modeled with the typical Darcy law for porous media. Therefore, a nonlinear flow equation must be used instead. The Forchheimer equation has a theoretical development for nonlinear flow effects due to large velocity. The original Forchheimer equation is provided as:

$$I = aq + bq^2 \quad (1)$$

Equation (1) is a nonlinear relationship between the hydraulic gradient $I$ and the specific discharge $q$. The two original Forchheimer coefficients are $a$ and $b$. $a = 1/K$ where $K$ is the hydraulic conductivity. Due to the two-dimensional flow pattern

observed during testing, the hydraulic gradient and specific discharge cannot be measured directly. Therefore, a modified Forchheimer equation must be used for the global conditions of the core specimen which is given as:

$$h_s = \alpha Q + \beta Q^2 \quad (2)$$

Equation (2) relates the standpipe head to the flow rate which can be measured during testing. The two modified Forchheimer coefficients are $\alpha$ and $\beta$. These coefficients can be determined from experimental data but give no indication of the original Forchheimer coefficients. Numerical modeling can be used to estimate the original Forchheimer coefficients from the known modified Forchheimer coefficients in order to determine the true hydraulic conductivity.

Hydraulic conductivity tests can also be conducted in the field. This allows for measurement of the in-situ hydraulic conductivity as a way to assess the performance of the PFC layer. The test apparatus developed specifically for this research is shown in Figure 6. A falling head test is conducted in the field to determine the modified Forchheimer coefficients. A no flow boundary is created under the metal base plate of the test apparatus so that the same boundary conditions created in the laboratory are imposed in the field. This allows for use of the modified Forchheimer Equation (2). The falling head equation for nonlinear two-dimensional flow is given as:

$$-\pi R_s^{\,2} \frac{dh_s}{dt} = Q = \frac{\alpha}{2\beta}\left[\sqrt{1+\frac{4\beta h_s}{\alpha^2}} - 1\right] \quad (3)$$

Equation (3) can be integrated from an initial head to some arbitrary head in order to measure the two modified Forchheimer coefficients from three time-depth pairs. The three time-depth measurements can be made with a stopwatch which has a split function. Therefore, no advanced instrumentation is needed in order to conduct the field test.

The two modified Forchheimer coefficients, $\alpha$ and $\beta$, can be determined either through constant head tests in the lab or a falling head test in the field. However, this gives no direct indication of the original Forchheimer coefficients, $a$ and $b$, or the hydraulic conductivity. In order to determine the original Forchheimer coefficients, numerical modeling of two-dimensional nonlinear flow must be conducted.

**Figure 6 – Field test apparatus using falling head permeability test**

## NUMERICAL MODELING OF NONLINEAR FLOW

Due to the two-dimensional nonlinear flow patterns created during both laboratory and field hydraulic conductivity testing, a representative one-dimensional hydraulic conductivity cannot be directly measured. Therefore, use of a modified Forchheimer equation is used to describe the hydraulic characteristics of the PFC during testing. Numerical modeling is needed to relate the modified Forchheimer coefficients determined from experiments to the original Forchheimer coefficients which ultimately relate to the true hydraulic conductivity.

A finite difference numerical model for two-dimensional isotropic nonlinear flow is described by Klenzendorf et al. (2010). The inputs to the model are the original Forchheimer coefficients ($a$ and $b$), standpipe head ($h_s$), and core geometry ($R_s$, $R_c$, and $b_c$). The model solves the steady state continuity equation in two-dimensional cylindrical coordinates using the original Forchheimer equation to calculate the outflow rate from the core specimen. Multiple values of $h_s$ are simulated and the corresponding $Q$ is calculated. The two modified Forchheimer coefficients ($\alpha$ and $\beta$) are determined using regression equations from the resulting $h_s$ vs. $Q$ relationship. Therefore, the model has inputs of the two original Forchheimer coefficients and the core geometry, with outputs of the two modified Forchheimer coefficients.

In order to estimate the original Forchheimer coefficients from the known modified Forchheimer coefficients, regression equations were developed based on 64 different combinations of the core geometry. Details are provided in Klenzendorf et al. (2010) with the following regression equations determined using a power law relationship:

$$a = \frac{1}{K} \approx 5.8 R_s \left( \frac{b_c}{R_c} \right)^{0.33} \alpha \tag{4}$$

$$b \approx 18.6 R_s^{\,2.5} \left( R_c b_c \right)^{0.25} \beta \tag{5}$$

The hydraulic conductivity can be estimated using Equation (4) based on the measured linear modified Forchheimer coefficient $\alpha$ and the core geometry. Similarly, the nonlinear coefficient $b$ can be determined using Equation (5).

The numerical model was also applied specifically to the field test apparatus since the $R_s$ and $R_c$ remain constant for all tests ($R_s$ = 5.08 cm and $R_c$ = 22.86 cm). The following regression equations were developed which require the units of measurement to be in cm:

$$a = \frac{1}{K} \approx 5{b_c}^{0.75}\alpha \quad (6)$$

$$b \approx 482{b_c}^{1.25}\beta \quad (7)$$

Equations (6) and (7) can be used to estimate the original Forchheimer coefficients from the measured modified Forchheimer coefficients only for the specific field test apparatus developed for this research.

## HYDRAULIC CONDUCTIVITY RESULTS

The hydraulic conductivity of PFC core specimens in the laboratory can be estimated using the modified Forchheimer equation for two-dimensional nonlinear flow and the regression Equation (4) developed from a finite difference numerical model. Figure 7 shows the resulting average hydraulic conductivity plus/minus one standard deviation for the PFC core specimens.

Nonparametric statistical tests were conducted on the hydraulic conductivity data which show a constant hydraulic conductivity over time for each of the three roadways (Klenzendorf et al., 2010). Therefore, although the PFC porosity is decreasing over time, the trapped sediment has not significantly decreased the drainage capacity of the PFC. This is due in part to the large variability observed in the hydraulic conductivity measurements and that sediment may be trapped in closed pore spaces not connected to the main flow channels. Because there has been no observed statistical decrease in drainage capacity, it is expected that the PFC layer is continuing to provide driver safety and water quality benefits. The improved water quality results obtained from stormwater monitoring supports this claim that the drainage capacity has not yet begun to degrade. No maintenance on the PFC layer has been conducted to date.

Several field tests were conducted at each roadway site to estimate the in-situ hydraulic conductivity using Equation (6) with the following results: Loop 360 $K$ = 3.0 cm/s, FM 1431 $K$ = 0.6 cm/s, and RR 620 $K$ = 1.5 cm/s. There are slight differences in the field results when compared to the lab results. This is due in part to the disturbance of the core specimen during the extraction process as well as the smaller sample size for in-situ measurements. The following average annual daily traffic (AADT) loads were measured in 2005 at each site: Loop 360 AADT = 48,000; FM 1431 AADT = 18,200; and RR 620 AADT = 32,000. The field test is considered the most desirable for assessing the performance of the PFC layer in order to

determine whether driver safety and water quality benefits are expected to persist in the future due to the simplicity and quickness of the test as well as the elimination of the coring process which may disturb the specimen.

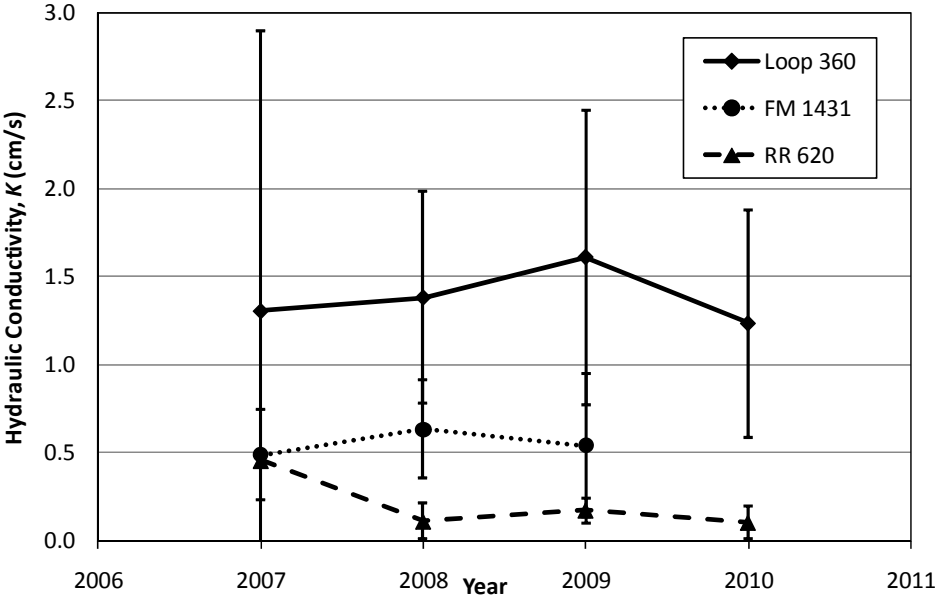



**Figure 7 – Average laboratory hydraulic conductivity plus/minus one standard deviation of PFC core specimens for three roadways over four years**

## DISCUSSIONS AND CONCLUSIONS

The use of Permeable Friction Course (PFC) on conventional impervious roadway surfaces provides both driver safety and water quality benefits. PFC has been approved as a best management practice (BMP) in the State of Texas. These benefits occur primarily due to the reduction of surface runoff during rainfall events. In order to properly assess whether these benefits are expected to persist in the future, measurements of the hydraulic characteristics of the PFC layer are needed.

Both porosity and hydraulic conductivity measurements were conducted in the laboratory on PFC core specimens. The results show a statistical decrease in porosity over time suggesting the pore space is becoming clogged with sediment. However, the hydraulic conductivity has remained statistically constant over time suggesting the drainage benefits of the PFC are expected to continue. Porosity measurements range from 12% to 23%, whereas hydraulic conductivity measurements range from 0.02 cm/s to nearly 3 cm/s.

A field test using a new test apparatus developed for this research study is used to determine the in-situ hydraulic conductivity. Under all hydraulic conductivity measurements, nonlinear flow is observed for the two-dimensional flow pattern imposed during testing. Therefore, in order to relate the measurable coefficients

obtained during testing to the theoretical nonlinear coefficients, numerical modeling must be conducted. The numerical model solves the two-dimensional nonlinear continuity equation. Multiple simulations were conducted to obtain regression equations which can be used to estimate the hydraulic conductivity based on the measureable coefficients obtained in the field. This provides for an accurate determination of the hydraulic conductivity.

## ACKNOWLEDGEMENTS

This research project was funded by the Texas Department of Transportation (TxDOT) project number 0-5220. This paper has not been subject to TxDOT's required peer and policy review, and therefore, does not necessarily reflect the views of TxDOT. Research was conducted at the Center for Research in Water Resources at The University of Texas at Austin.

## REFERENCES

Barrett, M.E., P. Kearfott, and J.F. Malina (2006): "Stormwater Quality Benefits of a Porous Friction Course and Its Effects on Pollutant Removal by Roadside Shoulders," *Water Environmental Research*, Vol. 78, No. 11, pg. 2177-2185.

Candaele, R.M., M.E. Barrett, and R.J. Charbeneau (2008): *Porous Friction Course: A Laboratory Evaluation of Hydraulic Properties*, CRWR Online Report 08-06, Center for Research in Water Resources, The University of Texas at Austin, <http://www.crwr.utexas.edu/online.shtml>.

Eck, B.J., R.J. Charbeneau, and M.E. Barrett (2010): *Drainage Hydraulics of Porous Pavement: Coupling Surface and Subsurface Flow*, CRWR Online Report 10-02, Center for Research in Water Resources, The University of Texas at Austin, <http://www.crwr.utexas.edu/online.shtml>.

Klenzendorf, J.B., R.J. Charbeneau, and M.E. Barrett (2010): *Hydraulic Conductivity Measurement of Permeable Friction Course (PFC) Experiencing Two-Dimensional Nonlinear Flow Effects*, CRWR Online Report 10-01, Center for Research in Water Resources, The University of Texas at Austin, <http://www.crwr.utexas.edu/online.shtml>.

NCHRP (2009): *Construction and Maintenance Practices for Permeable Friction Courses*, National Cooperative Highway Research Program, Report 640, Transportation Research Board of the National Academies, Washington, D.C.

Stanard, C.E., M.E. Barrett, and R.J. Charbeneau (2008): *Stormwater Quality Benefits of a Permeable Friction Course*, CRWR Online Report 08-03, Center for Research in Water Resources, The University of Texas at Austin, <http://www.crwr.utexas.edu/online.shtml>.

Stotz, G. and K. Krauth (1994): "The Pollution of Effluents from Pervious Pavements of an Experimental Highway Section: First Results," *The Science of the Total Environment*, Vol. 146-147, pg. 465-470.

# The Influence of CTE on Jointed Concrete Pavement Performance Based on M-E Rigid Pavement Analysis

Raphael Kampmann[1], W. Virgil Ping[2], and Michelle Roddenberry[3]

[1]Ph.D.-Candidate, Department of Civil and Environmental Engineering, FAMU-FSU College of Engineering, Florida State University, 2525 Pottsdamer Street, Tallahassee, FL 32310, USA, PH +1 (850) 410-6146; FAX +1 (850) 410-6142; kampmann@eng.fsu.edu
[2]Professor, Department of Civil and Environmental Engineering, FAMU-FSU College of Engineering, Florida State University, 2525 Pottsdamer Street, Tallahassee, FL 32310, USA, PH +1 (850) 410-6129; FAX +1 (850) 410-6142; ping@eng.fsu.edu,
[3]Assistant Professor, Department of Civil and Environmental Engineering, FAMU-FSU College of Engineering, Florida State University, 2525 Pottsdamer Street, Tallahassee, FL 32310, USA, PH +1 (850) 410-6125; FAX +1 (850) 410-6142; roddemi@eng.fsu.edu

## ABSTRACT

Three typical Florida Portland cement concrete (PCC) pavement mix designs have been evaluated to capture the input requirements for the new Mechanistic-Empirical Rigid Pavement Design Guide (M-E RPDG). The essential engineering properties were experimentally determined for all required maturity levels, with an emphasis on the coefficient of thermal expansion (CTE) according to AASHTO TP-60 test protocol.

Traffic loads, environmental conditions, structural parameters, and analysis/acceptance criteria were defined on account of local requirements. Due to the introduced hierarchy system, nine diverse JPCP models were established and iteratively idealized for PCC top layer thickness. The results demonstrated that cracking is the critical performance criterion for Florida JPCP. Moreover, top-down fatigue damage was isolated to be the controlling failure mechanism. Based on the nine Florida pavement models, a sensitivity matrix was developed to account for CTE as a control variable. Despite wide-ranging PCC, CTE, and thickness properties, clear resemblances were exposed for all scenarios under evaluation. With regard to CTE variation, the JPCP analysis according to the AASHTO M-E RPDG is minimally sensitive to faulting, sensitive to bottom-up damage, and extremely sensitive to top-down damage, cracking, and smoothness/roughness.

## INTRODUCTION

The American Association of State Highway and Transportation Officials (AASHTO) recently adopted the newly developed Mechanistic-Empirical Pavement Design Guide (M-E PDG) as the interim design guide for new and rehabilitated pavements (AASHTO 2008). To properly implement the AASHTO M-E PDG for the design of rigid pavement in Florida, essential engineering properties of local concrete paving mixtures will have to be further explored and evaluated (Coree et al. 2005; Guclu and Ceylan 2005; Hall and Beam 2004; Nantung et al. 2005; Salman et al. 2007). Specifically, the coefficient of thermal expansion (CTE) is an essential engineering property for concrete, affecting the performance of rigid pavement significantly (Guclu and Ceylan 2005; Hall and Beam 2004; Salman et al. 2007; Hossain et al. 2006; Mallela et al. 2005; Won 2005) when using the Mechanistic-Empirical Rigid Pavement Design Guide (M-E RPDG).

The primary objectives of this research were to evaluate the thermal engineering properties of common Florida concrete paving mixtures and to study the CTE sensitivity when using the AASHTO M-E RPDG. The research targeted typical Florida Portland cement concrete (PCC) mixtures and particularly their performance in representative pavement structures under local climate conditions and traffic loads. The CTE sensitivity analysis was aimed at studying the effect of varying PCC thermal expansion properties and their significance for the predicted distresses and smoothness/roughness models.

## LABORATORY EXPERIMENTAL PROGRAM

Three typical Florida concrete paving mixtures were thoroughly evaluated in the laboratory. Two concrete mixtures contained dissimilar quantities of limestone (MIX-01 using 67-Stone and MIX-02 using 57-Stone) mined from Quarry Mine #38-268 with a fineness modulus of 2.67 and 2.27 respectively and a specific gravity of 2.65, while one mixture contained granite aggregates (MIX-03) that were mined from Plant 022-Barin in Georgia. The fine aggregates were consistent in all three mixes and made of silica sand mined from Quarry Mine #47-314 with a specific gravity of 2.66 and a fineness modulus of 2.35. The proportions of concrete ingredients of the evaluated paving mixtures are summarized in TABLE 1.

**TABLE 1 Mix design composition**

| | Unit | MIX-01 | MIX-02 | MIX-03 |
|---|---|---|---|---|
| Cement | lb | 511 | 415 | 470 |
| Fly Ash | lb | 132 | 105 | - |
| Coarse Aggregates | lb | 1750 | 1900 | 1921 |
| Fine Aggregates | lb | 1191 | 1278 | 1235 |
| Air Entraining Admixture | oz | 1.5 | 1.0 | 2.5 |
| Admixture | oz | 35.4 | 45 | 18 |
| Water | lb | 279.1 | 258 | 267 |
| Water/Cementitious | - | 0.43 | 0.50 | 0.57 |

Considering the input parameters required for the mechanistic-empirical analysis procedure according to the M-E RPDG, the engineering properties of the three concrete mixtures were measured in the laboratory for all essential maturity levels (7, 14, 28, and 90 days). To account for all hierarchy levels in the design guide, the physical characteristics under empirical evaluation included compressive strength, flexural strength, splitting tensile strength, Young's modulus, Poisson's ratio, unit weight, air content, cement type, cement content, water-to-cement ratio, and coefficient of thermal expansion. Each property was experimentally measured according to their designated ASTM and AASHTO test methods before it was averaged for a series of seven test specimens per mix design and maturity level. Overall, more than 100 concrete specimens per mix design were produced (one batch per mix design) and tested for the required M-E RPDG input values.

The three concrete mixes demonstrated distinctive results for all engineering properties (Kampmann 2008), particularly with higher strength characteristics for MIX-02 (see TABLE 2). It should be emphasized that the CTE test results reported herein were based on test protocol AASHTO TP-60 with a recommended CTE value of 17.3 μin/in/ºF for the calibration sample. Since then a corrected test procedure (AASHTO T 336-09) has been introduced to adjust the CTE measurement (Tanesi et al. 2010; Crawford et al. 2010).

## MODELING OF RIGID PAVEMENT STRUCTURE

The measured laboratory test results were statistically reduced in order to obtain PCC design input parameters suitable for pavement analysis according to the AASHTO M-E PDG Software (compatible with AASHTO TP-60). To evaluate the three concrete mix designs for each hierarchy level, a total of nine different top concrete layer designs were generated; the resulting input parameters for the M-E RPDG analysis are summarized in TABLE 2.

**TABLE 2 PCC properties for pavement design models**

| | Unit | MIX-01 | MIX-02 | MIX-03 |
|---|---|---|---|---|
| **General and Thermal** | | | | |
| Unit weight | pcf | 143.1 | 146.6 | 144.2 |
| Thermal conductivity[2] | BTU/hr-ft-°F | 1.25 | | |
| Heat capacity [2] | BTU/lb-°F | 0.28 | | |
| **Mix** | | | | |
| Cement Type | - | Type I | | |
| Cementitious material content | lb/yd$^3$ | 643 | 520 | 470 |
| Water/cement ratio | - | 0.546 | 0.622 | 0.568 |
| Aggregate type | - | Limestone | | Granite |
| Zero-stress temperature[2] | °F | 110 | 103 | 100 |
| Ultimate shrinkage at 40% R.H.[2] | µin/in | 903 | 780 | 680 |
| Reversible shrinkage | % | 50 | | |
| Time to develop 50% shrinkage | days | 35 | | |
| Curing method | - | Curing compound | | |
| **Level 1** | | | | |
| Coefficient of thermal expansion | µin/in/°F | 6.8 | 6.8 | 6.0 |
| Poisson's ratio | - | 0.23 | 0.27 | 0.21 |
| E-modulus (7-day) | psi | 4370000 | 4530000 | 3290000 |
| E-modulus (14-day) | psi | 4590000[1] | 4590000[1] | 3470000 |
| E-modulus (28-day) | psi | 4420000[1] | 4360000[1] | 3680000 |
| E-modulus (90-day) | psi | 4180000[1] | 4905000 | 4020000 |
| 20 year/28 day ratio[2] | - | 1.2 | | |
| Modulus of rupture (7-day) | psi | 714 | 777 | 545 |
| Modulus of rupture (14-day) | psi | 745[3] | 825[3] | 603.5[3] |
| Modulus of rupture (28-day) | psi | 776 | 873 | 662 |
| Modulus of rupture (90-day) | psi | 844 | 831[1] | 707 |
| 20 year/28 day ratio[2] | - | 1.2 | | |
| **Level 2** | | | | |
| Coefficient of thermal expansion[4] | µin/in/°F | 6.4 | 6.7 | 6.9 |
| Poisson's ratio | - | 0.23 | 0.27 | 0.21 |
| Compressive Strength (7-day) | psi | 4832 | 6183 | 3489 |
| Compressive Strength (14-day) | psi | 6067 | 7442 | 4225 |
| Compressive Strength (28-day) | psi | 6908 | 8256 | 4883 |
| Compressive Strength (90-day) | psi | 7856 | 9188 | 5794 |
| 20 year/28 day ratio[2] | - | 1.2 | | |
| **Level 3** | | | | |
| Coefficient of thermal expansion | µin/in/°F | 7.0 | | |
| Poisson's ratio [5)] | - | 0.25 | | |
| Modulus of rupture (28-day) | psi | 776 | 873 | 662 |
| E-Modulus (28-day) | psi | 4420000[1] | 4360000[1] | 3680000 |

[1)] Measured values not increasing as expected by theory
[2)] Non empirical value (suggested by ME-PDG)
[3)] Non empirical value (linearly interpolated)
[4)] Calculated according to Rule of Mixture

**JPCP Modeling.** A typical profile of rigid pavement structure in Florida is illustrated in FIGURE 1. The rigid pavement structure (referred to as model) consists of a top layer concrete slab (made of the evaluated concrete and variable in thickness) and underlying layers of base, compacted subbase, and embankment (roadbed material).

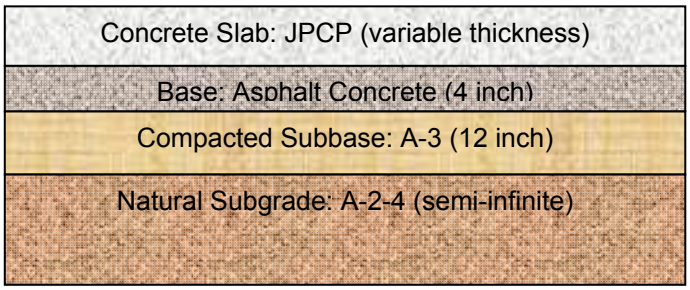



**FIGURE 1 Typical profile of rigid pavement structure in Florida**

**JPCP Design Features.** In the models, the concrete slab aspect ratio was fixed at 0.86 resulting from an assumed joint spacing of 15 feet and a slab width of 13 feet (widened slab). The transverse joints were liquid sealed and doweled using 1.5-inch diameter dowel bars spaced at 12 inches. Tied PCC shoulders were assumed with 40 percent of long-term load transfer efficiency (separately constructed). The base type was asphalt treated, causing an intermediate erodibility index of 3 (erosion resistant). Friction contact at the PCC-Base interface was assumed to be inapplicable (with zero friction contact). Interstate 4 ranging from milepost 61.747 to 68.102 (near Disney World) was considered appropriate for this evaluation. All rigid pavement structure models were subject to identical local traffic and the same environmental conditions (FDOT 2005).

**Analytical Parameters.** The M-E JPCP analysis evaluates transverse cracking, transverse joint faulting, and pavement roughness/smoothness (International Roughness Index, IRI). On account of local requirements (FDOT 2009) initial roughness/smoothness (International Roughness Index, IRI) for the project was predefined at 58 inches per mile and terminal IRI was limited to 160 inches per mile. Maximum transverse cracking was assumed to be adequate at less than 10% slabs cracked while mean joint faulting had to maintain values below 0.12 inches. The reliability factor for all three performance criteria was fixed at 90%.

## RESULTS OF JPCP THICKNESS ANALYSIS

The requirement for the thickness analysis was a minimum PCC layer thickness in conformity with the Florida distress and smoothness criteria (i.e., analytical parameters, including reliability). The ideal design thickness was determined through the M-E JPCP analysis that just satisfied all performance criteria within the specified design life. The results of the thickness analysis are summarized in TABLE 3 for the nine JPCP models.

**TABLE 3 Analytical results of JPCP layer thickness**

| | Unit | MIX-01 | MIX-02 | MIX-03 |
|---|---|---|---|---|
| Level-1 thickness (h) | In | 8.75 | 8.25 | 9.00 |
| Level-2 thickness (h) | In | 10.25 | 10.75 | 13.50 |
| Level-3 thickness (h) | In | 10.50 | 7.50 | 12.00 |

Other than MIX-02 and MIX-03, MIX-01 followed the expected pattern, increasing its thickness with raised hierarchy level. For MIX-02, the 28-day flexural strength was unexpectedly high and even higher than the 90-day flexural strength. However, the 28-day flexural strength value (and the 28-day Young's modulus) was the only strength input for level three; whereas, the level two relied on compressive

strength for all maturity levels (7, 14, 28, and 90-day). Thus, due to the 28-day modulus of rupture input, the generally expected pattern in layer thickness was not followed for MIX-02. For MIX-03, the layer thickness at level three was slightly lower than at level two due to the higher modulus of rupture.

The thickness analysis proved cracking to be the most critical performance criterion for the selected Florida JPCP models as all mix designs attained the limit of 10 percent slabs-crack far before any other performance criteria became problematic in accordance with the chosen (Florida) analysis parameters. However, the results were highly dependent on the selected JPCP design features (Kampmann 2008). The widened slabs with tied PCC shoulders and doweled transverse joints supported by an asphalt-treated base maintained faulting values within moderate ranges, below the designated Florida limit of 0.12 inches. Less faulting results in better driving quality and smoother pavement surface, causing almost negligible IRI augmentation throughout the design life. Furthermore, because of its failure mechanism, bottom-up cracking benefits from widened slabs and tied PCC shoulders as well. In essence, top-down fatigue damage was isolated to be the controlling failure mechanism in this research; the percentage of cracked slabs was directly proportional to top-down damage for all scenarios under evaluation, turning the predicted cracking into the most critical pavement performance in JPCP analysis.

## CTE SENSITIVITY ANALYSIS

To properly evaluate the sensitivity of pavement performance and the mechanistic-empirical analysis approach in response to CTE variations, the CTE was selected as the only control variable while all other parameters were held unchanged. Although it is unlikely for one single concrete property to change its magnitude without affecting any other characteristics, it was decided to vary CTE only as it shows the susceptibility of the M-E RPDG approach in response to CTE. The CTE values were ascended/descended by 0.1 μin/in/°F increments within a ± 10 % array causing an analysis matrix of 15 (CTE values) by 6 (5 analysis criteria) per mix design and hierarchy level. The sub-matrices were populated through their CTE array and the predicted performance criteria at the end of the design life. A typical CTE sensitivity sub-matrix for MIX-02 at level one is presented in TABLE 4 for illustration.

**TABLE 4 CTE sensitivity sub-matrix for MIX-02 at level one**

| MIX-02 | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Level | H | CTE | | IRI | | Faulting | | LTE | Fatigue Damage | | Cracking | |
| | In | μin/in/F | % | in/mile | | In | | % | (fraction) | | % slabs cracked | |
| | | | | | at Spec. Reliab. | | at Spec. Reliab. | | Bottom-up | Top-down | | at Spec. Reliability |
| 1 | 8.25 | 6.1 | 89.8 | 58.6 | 81.3 | 0.000 | 0.018 | 95.0 | 0.0077 | 0.0587 | 0.4 | 5.9 |
| | | 6.2 | 91.2 | 58.7 | 81.5 | 0.000 | 0.018 | 95.0 | 0.0081 | 0.0685 | 0.5 | 6.2 |
| | | 6.3 | 92.7 | 58.9 | 81.7 | 0.000 | 0.018 | 95.0 | 0.0086 | 0.0796 | 0.7 | 6.6 |
| | | 6.4 | 94.2 | 59.1 | 82.0 | 0.000 | 0.018 | 95.0 | 0.0090 | 0.0921 | 0.9 | 7.1 |
| | | 6.5 | 95.6 | 59.3 | 82.4 | 0.000 | 0.018 | 95.0 | 0.0095 | 0.1061 | 1.2 | 7.6 |
| | | 6.6 | 97.1 | 59.6 | 82.9 | 0.000 | 0.018 | 95.0 | 0.0100 | 0.1216 | 1.5 | 8.3 |
| | | 6.7 | 98.5 | 60.0 | 83.5 | 0.000 | 0.018 | 95.0 | 0.0106 | 0.1389 | 2.0 | 9.0 |
| | | **6.8**[1] | **100.0** | **60.4** | **84.3** | **0.000** | **0.018** | **95.0** | **0.0110** | **0.1587** | **2.6** | **9.9** |
| | | 6.9 | 101.5 | 61.0 | 85.3 | 0.000 | 0.018 | 95.0 | 0.0116 | 0.1808 | 3.3 | 11.0 |
| | | 7.0 | 102.9 | 61.8 | 86.5 | 0.000 | 0.018 | 95.0 | 0.0122 | 0.2055 | 4.2 | 12.3 |
| | | 7.1 | 104.4 | 62.7 | 88.0 | 0.000 | 0.018 | 95.0 | 0.0128 | 0.2327 | 5.3 | 13.8 |
| | | 7.2 | 105.8 | 63.8 | 89.7 | 0.000 | 0.018 | 95.0 | 0.0135 | 0.2633 | 6.7 | 15.7 |
| | | 7.3 | 107.3 | 65.3 | 92.0 | 0.000 | 0.018 | 95.0 | 0.0142 | 0.2994 | 8.4 | 17.9 |
| | | 7.4 | 108.8 | 66.7 | 94.2 | 0.000 | 0.018 | 95.0 | 0.0149 | 0.3330 | 10.2 | 20.1 |
| | | 7.5 | 110.2 | 68.5 | 97.1 | 0.000 | 0.018 | 95.0 | 0.0156 | 0.3730 | 12.4 | 22.9 |

[1] CTE value used for thickness analysis (100% value)

The results of the CTE sensitivity analysis are presented in FIGURES 2 through 6 to illustrate the CTE effect on every performance criterion separately.

**Faulting Response.** Pavement faulting is a direct result of load transfer efficiency (LTE), and generally shows a similar behavior in JPCP. Faulting is shown in FIGURE 2 to be insensitive to CTE magnification for thin concrete slabs (less than 9 inches); all mix designs under hierarchy level one and MIX-02 at level three considerations maintain constant values throughout the CTE array.

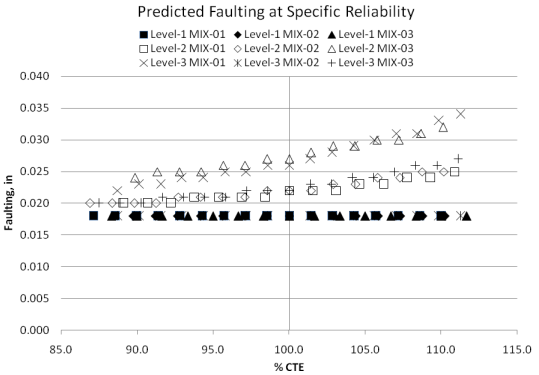



**FIGURE 2 Predicted faulting at 90% reliability vs. percentage of CTE**

For the thicker slabs, the effect is generally insignificant due to the chosen JPCP design features, and even 10 percent amplification of CTE could not significantly influence the predicted faulting (at specific reliability); the results fell consistently below the Florida analysis limit of 0.12 inches.

**Bottom-Up Damage Response.** The effect of CTE variation on bottom-up damage is shown in FIGURE 3. It is illustrated that bottom-up damage is not problematic for thick (more than 9 inch) PCC layers since no effect is observed for any change in CTE. However, the graph shows a linear increasing response to CTE if bottom-up cracking is initially present in thinner PCC layers. The slopes of the three mix designs under level one consideration are comparable and they show parallel bottom-up damage reactions. MIX-02 under hierarchy level three features the thinnest PCC slab and the steepest. The increased slope may be a result of the very thin PCC layer but might be caused by the different hierarchy level as well.

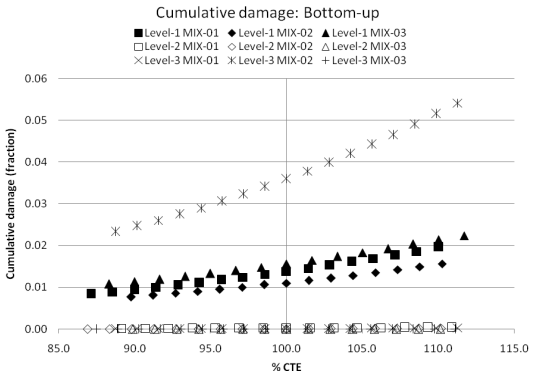



**FIGURE 3 Cumulative bottom-up damage versus percentage of CTE**

**Tow-Down Damage Response.** During thickness analysis, cumulative top-down damage was more prominent than bottom-up damage. Consequently, the CTE sensitivity is initially more pronounced at its 100 percent CTE value as shown in FIGURE 4.

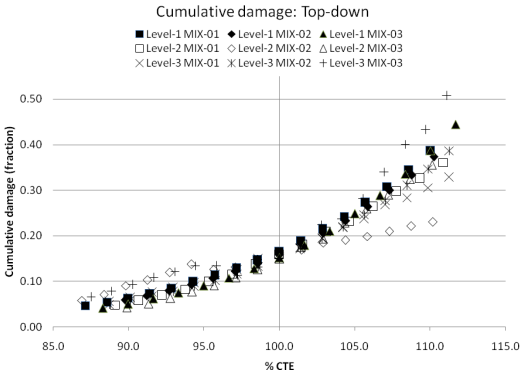



**FIGURE 4 Cumulative top-down damage versus percentage of CTE**

It may be seen that all pavement models follow a similar response for top-down damage. Regardless of layer thickness or any other factors, the cumulative damage pattern is consistently described by a quadratic function. Top-down damage shows a similar response for all scenarios and is very sensitive to CTE variation.

**Cracking Response.** Because predicted percentage of slabs cracked (at specific reliability) was the determining factor for the layer thickness analysis, it is the most important consideration during CTE sensitivity analysis as higher responses would impact the mechanistic-empirical analysis significantly. FIGURE 5 is presented to demonstrate the sensitivity of cracking performance due to CTE variation.

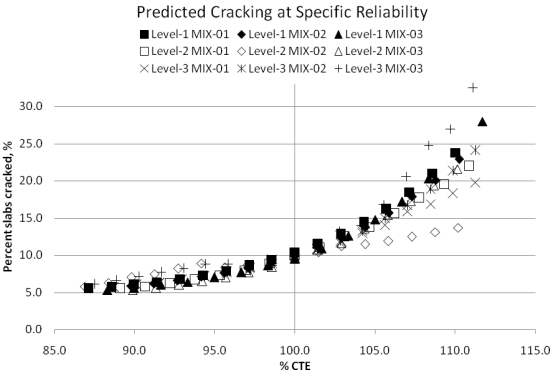



**FIGURE 5 Predicted cracking at 90% reliability versus percentage of CTE**

Comparing FIGURE 4 to FIGURE 5, predicted cracking is a direct result of top-down damage but with a more enhanced quadratic coefficient causing a rapid response for pavement structures among higher CTE values. Top-down damage is one of the most dominating factors for cracking computations according to the new M-E RPDG concept. Hence, the two graphs are congruent to one another. It is shown that a 10 percent CTE increase may cause an additional 20 percent slabs cracked (at specific reliability). In general, predicted transverse cracking (percent slabs cracked) shows a high sensitivity to CTE variation. An overestimation of CTE or a concrete mixture

with high CTE properties will cause augmented pavement cracking for the evaluated JPCP structure.

**Smoothness/Roughness Response.** Predicted smoothness/roughness is the most important performance to the end-user, and therefore its CTE sensitivity is an essential behavior in question. FIGURE 6 illustrates the predicted IRI at specific reliability in response to the CTE array.

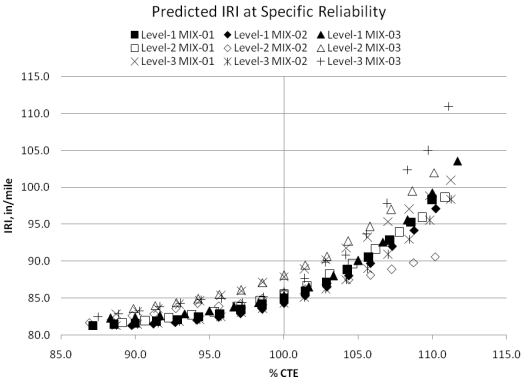



**FIGURE 6 IRI at specific reliability versus percentage of CTE**

The smoothness/roughness responses are shown in FIGURE 6 to mirror the cracking behavior which in turn is a consequence of top-down damage. The similar response between the international roughness index (IRI) and cracking is assumed to be caused by the minor faulting response based on the chosen structure type and the predominant cracking. The IRI is a direct result of cracking and is sensitive to CTE variation. Nevertheless, the predicted smoothness/roughness responses throughout the 20 percent CTE range maintain IRI values far below the Florida analysis limit of 160 inches per mile. Consequently, IRI is not a concern for Florida JPCP under the evaluated climate and traffic conditions, but its CTE sensitivity should not be underestimated in general.

## SUMMARY AND CONCLUSIONS

To properly implement the AASHTO Mechanistic-Empirical Pavement Design Guide (M-E PDG) in Florida, three common Florida concrete paving mixtures were evaluated to measure the engineering properties and M-E analysis input parameters with an emphasis on the coefficient of thermal expansion (CTE) according to the AASHTO TP-60 test protocol (AASHTO 2005). Various pavement models were generated in consideration of the local conditions. Nine different Jointed Plain Concrete Pavement (JPCP) models were established and analyzed for top layer thickness. The thickness analysis results demonstrated that cracking is the critical performance criterion, and top-down fatigue damage was identified to be the controlling failure mechanism. The CTE sensitivity study was carried out for the nine rigid pavement models.

Predicted faulting performance was illustrated to be slightly affected by CTE increase and its sensitivity was demonstrated to be negligible. Bottom-up damage was affected in thin concrete slabs only. As for top-down damage, regardless of layer thickness or any other factors, the cumulative damage pattern could be consistently described by a quadratic function for interchanged thermal expansion values. This

finding was significant because predicted cracking and IRI could be derived from top-down damage and showed related response patterns. The study will provide useful reference for pavement analysis according to the AASHTO M-E RPDG.

In conclusion, with regard to the CTE variation, the JPCP analysis according to the AASHTO M-E RPDG (considering the given design features and local climate and traffic conditions) is minimally sensitive to faulting, sensitive to bottom-up damage (for thin PCC layers), and extremely sensitive to top-down damage, cracking, and smoothness/roughness. Over all, two (cracking and IRI) out of three pavement performance criteria are highly susceptible to CTE.

## ACKNOWLEDGEMENT

Funding for this research was provided by the Florida Department of Transportation (FDOT). Emmanuel Uwaibi and Bruce Dietrich were the FDOT project managers and provided excellent support. Charles Ishee with the FDOT Materials Office was extremely helpful throughout the experimental phase. The technical problems encountered during CTE measurement would not have been solved without the outstanding support of Jussara Tanesi and Jagan Gudimettla, contractor working for the Turner-Fairbank Highway Research Center and for FHWA headquarters office, respectively. Marc Ansley with the FDOT Structures Research Laboratory provided strong support for the testing program.

## REFERENCES

American Association of State and Highway Transportation Officials (AASHTO) (2005). "AASHTO TP-60: Standard Method of Test for Coefficient of Thermal Expansion of Hydraulic Cement Concrete." AASHTO, Washington, D.C.

American Association of State and Highway Transportation Officials (AASHTO) (2008). Executive Committee 2007/2008. *Mechanistic-Empirical Pavement Design Guide, Interim Edition: A Manual of Practice*. AASHTO, Washington, D.C.

Coree, B., Ceylan, H., and Harrington, D. (2005). "Implementing the Mechanistic-Empirical Pavement Design Guide: Implementation Plan." *Report* No. IHRB Project TR-509. 38p.

Crawford, G.L., Gudimettla, J.M., Tanesi, J. (2010). "Inter-Laboratory Study on Measuring Coefficient of Thermal Expansion of Concrete." *Transportation Research Board (TRB) 89th Annual Meeting Compendium of Papers CD-ROM*, TRB, Washington, D.C. 15p.

Florida Department of Transportation (FDOT). (2005). *2004 Florida Traffic Information*. CD-ROM, Tallahassee, Florida.

Florida Department of Transportation (FDOT). (2009). *Rigid Pavement Design Manual*. Pavement Management Office, Document No. 625-010-006-e, Tallahassee, Florida.

Guclu, A., and Ceylan, H (2005). "Sensitivity Analysis of Rigid Pavement Systems Using Mechanistic-Empirical Pavement Design Guide." *Proceedings of the 2005 Mid-Continent Transportation Research Symposium*, Ames, Iowa. pp. 555-562

Hall, K. D., and Beam, St. (2004). "Estimation of the Sensitivity of Design Input Variables for Rigid Pavement Analysis Using the Mechanistic-Empirical Design Guide." *Transportation Research Board (TRB) 83rd Annual Meeting Compendium of Papers CD-ROM*, TRB, Washington, D.C. 9p.

Hossain, M., Khanum, T., Tanesi, J., and Schieber G. (2006). "Portland Cement Concrete Coefficient of Thermal Expansion Input for Mechanistic-Empirical Pavement Design Guide." *Transportation Research Board (TRB) 85th Annual Meeting Compendium of Papers CD-ROM*, TRB, Washington, D.C. 21p.

Kampmann, R. (2008). "Engineering Properties of Florida Concrete Mixes for Implementing the AASHTO Recommended Mechanistic-Empirical Rigid Pavement Design Guide", M.S. Thesis, Florida State University, Tallahassee, Florida. p210.

Mallela, J., Abbas, A., Harman, T., Rao, C. Liu, R., and Darter, M. I. (2005). "Measurement and Significance of Coefficient of Thermal Expansion of Concrete in Rigid Pavement Design." *Transportation Research Board (TRB) 84th Annual Meeting Compendium of Papers CD-ROM*, TRB, Washington, D.C. 9p.

Nantung, T., Chehab G., Newbolds S., Galal, K., Li, K., and Kim, D. H. (2005). "Implementation initiatives of the Mechanistic-Empirical Pavement Design Guide in Indiana." *Transportation Research Board (TRB) 84th Annual Meeting Compendium of Papers CD-ROM*, TRB, Washington, D.C. 10p.

National Cooperative Highway Research Program (NCHRP) 1-37A and 1-40D (2007). *Mechanistic-Empirical Pavement Design Guide Software*, Version 1.003. (Build May 24, 2007) Software for review

Salman, H.K., Buck, N., and Chatti, K. (2007). "Significant M-E PDG Design Inputs for Joined Plain Concrete Pavements in Michigan." *Proceedings of the Fifth International Conference on Maintenance and Rehabilitation of Pavements and Technological Control*, Park City, Utah. pp. 79-84.

Tanesi, J., Crawford, G.L., Nicolaescu, M., Meininger, R., Gudimettla, J.M. (2010). "How will the new AASHTO T336-09 CTE Test Method Impact You?" *Transportation Research Board (TRB) 89th Annual Meeting Compendium of Papers CD-ROM*, TRB, Washington, D.C. 14p.

Won, M. (2005). "Improvements of Testing Procedures for Concrete Coefficient of Thermal Expansion." *Transportation Research Board (TRB) 84th Annual Meeting Compendium of Papers CD-ROM*, TRB, Washington, D.C. 6p.

# SPEED AND TEMPERATURE EFFECTS ON FULL-SCALE PAVEMENT RESPONSES IN NON-CONVENTIONAL FLEXIBLE PAVEMENTS

April L. Ellison[1] and Dr. David H. Timm[2], P.E., A.M. ASCE

[1]Graduate Research Assistant, Department of Civil Engineering, Auburn University, Auburn, AL 36849; ellisap@auburn.edu; (334) 844-4320.
[2]Gottlieb Associate Professor of Civil Engineering, Auburn University, 238 Harbert Engineering Center, Auburn University, AL 36849; timmdav@eng.auburn.edu; (334) 844-6282.

## ABSTRACT

The importance of speed and temperature effects on flexible pavement response and performance are well-recognized within mechanistic-empirical pavement design and analysis procedures. While much laboratory testing and some field testing has been conducted to quantify these relationships on traditional mixtures (i.e., Superpave mixes with unmodified or polymer modified asphalt), there has been little investigation until now examining non-conventional materials, particularly in a full-scale test. Nine asphalt concrete (AC) test sections were constructed that included sulfur-modified material, high polymer content, warm-mix asphalt (WMA), high recycled asphalt pavement (RAP) content, Trinidad Lake Asphalt pellets, an open-graded friction course, and a hot-mix asphalt (HMA) control section. The sections were instrumented with an array of strain gauges to measure tensile strain at the bottom of the asphalt concrete. An equation was developed that adequately modeled seven of the nine test sections. The similar response characteristics of the seven sections indicate that they could be modeled in a similar fashion. Laboratory dynamic modulus testing should be conducted to further show how the different materials respond to temperature and speed and to explain why two of the sections did not conform to the model.

## INTRODUCTION

It is well known that temperature plays a significant role in flexible pavement response to loading. In general, pavement stiffness decreases as temperature increases. The reduction in modulus due to temperature increase leads to higher pavement strains under loading. Willis et al. (2009) showed that an increase in mid-depth pavement temperature resulted in a very large increase in tensile strain. It is important to consider pavement temperature when designing flexible pavements. Understanding the temperature effect on pavement response can lead to better designs and longer lasting pavement structures.

It is also well known that load duration (i.e., vehicle speed) has a significant effect on flexible pavement response. Many pavement design models ignore the dynamic loading effects that are inherent to any pavement structure. Several studies have shown that speed, or load duration, can have a significant effect on the strain response of flexible pavements (Sebaaly and Tabatabaee 1993; Siddharthan et al., 1996; Chatti et al.,1996; Dai et al., 1997; Mateos and Snyder 2002; Elseifi et al., 2006; Garcia and Thompson 2008). The strain response dependency on speed can be

explained by the viscoelastic nature of the hot mix asphalt (HMA). Because of the viscoelastic nature of the AC material, the material will show stiffer behavior under shorter loading times (Sebaaly and Tabatabaee 1993).

According to Garcia and Thompson (2008), frequency significantly affects the modulus of the hot-mix asphalt ($E_{HMA}$). $E_{HMA}$ can be calculated based on theoretical pavement response using laboratory tests and pavement modeling. Garcia and Thompson (2008) stated that to estimate $E_{HMA}$, it is necessary to simulate the frequency of loading (or pulse duration) that actually occurs in the field at the depth of interest. Theoretical pavement models are used to determine pavement response to wheel loads and, ultimately, predict pavement life. These models often fail to accurately characterize field conditions and result in erroneous pavement life prediction. According to Mateos and Snyder (2002), the inability of the multilayer linear elastic model to consider dynamic effects is a serious deficiency. There is added uncertainty in the case of modeling non-conventional materials. Some of these non-conventional materials include sulfur modified mixes, polymer modified mixes, and new warm-mix technologies. These and other non-conventional materials are growing in popularity within the paving industry. Research is needed to characterize these new materials for use in both pavement modeling and pavement life prediction.

Theoretical pavement modeling can be performed if the material properties are adequately characterized. In most cases, material properties are very difficult to define and loading conditions are assumed static for the purpose of simplifying the analysis (Sebaaly and Tabatabaee, 1993). An alternative approach to theoretical pavement modeling is direct response field-measurement through embedded instrumentation. This approach overcomes the need for material and loading assumptions, and provides a relatively high degree of certainty in response measurements. Tensile strain at the bottom of the AC layer is generally used as the critical pavement fatigue response.

The increase in use of non-conventional materials poses many questions about their response to vehicle loading and temperature changes. Although traditional HMA pavements respond in a predictable way to both variables, it cannot be assumed that the same relationships apply to these new materials. Therefore, a need for proper response characterization for these materials was needed and addressed through direct measurement in this investigation.

## OBJECTIVES

The first objective of this research was to compare the non-conventional materials' response to both pavement temperature and vehicle speed with the response of conventional HMA. The second objective was to provide a validation data set for model validation and refinement.

## SCOPE

This study was conducted at the National Center for Asphalt Technology (NCAT) Test Track located in Opelika, Alabama. The 1.7-mile long track is comprised of 46 unique 200-ft test sections. Nine sections were used for this investigation that included sulfur-modified material, high polymer content, WMA, high RAP content and Trinidad Lake Asphalt pellets. Five semi-trailer trucks were

used to apply the dynamic pavement loading. Testing was conducted on four test dates between December 2009 and May 2010. Four testing speeds were utilized to study the effect of loading rate on pavement response. Pavement instrumentation was used to measure pavement response to truck loading and mid-depth pavement temperature. Regression analyses were conducted on the data to ascertain the relative influence of speed and temperature on each section. It must be noted that this particular investigation was not intended to quantify relative performance of sections. That would require measurement of field distress and laboratory characterization of material properties. These tests are currently in-progress and the results will be forthcoming, but are not included in this paper since they are incomplete.

## METHODOLOGY

### Test Facility

Figure 1 provides a schematic of the nine Test Track sections and a general description of the different materials tested. The Thiopave® section (N6) is a sulfur-modified material developed by Shell Oil that features pelletized sulfur and a warm-mix additive. Section N7 is a highly polymer-modified mixture (7.5% SBS polymer) developed by Kraton Polymers. It was designed thinner than the other sections to better evaluate its strain capabilities. Sections N10 and N11 are both 50% RAP sections with N11 using a warm-mix technology for production. Section S8 was the same as the control section (S9), except that it featured an open-graded friction course surface. S10 and S11 were designed to be the same as the control, but were produced with foamed warm mix (S10) and additized warm mix (S11). Finally, the S12 asphalt binder was modified with 25% Trinidad Lake Asphalt pellets. Each section was constructed on a crushed granite base of approximately 6 in. depth over a common roadbed. The thicknesses in Figure 1 represent surveyed depths from the center of the instrumentation array.

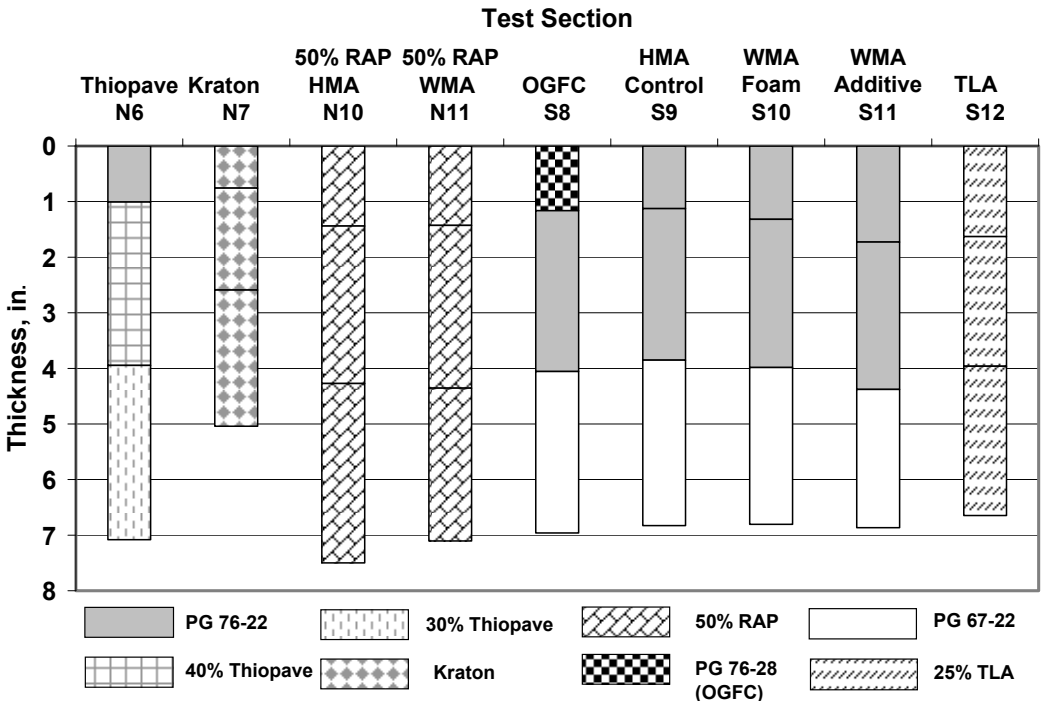



**Figure 1. Test Section Composition.**

A fleet of five tractor-triple-trailer trucks were used to load the pavement. Each truck was comprised of a steer axle (11,000 lb), tandem axle (40,000 lb), and 5 trailing single axles (21,000 lb per axle). Each axle type was examined separately and regression equations were developed for each. However, since similar trends were noted between all axle types, over 70% of the data collected represented single axles and a previous study at the Test Track (Willis et al., 2009) found single axles consistently produced higher strain levels. This paper focuses only on the single axle response to loading.

Pavement responses were measured with embedded instrumentation. A total of 12 CTL asphalt strain gauges were placed at the bottom of the AC. More information about the strain gauges can be found elsewhere (Priest and Timm, 2006). Six were oriented in the longitudinal direction while the other six were oriented in the transverse direction. The data analysis showed that longitudinal strains were consistently higher than transverse strains. For this reason, longitudinal strain responses are presented herein. The gauge array was centered in the right wheel path in each section.

Temperature probes were installed at selected depths within the pavement structure. More information about the temperature probes can be found elsewhere (Priest and Timm, 2006). The mid-depth AC temperature was selected for this study to predict pavement strain. Previous studies have shown mid-depth pavement temperature to be a good predictor of induced pavement strain (Priest and Timm, 2006; Willis et al., 2009).

**Testing Scheme**

Testing was carried out on four separate test dates ranging from December 2009 to May 2010 to capture a broad range of pavement temperatures. Table 1 provides the testing dates along with the average mid-depth pavement temperatures for each test section. Though general consistency was observed amongst temperatures on a given date, there were some large disparities especially on the latter dates. The reason for this was beyond the scope of this investigation, but was likely due to differences in materials and compaction. For example, S12 achieved the highest temperatures on the last three test dates and also had the highest density. Despite these differences, the effects of temperature on strain *were* incorporated in this investigation.

**Table 1. Average Temperature by Section and Date.**

| Section | Average AC Density ($g/cm^3$) | Average Mid-Depth Temperature, F | | | |
|---|---|---|---|---|---|
| | | 16-Dec-09 | 19-Feb-10 | 16-Apr-10 | 27-May-10 |
| N6: Sulfur-Modified WMA | 2.37 | 49.4 | 65.6 | 89.0 | 115.9 |
| N7: Highly-Modified SBS | 2.34 | 48.9 | 67.0 | 93.4 | 118.9 |
| N10: 50% RAP HMA | 2.39 | 49.6 | 66.4 | 90.5 | 116.4 |
| N11: 50% RAP WMA | 2.38 | 49.6 | 66.1 | 90.5 | 116.9 |
| S8: OGFC | 2.35 | 45.8 | 68.1 | 87.5 | 116.0 |
| S9: HMA Control | 2.29 | 48.7 | 67.1 | 84.0 | 112.9 |
| S10: WMA Foam | 2.36 | 46.9 | 70.6 | 90.6 | 120.9 |
| S11: WMA Additive | 2.35 | 49.4 | 66.9 | 84.0 | 113.5 |
| S12: TLA | 2.41 | 45.6 | 73.2 | 94.9 | 125.9 |

Testing was conducted at four different speeds: 15, 25, 35, and 45-mph. At least 3 passes of each of five trucks were recorded, providing typically 15 truck passes for each test date, section, speed, and axle type.

**Analysis**

The "best-hit" longitudinal strain magnitude under single-axle loading was determined for each speed, test date, and test section. Figure 2 provides a typical longitudinal strain trace and key points of the signal are noted. The strain magnitude was defined as the difference between the peak tensile and peak compressive strains. Since multiple readings were taken at a given speed on a given date (3 passes of each truck with 5 single axles over 6 longitudinal gauges), there were approximately 90 strain magnitudes per section and date from which the maximum or "best-hit" was determined.

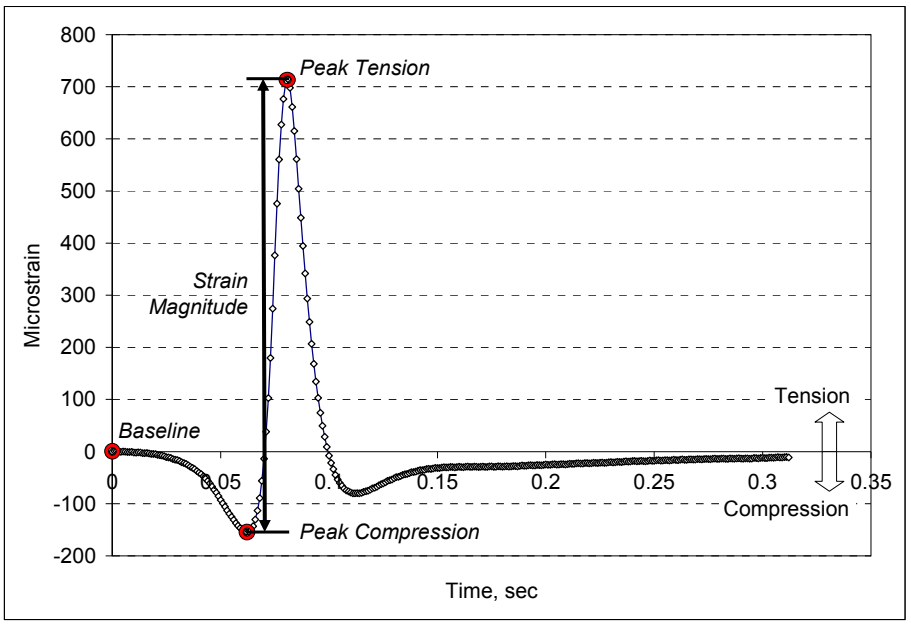



**Figure 2. Processed Longitudinal Strain Trace.**

## RESULTS AND DISCUSSION

A main objective of this research was to determine the effects of both vehicle speed and pavement temperature on flexible pavement response in the non-conventional materials. To accomplish this, a non-linear regression was performed for each test section. Section-specific equations were developed to predict pavement strain in each section based on mid-depth pavement temperature and vehicle speed:

$$\varepsilon = a * v^{b} * c^{T} \quad (1)$$

where:

ε = longitudinal microstrain at bottom of asphalt concrete
v = vehicle speed, mph
T = mid-depth pavement temperature, ˚F
a,b,c = section-specific regression coefficients

Equation 1 provided very good correlation for longitudinal strain under single axle load in each of the test sections. Table 2 shows the regression coefficients for Equation 1, coefficients of determination ($R^2$), and significance values (p-values) for each test section. It should be noted that all coefficients were significant (p-value <.05) with the exception of N10-a, S12-a, S12-b.

**Table 2. Correlation Coefficients and $R^2$ Values.**

| Section | a | p-value(a) | b | p-value(b) | c | p-value(c) | $R^2$ |
|---|---|---|---|---|---|---|---|
| **N6** | 128.772 | 0.00 | -0.209 | 0.00 | 1.027 | 0.00 | 0.997 |
| **N7** | 173.651 | 0.00 | -0.167 | 0.00 | 1.022 | 0.00 | 0.983 |
| **N10** | 1373.058 | **0.20** | -0.843 | 0.00 | 1.019 | 0.00 | 0.737 |
| **N11** | 135.253 | 0.00 | -0.168 | 0.00 | 1.020 | 0.00 | 0.989 |
| **S8** | 154.340 | 0.00 | -0.168 | 0.00 | 1.023 | 0.00 | 0.995 |
| **S9** | 167.563 | 0.00 | -0.243 | 0.00 | 1.025 | 0.00 | 0.988 |
| **S10** | 332.815 | 0.00 | -0.353 | 0.00 | 1.019 | 0.00 | 0.949 |
| **S11** | 277.173 | 0.03 | -0.250 | 0.04 | 1.018 | 0.00 | 0.872 |
| **S12** | 58.318 | **0.14** | 0.242 | **0.18** | 1.017 | 0.00 | 0.827 |

As shown in Table 2, six of the nine sections had $R^2$ values greater than 0.94. The lowest $R^2$ value was found to be 0.737 for section N10. Due to the high correlation values given here, Equation 1 was determined adequate for prediction of strain response of the flexible pavements. Sections N10 and S12 did not correlate well to the model and, therefore, required further investigation.

These sections were investigated to determine cause for their deviation from the model. A closer look revealed that both N10 and S12 data were erratic for the last data collection date which corresponded to the highest test temperatures. Examination of additional Test Track data, gathered on a weekly basis since August 2009 at variable temperatures but speeds approximately equal to 45 mph, indicated highly variable readings for these two sections with data collected after April, 2010. Due to the variability of the data for sections N10 and S12, and their corresponding insignificant p-values of regression constants, these sections were excluded from further analyses.

The negative sign of the b-coefficients indicated that an increase in speed resulted in a decrease in strain. This relationship was expected and was supported by the literature review (Sebaaly and Tabatabaee 1993; Siddharthan et al., 1996; Chatti et al.,1996; Dai et al., 1997; Mateos and Snyder 2002; Elseifi et al., 2006; Garcia and Thompson 2008). The positive b-coefficient for S12 can be explained by the fact that b was not statistically-significant for that section. Table 2 indicated that speed has the greatest effect on section N10. However, it corresponded to an extremely high a-term which was not statistically-significant. It can be noted from the c-values in Table 2 that temperature had a nearly identical effect on each section. The conclusion that these non-conventional materials show similar response to temperature was a promising finding since it would imply that they can be modeled like conventional materials.

Equation 1 was used to predict longitudinal strain in each section for mid-depth temperatures ranging from 50˚F to 120˚F. Data were generated for each of the four testing speeds: 15, 25, 35, and 45 mph. The effect of speed on pavement strain was analyzed first. Figure 4 shows the results for a reference temperature of 70˚F.

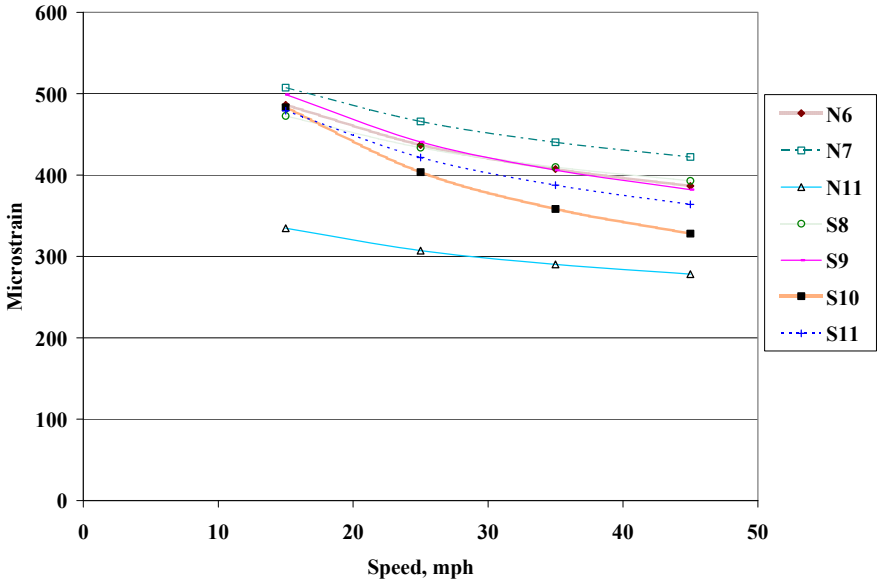



**Figure 4. Strain versus Speed at 70˚F.**

Figure 4 showed that strain decreases as speed increases for each section, as expected. Similar trends were noted for the other reference temperatures as well. It can be seen from Figure 4 that Section S10 is most affected by speed at this temperature while N11 is the least affected. This was also the case for other reference temperatures.

Referring to the b-regression constants in Table 2, it appears that the warm-mix sections (S10, S11) are slightly more sensitive to speed relative to the control section (S9). The other sections were slightly less-sensitive to speed relative to the control. However, overall, it could be said that they have similar responses to speed and could be modeled similarly.

In terms of strain magnitude, it was expected and shown in Figure 4 that section N7 would have the highest overall strain level due to its thinner HMA cross-section relative to the other sections. Section N11 was expected to have the lowest strain levels since it was 50% RAP and logically a stiffer material resulting in lower strain levels. These strain levels, and those of the other sections, are not directly related to performance. For example, the Kraton section (N7) was intentionally designed to sustain higher strain levels, so the fact that the strains are higher than the others does not necessarily imply it will have poorer performance.

The next step in the study was to investigate the effect of temperature on pavement strain. Plots of predicted pavement strain versus mid-depth temperature were generated for each of the four test speeds. For brevity, Figure 5 shows only the plot for the 45 mph test speed.

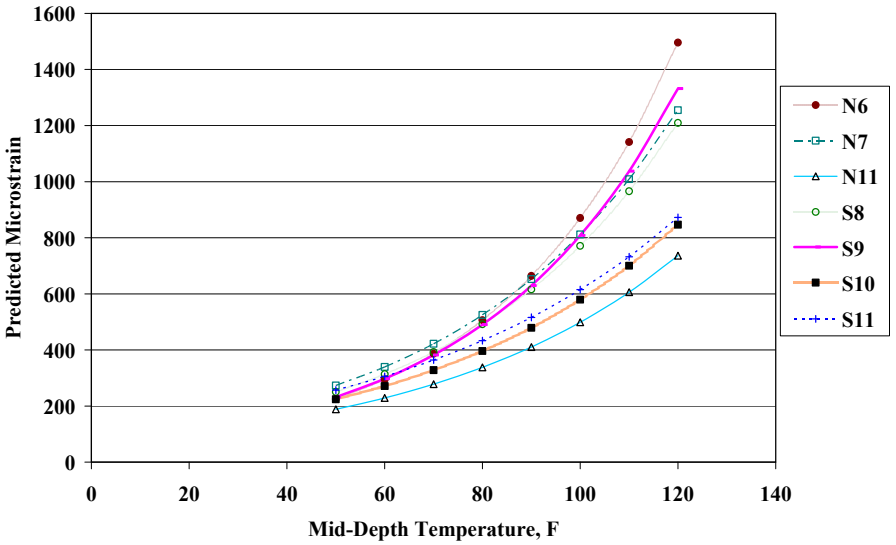



**Figure 5. Strain versus Temperature at 45-mph.**

Figure 5 indicated that pavement strain increases as mid-depth temperature increases. This was expected as higher temperatures result in softer pavements. It can be seen that N6 experienced the most rapid increase in strain due to temperature and the highest overall predicted strain at the highest temperature. It is interesting to note that the Kraton section tracks lower than the control section at temperatures greater than 100F. The high polymer content may be contributing toward a stiffening at higher temperatures. Section N11 was again lower, as expected, since it was the high RAP section. These relationships were also observed for the 15, 25, and 35-mph test speeds.

## CONCLUSIONS AND RECOMMENDATIONS

The overall objective of this research was to compare a variety of materials' responses to a conventional HMA cross-section and provide field-validation data sets for model development and refinement. Though laboratory dynamic modulus and beam fatigue testing is in progress on these materials, the testing is not yet complete. Therefore, the differences in strain seen above can not be objectively transformed into predictions of performance. With that in mind, the following conclusions and recommendations can be made:

1. Sections N10 and S12 should be further investigated. Their deviation from conforming to the speed-temperature model could result from simple testing variability, or from differences in material behavior. Laboratory dynamic modulus testing should help answer this question.
2. The fact that the remaining sections conformed to the same speed-temperature model implies they have similar response characteristics and could be modeled in pavement design in a similar fashion. Differences in how materials respond to temperature and speed should be captured by dynamic modulus testing, when complete.

3. When laboratory dynamic-modulus and beam fatigue testing are completed, response modeling and performance predictions should be conducted to compare the measured responses to predicted responses and estimate the performance of all sections.

**ACKNOWLEDGEMENTS**

The authors wish to thank the following state departments of transportation for their continued cooperation and support of this research: Alabama, Florida, North Carolina, Oklahoma and Tennessee. The authors are also appreciative of Kraton Polymers LLC, Lake Asphalt of Trinidad and Tobago, Ltd. and Shell Sulphur Solutions for their support and cooperation. Finally, the Federal Highway Administration also deserves special recognition for their support and cooperation.

**REFERENCES**

Chatti, K., Kim, H. B., Yun, K. K., Mahoney, J. P., Monismith, C. L. (1996). "Field Investigation into Effects of Vehicle Speed and Tire Pressure on Asphalt Concrete Pavement Strains." *Transportation Research Record*, 1539, pp. 66-71.

Dai, S.T., Van Deusen, D., Beer, M., Rettner, D., Cochran, G. (1997). "Investigation of Flexible Pavement Response to Truck Speed and FWD Load Through Instrumented Pavements."

Elseifi, M. A., Al-Qadi, I. L., Yoo, P. J. (2006). "Viscoelastic Modeling and Field Validation of Flexible Pavements." *Journal of Engineering Mechanics*, ASCE, pp. 172-178.

Garcia, G., Thompson, M. R. (2008). "Strain Pulse Duration Considerations for Extended Life Hot Mix Asphalt Pavement Design." *Proceedings of the 87th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Hornyak, N., Crovetti, J. A. (2009). "Analysis of Load Pulse Durations for Marquette Interchange Instrumentation Project." *Transportation Research Record,* 2094, pp. 53-61.

Mateos, A., Snyder, M. B. (2002). "Validation of Flexible Pavement Structural Response Models with Data from the Minnesota Road Research Project." *Transportation Research Record,* 1806, pp. 19-29.

Priest, A. L, Timm, D. H. (2006). "Methodology and Calibration of Fatigue Transfer Functions for Mechanistic-Empirical Flexible Pavement Design." *National Center for Asphalt Technology*, Report No. 06-03.

Sebaaly, P. E., Tabatabaee, N. (1993). "Influence of Vehicle Speed on Dynamic Loads and Pavement Response." *Transportation Research Record*, 1410, pp. 107-114.

Siddharthan, R., Yao, J., Sebaaly, P. E. (1996). "Field Verification of Moving Load Model for Pavement Response." *Transportation Research Record*, 1540, pp. 125-131.

Willis, R., Timm, D., West, R., Powell, B., Robbins, M., Taylor, A., Smit, A., Tran, N., Heitzman, M., Bianchini, A. (2009). "Phase III NCAT Test Track Findings." *National Center for Asphalt Technology*, Report No. 09-08.

# Effect of Construction Environment on JPCP Performance

Daba S. Gedafa, Ph.D., M.ASCE[1], Taslima Khanum, M.ASCE[2], Mustaque Hossain, Ph.D., P.E., F.ASCE[3], and Gregg Schieber, P.E., M.ASCE[4]

**Abstract:** Some properties of newly paved Jointed Plain Concrete Pavements (JPCP) are known to influence the long-term performance. The traditional empirical design procedures for JPCP were unable to take into account most of these factors. However, the new Mechanistic-Empirical Pavement Design Guide (MEPDG) accounts for climatic conditions, local materials, selected construction practices, and actual highway traffic distribution. In this study, performance (in terms of International Roughness Index (IRI), faulting, and percent slab cracked) of six typical JPCP pavements in Kansas corresponding to alternative inputs of Portland Cement Concrete (PCC) strength development, PCC shrinkage, and "zero-stress" temperature has been evaluated using MEPDG.

The results show that predicted JPCP roughness (IRI) and faulting by MEPDG are not very sensitive to the PCC strength. However, slab cracking decreases with higher PCC strength. In general, PCC shrinkage does not affect predicted IRI. Higher shrinkage strain results in higher faulting. Long term cracking appears to be fairly insensitive to the shrinkage strain. MEPDG-predicted IRI and percent slabs cracked are fairly insensitive to the zero-stress temperature but the faulting is severely affected except on a JPCP pavement section with widened lane and tied PCC shoulder. Percent slab cracked highly depends on the PCC slab thickness. April and October are the best months for JPCP construction (paving) in Kansas.

**Keywords:** Construction environment; JPCP performance; MEPDG.

## Introduction

A number of variables are known to influence the long-term performance of Jointed Plain Concrete Pavement (JPCP). These variables can be grouped into the following categories: (1) Site conditions, (2) Pavement design features, and (3) Construction practices (Owusu-Antwi et al. 1998). The site conditions include traffic loading, climate, and the foundation type, and support. The design features that influence JPCP performances include joint spacing, and load transfer design. Construction factors that are important in determining JPCP performance are Portland cement concrete (PCC) mix design and properties, method of paving, method of dowel installation, method of finishing, texturing and curing, etc. (Owusu-Antwi et al. 1998). For longer lasting JPCP design, a clear understanding and consideration of these factors are needed.

[1]University of Connecticut, Department of Civil & Environmental Engineering, 261 Glenbrook Rd Unit 2037, Storrs, CT 06269. E-mail:dgedafa@engr.uconn.edu

[2]Caltrans District 4, 111 Grand Avenue, Oakland, CA 94623. E-mail:taslima_khanum@dot.ca.gov

[3]Kansas State University, Department of Civil Engineering, Manhattan, KS 66506. E-mail: mustak@ksu.edu

[4]Kansas Department of Transportation, Bureau of Materials & Research., 2300 Van Buren, Topeka, KS 66611. E-mail: gregs@ksdot.org

## Problem statement

Historically only the mechanical properties of PCC mixture have been used in JPCP design. For example, the 1986 and 1993 American Association of State Highway and Transportation Officials (AASHTO) Guide procedures do not consider any PCC mixture properties other than the modulus of rupture and the modulus of elasticity for JPCP design (AASHTO 1986, AASHTO 1993). However, the new MEPDG accounts for climatic conditions, local materials, selected construction practices, and actual highway traffic distribution by means of axle load spectra (NCHRP 2004). Thus MEPDG is a very useful tool for studying the effect of construction environment on predicted performance of existing JPCPs.

## Objective

The objective of this study was to evaluate the performance of typical JPCP pavements in Kansas due to alternative PCC inputs corresponding to (1) PCC strength, (2) PCC shrinkage, and (3) "zero-stress" temperature.

## Test sections

Six in-service JPCP projects were selected for the MEPDG analysis. Three of these projects were the experimental sections chosen from the Kansa Department of Transportation (KDOT) Specific Pavement Studies (SPS-2) project located on Interstate route 70 (I-70). Two other projects are also located on I-70, and one on route K-7. Table 1 tabulates the features of these sections. The SPS-2 test sections are each 152.4 m long, and the rest are 1.6 kilometer to several kilometers long.

All sections have 4.6 m joint spacing with dowelled joints. SPS-2 sections have 37 mm diameter steel dowels. On KDOT sections, dowel diameter was one-eighth of the PCC slab thickness. All sections have 3.7 m lanes with tied concrete shoulders except SPS-2 Section 6. That section has a widened lane of 4.3 m with tied PCC shoulder. The sections were constructed on stabilized base and treated subgrade. Base stabilization was done with Portland cement. Depending on the cement content and gradation, the bases were designated as Portland cement-treated base (PCTB) or lean concrete base (LCB).

Base thickness ranged from 10 to 15 cm. The projects have primarily silty clay as subgrade. The top 15 cm of the natural subgrade were treated with lime or fly ash to reduce the plasticity and/or control moisture during construction. The PCC slab thickness, designed according to the 1986 or 1993 AASHTO design guide, ranged from 22.5 to 30 cm. The strength (modulus of rupture) shown in Table 1 are the actual average values obtained during construction. The as-constructed International Roughness Index (IRI) values on these projects varied from 0.94 m/km to 1.94 m/km.

The annual average daily traffic (AADT) on these sections ranged from 9,200 for the I-70 Gear County project to 36,000 for the I-70 Shawnee County project. Very low percentage of truck traffic was observed on the I-70 Shawnee County project.

**Table 1.** Project Features of the Study Sections

| Project ID | Route | County | Year Built | Mile post Limit | PCC Thickness (cm) | PCC 28-day Tensile Strength (MPa) | Subgrade Soil Type | Initial AADT | % Truck | Initial IRI (m/km) |
|---|---|---|---|---|---|---|---|---|---|---|
| K-2611-01* | I-70 | Geary | 1990 | 0 – 7 | 28 | 4.8 | A-6 | 9,200 | 18 | 0.95 |
| K-3344-01** | I-70 | Shawnee | 1993 | 9 – 10 | 26.5 | 3.3 | A-7-6 | 36,000 | 5 | 1.52 |
| SPS-2 (Sec-5)† | I-70 | Dickinson | 1992 | 20 – 22.61 | 28 | 6.5 | A-6 | 11,970 | 22.3 | 1.93 |
| SPS-2 (Sec-6)† | I-70 | Dickinson | 1992 | 20 –22.61 | 28 | 4.3 | A-6 | 11,970 | 22.3 | 1.55 |
| SPS-2 (Control)* | I-70 | Dickinson | 1992 | 20 – 22.61 | 30.5 | 4.5 | A-6 | 11,970 | 22.3 | 1.51 |
| K-3382-01** | K-7 | Johnson | 1995 | 12 – 15 | 23 | 3.7 | A-7-6 | 13,825 | 7 | 1.28 |

*15 cm Portland Cement-Treated Base (PCTB)
** 10 cm Portland Cement-Treated Base (PCTB)
† 15 cm Lean Concrete Base (LCB)

## MEPDG Rigid Pavement Design Inputs

Input data used for the MEPDG analysis of concrete pavements are categorized as: a) General information; b) Traffic; c) Climate; d) Pavement structures; and e) Miscellaneous. Table 2 shows some of the inputs in this study. For this study, monthly and hourly truck distribution, truck class distribution, axle load distributions, and some other general traffic inputs were derived from Kansas weigh-in-motion (WIM) and/or Automatic Vehicle Classification (AVC) data. The design reliability used for all projects was 90%, and the corresponding failure criterion for IRI was 2.60 m/km, 3.05 mm for faulting, and 15% for slab cracking.

## Results and Discussions

The MEPDG analysis was done for 20 years. Typical JPCP distresses, IRI, faulting, and percent slabs cracked, were calculated and compared after 20 years.

### *Strength*

**Smoothness or IRI.** The MEPDG-predicted IRI for the sections was compared at three levels of strength - low (21 MPa), average (35 MPa), and high (55 MPa). Figure 1 (a) shows the results. In general, the strength does not affect predicted IRI. There is a slight effect on the K-7 Johnson County project. That section has the thinnest PCC slab among all sections. When the strength was increased from 21 MPa to 35 MPa, IRI decreased from 1.75 m/km to 1.70 m/km. When the strength was increased to 55 MPa, IRI was 1.66 m/km. These decreases are negligible for all practical purposes.

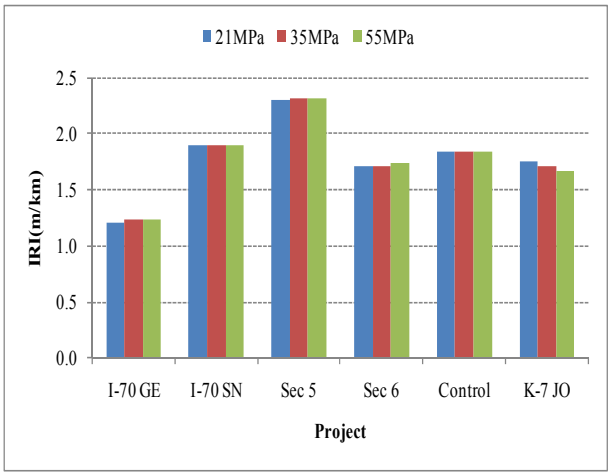
**Table 2.** Input Parameters for MEPDG Rigid Pavement Design Analysis

| INPUT PARAMETERS | Input Value | | | | | |
|---|---|---|---|---|---|---|
| | **I-70 GE** | **I-70 SN** | **SPS-2 (Sec 5)** | **SPS-2 (Sec 6)** | **SPS-2 (Control)** | **K-7 JO** |
| ***General Information*** | | | | | | |
| Pavement construction Date | Nov 90 | Oct 93 | July 92 | July 92 | July 92 | Sep 95 |
| Initial IRI (m/km) | 0.95 | 1.52 | 1.93 | 1.55 | 1.51 | 1.28 |
| ***Traffic*** | | | | | | |
| Initial two-way AADTT | 1,656 | 1,800 | 2,670 | 2,670 | 2,670 | 968 |
| No. of lanes in design direction | 2 | 2 | 2 | 2 | 2 | 2 |
| Traffic growth factor (%) | 1.2 | 3 | 3.5 | 3.5 | 3.5 | 6.7 |
| Design lane width (m) | 3.7 | 3.7 | 3.7 | 4.3 | 3.7 | 3.7 |
| ***PCC Layer*** | | | | | | |
| PCC Layer thickness (cm) | 27.9 | 26.7 | 27.9 | 27.9 | 30.5 | 22.9 |
| Material Unit Weight ($kN/m^3$) | 22.0 | 22.3 | 22.5 | 21.9 | 22.9 | 22.3 |
| Cement Type | II | I | II | II | II | II |
| Cement Content ($kg/m^3$) | 387.5 | 373.6 | 511.2 | 315.5 | 355.8 | 369.4 |
| Poisson's ratio | 0.20 | 0.20 | 0.20 | 0.20 | 0.20 | 0.20 |
| Coeff. of thermal. expansion ($m/m/°C \times 10^{-6}$) | 9.9 | 9.9 | 9.9 | 9.9 | 9.9 | 9.9 |
| Derived Shrinkage Strain (μm)* | 621 | 727 | 596 | 423 | 456 | 670 |
| Water-cement ratio (w/c) | 0.44 | 0.411 | 0.35 | 0.35 | 0.42 | 0.46 |
| ***Base Material*** | | | | | | |
| Base Type | PCTB | PCTB | LCB | LCB | PCTB | PCTB |
| Base Thickness (cm) | 15 | 15 | 15 | 15 | 15 | 10 |
| Base material unit wt. ($kN/m^3$) | 21.2 | 21.2 | 21.3 | 21.3 | 21.2 | 21.2 |
| Base Modulus (MPa) | 3450 | 3450 | 13800 | 13800 | 3450 | 3450 |
| ***Treated Subgrade*** | | | | | | |
| Subgrade type | N/A | LTSG | FASG | FASG | FASG | LTSG |
| Subgrade modulus (MPa) | N/A | 345 | 345 | 345 | 345 | 345 |
| Unit weight ($kN/m^3$) | N/A | 19.6 | 19.6 | 19.6 | 19.6 | 19.6 |
| Poisson's ratio | N/A | 0.20 | 0.15 | 0.15 | 0.15 | 0.20 |
| ***Compacted Subgrade*** | | | | | | |
| Subgrade soil type | A-6 | A-7-6 | A-6 | A-6 | A-6 | A-7-6 |
| Subgrade Modulus (MPa) * | 67 | 43 | 48 | 48 | 52 | 50 |
| Plasticity index, PI | 15.8 | 25.7 | 26 | 26 | 23 | 19.9 |
| Percent passing # 200 sieve | 71.8 | 93.3 | 78.1 | 78.1 | 76.9 | 94.3 |
| % passing # 4 sieve | 100 | 100 | 100 | 100 | 98 | 100 |
| $D_{60}$ (mm) | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |

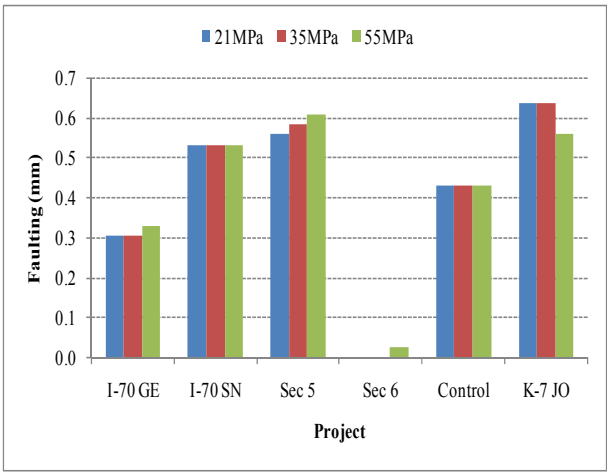
* computed by MEPDG

**Faulting.** Figure 1(b) shows the predicted faulting on all sections corresponding to three levels of strength. Almost no changes in faulting values were observed for all projects. The predicted faulting values were also negligible for all practical purposes. It appears that faulting is also insensitive to strength.

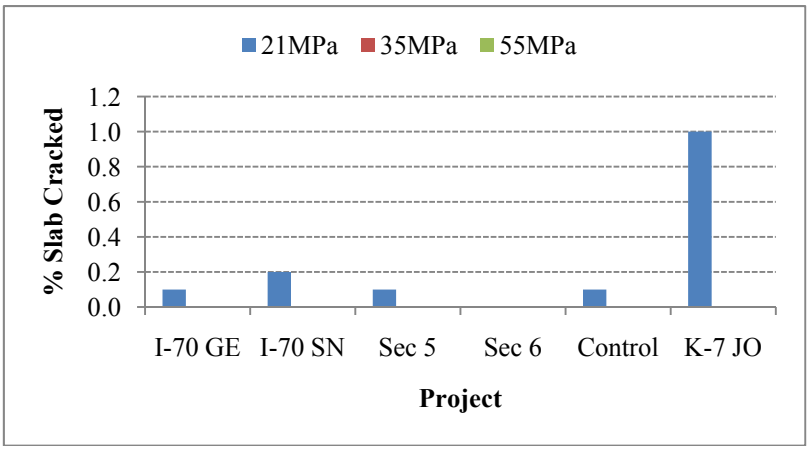
**% Slabs Cracked.** Figure 1(c) illustrates the effect of strength development on predicted percent slabs cracked. Although very small amounts of cracking were observed almost on all projects at 21 MPa level, no cracking was observed when the strength was increased to 35 MPa. The biggest change was for K-7.




(a) IRI




(b) Faulting

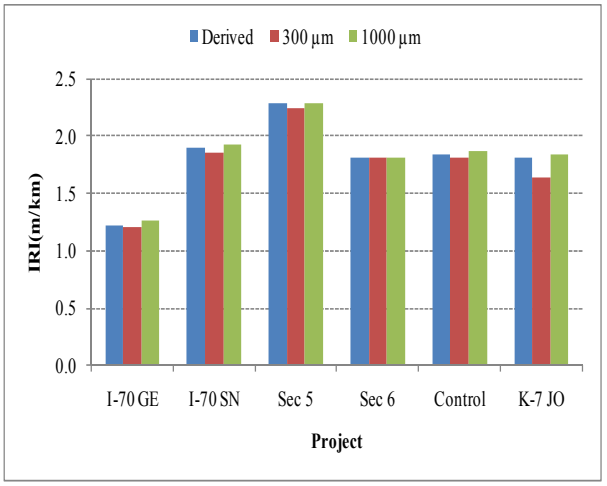



(c) % Slab Cracked

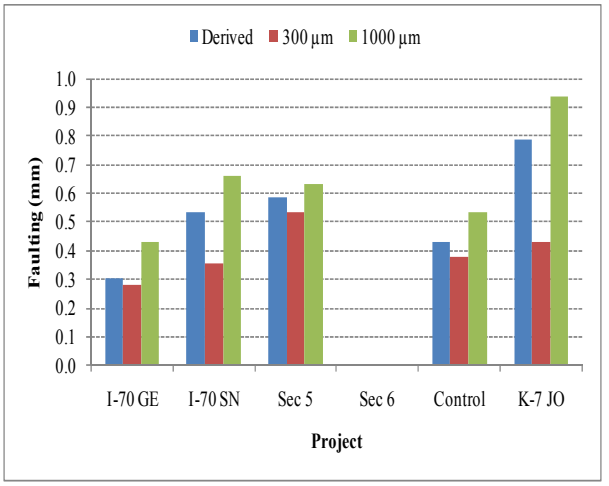
**Figure 1.** Predicted JPCP distresses for varying PCC strength

### *Shrinkage*

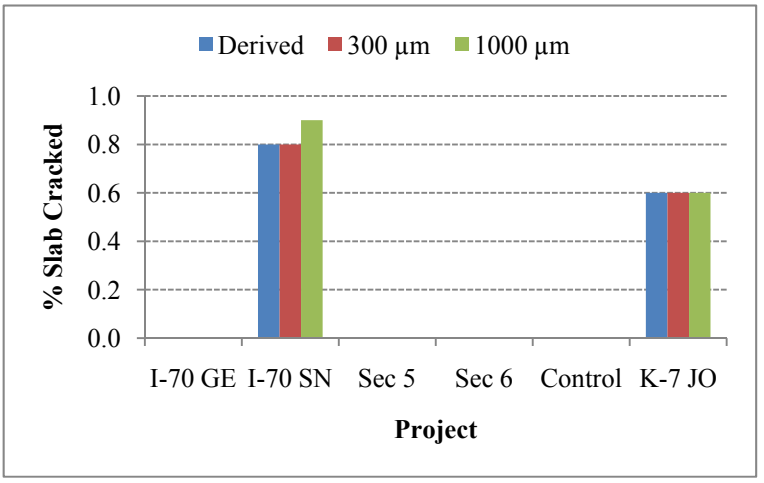
**Smoothness or IRI**. The MEPDG-predicted 20-year IRI values for the sections were compared at three levels of shrinkage strain –MEDPG derived, low (300 μm), and high (1,000 μm). Figure 2 (a) shows the results. In general, the shrinkage does not greatly affect IRI. The derived and lower shrinkage strain levels tend to predict similar IRI. There is a slight effect for the K-7 Johnson County project. This section has the thinnest PCC slab among all sections. When the shrinkage strain increased from 300 μm to 1,000 μm, IRI increased from 1.63 m/km to 1.80 m/km. According to the MEPDG algorithm, higher shrinkage strain results in higher faulting. That, in turn, is responsible for increased roughness.




(a) IRI




(b) Faulting




(c) % Slabs Cracked

**Figure 2.** Predicted JPCP distresses for different shrinkage strain input.

**Faulting.** Figure 2(b) shows the predicted faulting on all sections corresponding to three levels of shrinkage strain. Higher shrinkage strain results in higher faulting. The effect is most pronounced for the I-70 Shawnee and K-7 Johnson County projects. When the shrinkage strain increased from 300 μm to 1,000 μm, faulting almost doubled though the faulting values are negligible for all practical purposes. Nevertheless, faulting is very sensitive to the shrinkage strain.

**Percent Slabs Cracked.** Figure 2(c) illustrates the effect of shrinkage strain on predicted percent slabs cracked for the JPCP projects in this study. Only two projects showed cracking, and the effect of shrinkage strain is almost negligible. Only I-70 Shawnee County project showed a slight increase in cracking with higher strain. Cracking appears to be fairly insensitive to the shrinkage strain.

### *PCC Zero-Stress Temperature*

PCC zero-stress temperature, Tz, is an important parameter that affects the stress buildup in the PCC slab immediately after construction. This parameter is also related to the time of construction since it is computed based on the cement content and the mean monthly temperature (MMT) during construction as shown below (NCHRP 2004):

$$T_z = (CC*0.59328*H*0.5*1000*1.8/ ((1.1*2400) + MMT) \quad (1)$$

where,

$T_z$ = Temperature at which the PCC layer exhibits zero thermal stress;
CC = Cementitous material content, lb/yd$^3$;
H = -0.0787+0.007*MMT-0.00003*MMT$^2$; and
MMT = Mean monthly temperature for the month of construction, º F.

Since MMT varies widely in Kansas and JPCP's can be built any time from April to December, this parameter is of much interest to KDOT. Tz was computed for different months of the year for all projects based on the mean monthly temperatures of the year of construction from nearby weather stations available at the Kansas State University Weather Data Laboratory. It is obvious that MMT has a larger impact on the computed Tz values. The Tz values are higher for the summer months. Comparison of MEPDG-derived Tz versus computed Tz's in this study is shown in Table 3. $T_z$ was derived by MEPDG based on the month of construction and project-specific climatic file (.icm) from the weather station. It can be seen that computed and derived Tz values are similar for almost all projects, except the I-70 Geary and I-70 Shawnee County projects. Both projects show lower computed $T_z$ values compared to the derived ones. As $T_z$ plays an important role in prediction of pavement distresses, it is suggested that average MMT from 30 years of climatic data be used in Tz calculation for new JPCP design. Therefore, updated MEPDG climatic files will also be necessary for future design purposes.

**Table 3**. Computed and Derived $T_z$ for MEPDG Analysis

| | PCC Zero-Stress Temperature Tz (° C) | | | | | |
|---|---|---|---|---|---|---|
| | **I-70 GE** | **I-70 SN** | **SPS-2 (Sec 5)** | **SPS-2 (Sec 6)** | **SPS-2 (Control)** | **K-7 JO** |
| MEPDG | 17.2 | 30.0 | 54.4 | 43.9 | 46.1 | 38.9 |
| Computed | 12.2 | 17.8 | 54.4 | 43.3 | 46.1 | 37.8 |

Figure 3 shows the predicted JPCP distresses by MEPDG corresponding to three probable and one actual construction months for the projects in this study. Three probable construction months were chosen based on an analysis of the MMT values obtained from the weather database for the years of construction of these projects. The months of April, August, and December were selected to represent temperature, high and low MMT or Tz values. October was also chosen but later disregarded since MMT values for this month are very similar to those in April. Actual construction months for the projects, shown in Table 2, are as follows: I-70 Geary County: November 1990; I-70 Shawnee County: October 1993; SPS-2's: July 1992; and K-7 Johnson County: September 1995.

**Smoothness or IRI.** Figure 3(a) shows that construction month/Tz does not greatly affect the predicted IRI. For almost all projects that were constructed in April show slightly lower IRI. However, the initial IRI results in Table 2 show that the SPS sections, built in July 1992, have the highest initial or as-constructed IRI among all sections.

**Faulting.** Construction months/Tz tends to make the biggest difference in predicted faulting as shown in Figure 3(b). It is clear that the pavements constructed in August (with highest MMT/Tz) have much higher faulting than those constructed in temperate climate in April or even in colder time in December. The effect is very pronounced on I-70 Shawnee County and K-7 Johnson County - the projects with lower PCC strength and thinner PCC slab, respectively. The only project which is not affected by this parameter is SPS-2 Section 6. This pavement has a widened lane and that appears to address the higher faulting effect due to construction during the month with high MMT.

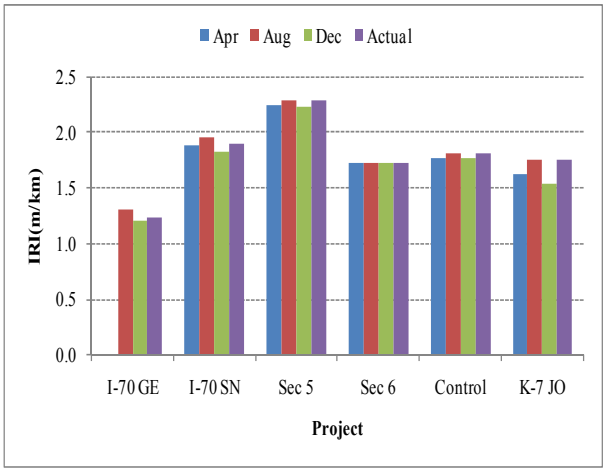
**% Slab Cracked.** Figure 3(c) shows that predicted slab cracking is not highly affected by the construction month. However, both I-70 Shawnee County and K-7 Johnson County, where some cracking was observed, showed slightly less slab cracking for construction during April.

Considering all results, it appears that April and October are the two best months for JPCP construction (paving) in Kansas.

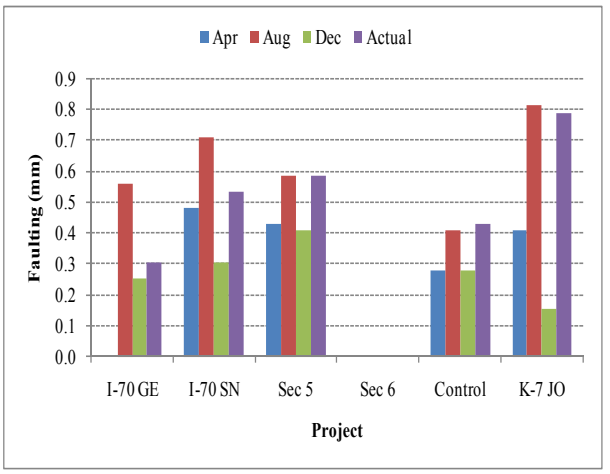
## Conclusions

Based on this study, the following conclusions can be made:

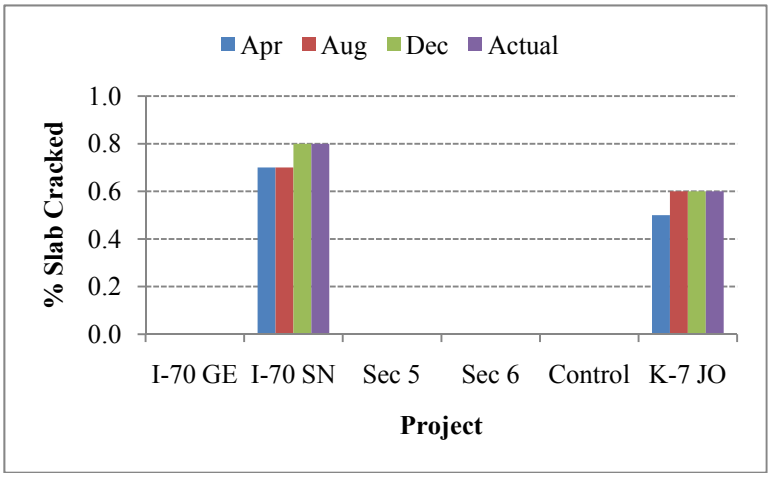
(1) Predicted JPCP roughness (IRI) and faulting by MEPDG are not very sensitive to the PCC strength. However, slab cracking is affected by strength, and cracking decreases with increasing strength.




(a) IRI




(b) Faulting




(c) % Slabs cracked

**Figure 3.** Predicted JPCP distresses for different pavement construction months

(2) In general, the shrinkage does not greatly affect predicted IRI. The higher shrinkage strain, the higher the faulting. Cracking appears to be fairly insensitive to the shrinkage strain.

(3) MEPDG predicted IRI and percent slabs cracked are fairly insensitive to the zero-stress temperature but faulting is severely affected. However, a JPCP section (SPS-2 Section 6) with widened lane and tied PCC shoulder did not show any faulting even for the highest zero-stress temperature.

(4) April and October appear to be the best months for JPCP construction (paving) in Kansas.

(5) In the MEPDG software, climatic database needs to be updated for accurate prediction of zero-stress temperature ($T_z$).

**Acknowledgement**

The project was sponsored by the Kansas Transportation and New Developments (K-TRAN) program. Assistance of Mr. Jose Villarreal, previously with Kansas State University, in this study is highly appreciated.

**References**

AASHTO. *Guide for Design of Pavement Structures.* American Association of State Highway and Transportation Official. Washington, D.C., 1986.

American Association of State Highway and Transportation Officials, *AASHTO Guide for Design of Pavement Structures*, Washington, DC, 1993.

NCHRP. *Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures.* Final Report for Project 1-37A, Part 1 & Part 3, Chapter 4. National Cooperative Highway Research Program, Transportation Research Board, National Research Council, Washington, D.C., March 2004.

Owusu-Antwi, E.B., Titus-Glover, L. and M.I. Darter. *Design and Construction of PCC Pavements. Volume I: Summary of Design Features and Construction Practices that Influence Performance of Pavements*. Report No. FHWA-RD-98-052, Federal Highway Administration, McLean, Virginia, April 1998.

# Impact of Non-Uniform Aircraft Tire Pressure on Airfield Pavement Responses

Hao Wang[1], Imad L. Al-Qadi[2]

**Abstract**: Due to the continual loading of flexible airport pavements with heavy and wide-bodied aircrafts, it is crucial to consider the impact of aircrafts' tire pressure on pavement damage. The measurements in previous literature have shown that the distribution of tire contact stresses under heavy aircraft is highly non-uniform. Therefore, this study developed a three-dimensional (3-D) finite element (FE) model to investigate the effect of non-uniform aircraft tire pressure on airfield pavement responses. An existing asphalt pavement section at the National Airport Pavement Test Facility (NAPTF) was used in the analysis. The FE model characterized the hot-mix asphalt (HMA) layer as a viscoelastic material and utilized an implicit dynamic analysis to predict the time-dependent pavement responses under moving aircraft tire loading. The tire loading was simulated as a continuously moving load having half-sinusoidal shape distribution along the contact length and non-uniform distribution along five ribs of the tire. The pavement responses (tensile and shear stresses/strains) under various tire pressure distributions were calculated and compared. The results showed that the non-uniform aircraft tire pressure increases the longitudinal tensile strains at the bottom of the HMA layer and the shear strains/stresses in the HMA layer, compared to the traditional uniform tire pressure distribution. The non-uniform contact stresses at high tire inflation pressure resulted in high shear stress near the pavement surface that could cause primary rutting or near-surface cracking; this shows the importance of establishing guidelines for asphalt mixtures stability when used for airfield pavements that support relatively heavy aircraft.

**Keywords**: Aircraft Tire Pressure, Non-Uniform Contact Pressure, Finite Element Model, Pavement Responses, Shear Stress

[1] Graduate Research Assistant, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 205 N Mathews MC-250, Urbana, IL 61801, E-mail: haowang4@illinois.edu

[2] Founder Professor of Engineering, Illinois Center for Transportation, Director, University of Illinois at Urbana-Champaign, 205 N Mathews MC-250, Urbana, IL 61801, E-mail: alqadi@illinois.edu

## Introduction

Traditional pavement analysis usually assumes that the contact stresses at the tire-pavement interface is equal to the inflation pressure and is uniformly distributed in a circular or rectangular contact area. However, this assumption is not consistent with the realistic tire loading condition because field measurements have clearly shown that the vertical contact pressure at the tire-pavement interface is non-uniformly distributed and the peak stress could be 1.5-2 times the tire inflation pressure for truck tires (Tielking and Abraham 1994; De Beer et al. 1997). Compared to truck tires, aircraft tires carry much greater wheel loads with higher inflation pressure. Under heavy aircraft weight, the bending stress in the bulged sidewall of the tire could cause high vertical contact stresses under the edge ribs; while the tire acts like a membrane in the center regions resulted in the vertical contact stresses approximately equal to tire pressure under the center ribs. Field measurements have found that the vertical contact stresses under the edge ribs of aircraft tires could be as high as 1.5-3 times the inflation pressure, depending on the load applied on the tire (Howell et al. 1986; Tielking 1989; Daugherty 2003; Rolland 2009).

Aircraft tire inflation pressure ranges from 1.2-1.5MPa depending on aircraft gross weights and landing gear configurations. The new generation of aircrafts, like Boeing 787 and Airbus 350/380, has tire pressure exceeding 1.5MPa. This creates a challenge for the traditional PCN rating that includes four pressure categories: W (no pressure limitation), X (1.5MPa limitation), Y (1.0MPa limitation), and Z (0.5MPa limitation) (ICAO, 1983). Roginski (2007) found that the rutting in the asphalt layer was the main failure mode under heavy aircraft loading, and high tire pressure had no adverse effect on flexible pavements that have stable asphalt layers and meet thickness requirement through a series of tire pressure tests performed at the FAA National Airport Pavement Test Facility (NAPTF) by Boeing Company. Fabre et al. (2010) emphasized the importance of considering non-uniform contact stress distribution in the analysis of high tire pressure effect on pavement responses in the upper layer.

Many studies have been conducted to analyze the effects of truck tire loading on highway pavement responses. It has been concluded that the truck tire configuration, load, and pressure could affect the potential of fatigue cracking and rutting of asphalt layer significantly (Al-Qadi et al. 2002; Elseifi et al. 2005; Machemehl et al. 2005). The critical pavement responses were found sensitive to the contact stress distributions at the tire-pavement interface; especially for the thin asphalt pavement or the pavement responses at near-surface (De Beer 2002; Al-Qadi and Yoo 2007; Yoo and Al-Qadi 2008). Compared to the uniform contact stress distribution, the non-uniform contact stress distribution under truck tires could induce greater or less pavement responses, depending on wheel load, inflation pressure, material stiffness, pavement thickness, and the type of response for comparison (Siddharthan et al. 2002; Wang and Al-Qadi 2009). However, limited research has been conducted to evaluate the effect of non-uniform aircraft tire pressure on airfield pavement responses.

## Objective and Scope

The objective of this study is to evaluate the effect of non-uniform aircraft tire pressure on airfield pavement responses using a developed three-dimensional (3-D) finite element (FE) model. An existing and already tested asphalt pavement section at the National Airport Pavement Test Facility (NAPTF) was used in the analysis. The FE model characterized the hot-mix asphalt (HMA) layer as a viscoelastic material and utilized an implicit dynamic analysis to predict the time-dependent pavement responses under moving aircraft tire loading. The tire loading was simulated as a continuously moving load with half-sinusoidal distribution along the contact length and non-uniform distribution along the five ribs of the tire. The pavement responses (tensile and shear stresses/strains) under various tire pressure distributions were calculated and compared.

**Pavement Structure and Material Characterization**

The modeled pavement structure is an existing section built in FAA NAPTF in the Boeing high tire pressure test program, as shown in Figure 1. The pavement section consists of a HMA wearing surface layer, a crushed aggregate base layer (P209), an Eco-concrete base layer (P306), an uncrushed aggregate subbase layer (P154), and a soil subgrade. The uncrushed aggregate (P154) was also used as the base layer in the analysis to evaluate the effect of base support on pavement responses. The dynamic modulus of the HMA and the elastic modulus of base/subbase layers and subgrade were provided by FAA NAPTF.

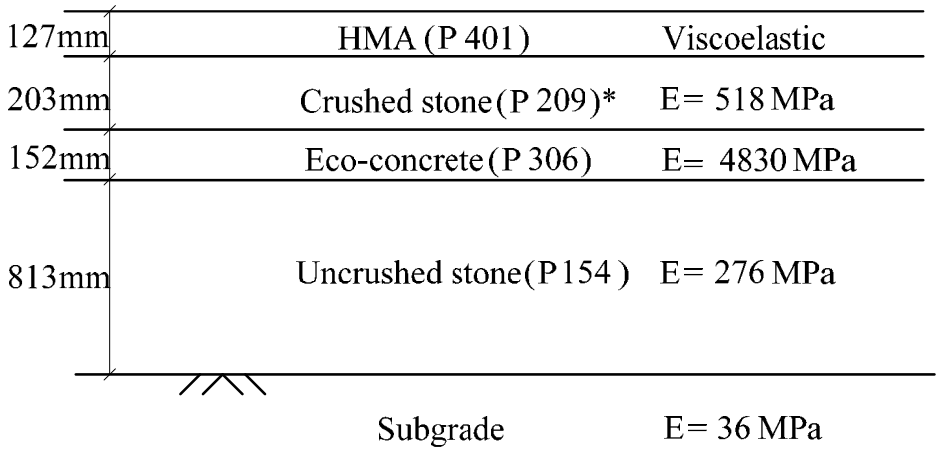



**Figure 1 Cross-section of the modeled pavement structure.**

Figure 2 shows the measured dynamic modulus and the fitted master curve using the sigmoid function at a reference temperature of 20 °C. As expected, under a constant loading frequency, the dynamic modulus decreases as the temperature increases; and under a constant testing temperature, the dynamic modulus increases as the frequency increases.

The relaxation modulus was inter-converted from the dynamic modulus using Equations 1 and 2 assuming that the linear viscoelascity of HMA was represented by a generalized Maxwell solid model. The relaxation modulus and relaxation times were determined by minimizing the sum of squares of the errors (Equation 3). The solver function in EXCEL was used to perform the curve fitting in an iterative

process. The bulk and shear relaxation moduli were calculated assuming a constant Poisson's ratio.

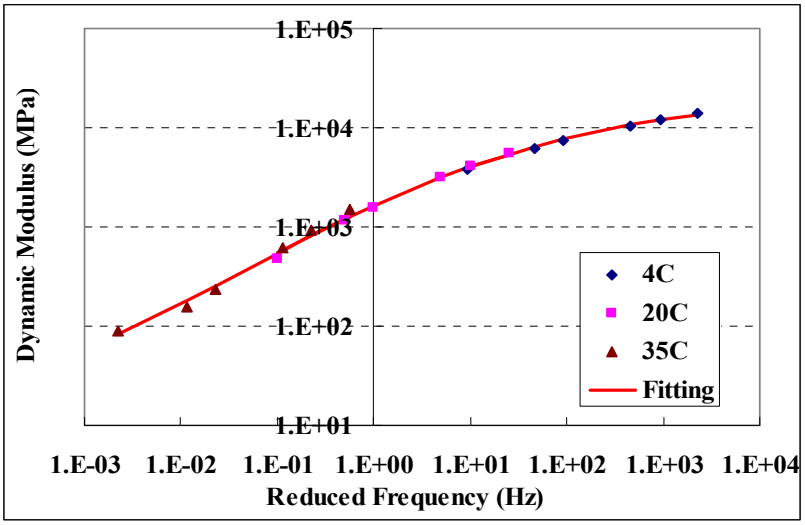



**Figure 2 Master curve of dynamic modulus**

$$E^{'}(\omega) = E_{\infty} + \sum_{i=1}^{n} \frac{\omega^2 \tau_i^2 E_i}{1 + \omega^2 \tau_i^2} \tag{1}$$

$$E^{''}(\omega) = \sum_{i=1}^{n} \frac{\omega \tau_i E_i}{1 + \omega^2 \tau_i^2} \tag{2}$$

$$\min \sum_{j=1}^{k} \left[ \left( \frac{E^{'}(\omega)_{calcualted}}{E^{'}(\omega)_{measured}} - 1 \right)^2 + \left( \frac{E^{''}(\omega)_{calculated}}{E^{''}(\omega)_{measured}} - 1 \right)^2 \right] \tag{3}$$

where,

$E^{'}(\omega)$ is real part of the dynamic modulus;

$E^{''}(\omega)$ is imaginary part of the dynamic modulus;

$E_{\infty}$ is equilibrium relaxation modulus at infinite time;

$\omega$ is angular frequency;

$E_i$, and $\tau_i$ are Prony series parameters for relaxation modulus;

$n$ is number of Maxwell elements; and

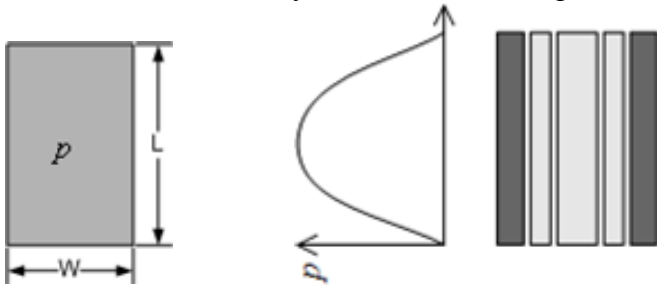
$k$ is number of data points from the measurements.

## Simulation of Tire Loading

Considering that the focus of this study is to evaluate the effect of tire pressure distribution on flexible pavement responses, a single tire loading was simulated. The applied tire load was 280kN, which is approximately equivalent to the single landing wheel load of Airbus A380 with the takeoff weight of 560 tons. Two pressure levels (1.45 and 1.69 MPa) were used to evaluate the effect of tire inflation pressure on

pavement responses due to the advent of new aircrafts exceeding the current ICAO tire pressure limit. For each tire inflation pressure, two different contact stress distributions were assumed in this study: one assumes the uniform vertical contact stress equal to the tire inflation pressure within the rectangular area (Figure 3a); and one assumes the non-uniform contact stress distribution in the tire imprint area with five ribs (Figure 3b).

For the non-uniform stress distribution, a half-sinusoidal pressure distribution was assumed along the contact length of each rib. The peak contact stresses under two edge ribs were assumed equal to 2.2 times the inflation pressure; while the peak contact stresses under three center ribs were assumed equal to 1.1 times the inflation pressure. Table 1 summarizes the contact stress distributions and contact areas at the inflation pressure of 1.45 MPa, respectively, for the uniform and non-uniform contact stress assumptions. As the inflation pressure increases, the contact length increases; while the contact width was assumed constant due to the relatively high lateral stiffness of the tire wall. The non-uniform contact stress assumption was based on the contact stress measurements under heavy aircraft tire load reported by Rolland (2009).




**(a)** **(b)**

**Figure 3 (a) Uniform and (b) non-uniform contact stress assumptions**

**Table 1 Contact Stress Distributions and Contact Areas**

| Contact pressure assumptions | | Contact Width (mm) | Contact Length (mm) | Peak pressure (MPa) |
|---|---|---|---|---|
| Uniform | | 320 | 560 | 1.45 |
| Non-uniform | Rib 1 | 70 | 560 | 3.19 |
| | Rib 2 | 45 | 560 | 1.62 |
| | Rib 3 | 90 | 560 | 1.62 |
| | Rib 4 | 45 | 560 | 1.62 |
| | Rib 5 | 70 | 560 | 3.19 |
| | Grooves | 10 | 560 | 0.00 |

**3D FE Model**

A 3-D FE model was built to predict pavement responses under moving tire loading with different contact stress distributions. The schematic illustration of the FE mesh is shown in Figure 4. In the FE model, the element thicknesses were selected at 10-20 mm for the HMA layers and 30-50 mm for the base/subbase layers and subgrade.

Infinite elements were used in the transverse and longitudinal boundaries of the model and at the bottom of subgrade to reduce the degrees of freedom at far field and absorb stress waves for dynamic analysis. The Coulomb friction model was used at the interface between each adjacent layer.

The loading area at a specific level of load and inflation pressure was considered by adjusting the number and dimensions of elements within the tire imprint area. The widths of the elements within the loading area were selected at 10-25 mm depending on the widths of tire ribs and grooves. The lengths of the elements were selected at 40 mm in the longitudinal (traffic) direction. The elements in the loading area were loaded with the uniform or non-uniform contact stress corresponding to their locations within the tire imprint area. For the non-uniform stress distribution, the loading amplitudes of contact stress continuously change at each step as the tire is moving. The dynamic transient analysis was used in this study considering the inertia associated with the moving load and the dependency of the material properties on the loading frequency. Additional details about the 3-D FE model, tire contact stress, moving load simulation and dynamic transient analysis can be found somewhere else (Al-Qadi et al. 2008; Yoo and Al-Qadi 2008).

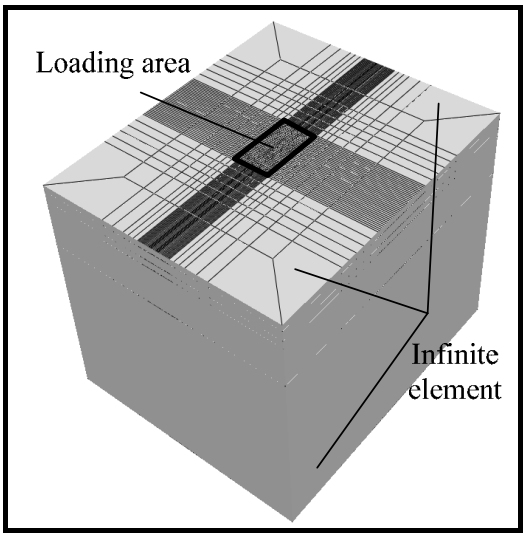



**Figure 4 Schematic illustration of FE mesh**

**Results and Analysis**

The critical responses considered in this study included the tensile strain at the bottom of the HMA layer that is responsible for the bottom-up fatigue cracking (for relatively thin HMA layers) and the shear strain/stress in the HMA layer that is responsible for the primary rutting and near-surface cracking. All the pavement responses were calculated at tire moving speed of 8 km/h and pavement temperature of 20 °C.

Table 2 compares the calculated pavement responses using different contact stress distributions. Compared to the uniform stress distribution, the non-uniform stress distribution induced 29-51% greater longitudinal tensile strain and 23-28% lower transverse tensile strain at the bottom of the HMA layer. This is probably because the

high contact stresses concentrated at the tire edge ribs could increase the longitudinal tensile strain while reduce the transverse tensile strain under the tire center. However, the uniform contact stress distribution was found to cause greater critical tensile strains than the non-uniform stress distribution. Therefore, using uniform stress distribution is a conservative approach for predicting fatigue cracking for the pavement structure. However, tensile strain related bottom-up cracking is not usually the primary distress in airfield pavements, unless weak bonding between asphalt layers exists or thin asphalt layers are used.

The results clearly show that the non-uniform contact stress distribution induces 20-36% greater shear strains and 52-61% greater shear stresses, compared to the uniform stress distribution. The changes are relatively significant. This suggests the necessity of using stable asphalt mixtures with high shear resistance that will support relatively heavy aircrafts. As expected, the relatively weaker base support (P154) causes greater pavement responses. Interestingly, the effect of non-uniform stress distribution on pavement responses became more significant as the base support was stronger.

**Table 2 Comparison of Pavement Responses Using Different Pressure Distributions**

| Base material | Crushed stone (P209) | | | Uncrushed stone (P154) | | |
|---|---|---|---|---|---|---|
| Contact pressure | Uniform | Non-uniform | Change | Uniform | Non-uniform | Change |
| Tire inflation pressure = 1.45 MPa | | | | | | |
| T. tensile strain (μ) | 493 | 360 | -27% | 624 | 482 | -23% |
| L. tensile strain (μ) | 251 | 380 | +51% | 336 | 479 | +43% |
| Shear strain (μ) | 541 | 735 | +36% | 680 | 815 | +20% |
| Shear stress (kPa) | 707 | 1136 | +61% | 818 | 1251 | +53% |
| Tire inflation pressure = 1.69 MPa | | | | | | |
| T. tensile strain (μ) | 522 | 377 | -28% | 660 | 492 | -25% |
| L. tensile strain (μ) | 317 | 443 | +40% | 412 | 532 | +29% |
| Shear strain (μ) | 597 | 811 | +36% | 714 | 893 | +25% |
| Shear stress (kPa) | 819 | 1307 | +60% | 942 | 1428 | +52% |

* T. is transverse and L. is longitudinal.

Table 3 compares the calculated pavement responses at different tire inflation pressures. It was concluded that, compared to the inflation pressure of 1.45 MPa, the higher inflation pressure of 1.69 MPa induced 11-26% greater longitudinal tensile strains and 2-6% greater transverse tensile strains at the bottom of HMA layer. The increase of inflation pressure could cause the decrease of contact length, while constant transverse distribution pattern was assumed for the contact stress under both

inflation pressures. In addition, the results show that the high inflation pressure induces 5-10% greater shear strains and 10-16% greater shear stresses. This indicates that increasing tire inflation pressure would reduce the service life of airfield pavements.

**Table 3 Comparison of Pavement Responses at Different Tire Inflation Pressures**

| Base material | Crushed stone (P209) | | | Uncrushed stone (P154) | | |
|---|---|---|---|---|---|---|
| Tire pressure (MPa) | 1.45 | 1.69 | Change | 1.45 | 1.69 | Change |
| Using uniform contact stress | | | | | | |
| T. tensile strain (μ) | 493 | 522 | +6% | 624 | 660 | +6% |
| L. tensile strain (μ) | 251 | 317 | +26% | 336 | 412 | +23% |
| Shear strain (μ) | 541 | 597 | +10% | 680 | 714 | +5% |
| Shear stress (kPa) | 707 | 819 | +16% | 818 | 942 | +15% |
| Using non-uniform contact stress | | | | | | |
| T. tensile strain (μ) | 360 | 377 | +5% | 482 | 492 | +2% |
| L. tensile strain (μ) | 380 | 443 | +17% | 479 | 532 | +11% |
| Shear strain (μ) | 735 | 811 | +10% | 815 | 893 | +10% |
| Shear stress (kPa) | 1136 | 1307 | +15% | 1251 | 1428 | +14% |

* T. is transverse and L. is longitudinal.

## Conclusions

The new generation of wide-body aircraft tends to have heavy gross weights and high tire inflation pressures. Therefore, understanding the realistic contact stress distribution at the aircraft tire-pavement interface under heavy load and high inflation pressure is very important for airfield pavement design.

In this study, the airfield pavement responses under different aircraft tire pressure distributions were calculated and compared using a 3-D FE model. The results showed that the non-uniform aircraft tire stress distribution increases the longitudinal tensile strains at the bottom of HMA layer and shear strains/stresses in the HMA layer as well, compared to the traditional uniform tire stress distribution. Especially, the non-uniform stress distribution at high tire inflation pressure resulted in high shear stress near the pavement surface that could cause primary rutting or near-surface cracking. Therefore, it is important to evaluate the stability of asphalt mixtures for airfield pavements that expected to support relatively heavy aircrafts.

## Acknowledgement

The authors would like to acknowledge Navneet Garg and Jeffrey Gagnon from National Airport Pavement Test Facility of Federal Aviation Administration for providing the pavement section structure and the material properties of each layer. The computing support provided by the National Center for Super Computing Applications (NCSA) at the University of Illinois at Urbana-Champaign is greatly appreciated.

**References:**

Al-Qadi, I.L., A. Loulizi, I. Janajreh, and T.E. Freeman (2002), "Pavement Response to Dual Tires and New Wide-Base Tires at Same Tire Pressure", *Transportation Research Record, No. 1816*, TRB, Washington, D.C., pp. 125-136.

Al-Qadi, I.L. and P.J. Yoo (2007), "Effect of Surface Tangential Contact Stress on Flexible Pavement Response", *Journal of Association of Asphalt Paving Technologists*, Vol. 76, pp. 663-692.

Al-Qadi, I.L., H. Wang, P.J. Yoo, and S.H. Dessouky (2008), "Dynamic Analysis and In-Situ Validation of Perpetual Pavement Response to Vehicular Loading", *Transportation Research Record No. 2087*, TRB, Washington, D.C., pp. 29-39.

Daugherty, R.H. (2003), *A Study of the Mechanical Properties of Modern Radial Aircraft Tires*, NASA Langley Research Center

De Beer, M.C., C. Fisher and F.J. Jooste (1997), "Determination of Pneumatic Tire pavement Interface Contact Stresses Under Moving Loads and Some Effects on Pavements with Thin Asphalt Surfacing Layers", *Proceedings of 8th International Conference on Asphalt Pavements* (Volume I), Seattle, Washington, pp 179-227.

De Beer, M.C., Fisher, and F.J. Jooste (2002), "Evaluation of Non-uniform Tire Contact Stresses on Thin Asphalt Pavements", *Proceedings of 9th International Conference on Asphalt Pavements* (ICAP 2002), Copenhagen, Denmark, 2002

Elseifi, M.A., I.L. Al-Qadi, P.Y. Yoo, and I. Janajreh (2005), "Quantification of Pavement Damage Caused by Dual and Wide-Base Tires", *Transportation Research Record No. 1940*, TRB, Washington, D.C., pp.125-135.

Fabre, C., J.M. Balay, and P. Lerat (2010), "Full-Scale Aircraft Tire Pressure Tests", *Proceedings of 2010 FAA Worldwide Airport Technology Transfer Conference* (in CD), Atlantic City, NJ

ICAO (1983), *Aerodrome Design Manual*, Part 3, Pavements, 2nd edition, Document 9157-AN/901, International Civil Aviation Organization

Machemehl, R.B., F. Wang, and J.A. Prozzi (2005), "Analytical Study of Effects of Truck Tire Pressure on Pavements with Measured Tire–Pavement Contact Stress

Data", *Transportation Research Record No. 1919*, TRB, Washington, D.C., pp. 111–120.

Roginski, M.J. (2007), "Effects of Aircraft Tire Pressures on Flexible Pavement", *Proceedings of the International Conference on Advanced Characterization of Pavement and Soil Engineering*, Athens, Greece, pp 1473-1481

Rolland, E. (2009), "Tire Pressure Test Effect on Pavement", *Airbus High Tire Pressure Workshop* (in CD), Toulouse, France

Siddharthan, R.V., N. Krishnamenon, M. El-Mously, and P.E. Sebaaly (2002), "Investigation of Tire Contact Stress Distributions on Pavement Response", *Journal of Transportation Engineering*, Vol. 128, No. 2, ASCE, pp. 136-144.

Tielking, J.T. (1989), *Tire/Pavement Pressure Distribution*, Final Report to Air Force Engineering and Service Center, Texas A&M University

Tielking, J.T., and M.A. Abraham (1994), "Measurement of Truck Tire Footprint Pressures", *Transportation Research Record No. 1435*, TRB, Washington, D.C., pp. 92–99.

Howell, W.E., S.E. Perez, W.A. Vogler (1986), "Aircraft Tire Footprint Forces", *ASTM Special Technical Publication (SPT) 929*, pp. 110-124

Yoo, P.J., and I.L. Al-Qadi (2008), "Truth and Myth of Fatigue Cracking Potential in Hot-Mix Asphalt: Numerical Analysis and Validation", *Journal of Association of Asphalt Paving Technologists*, Vol. 77, pp. 549-590.

Wang, H. and I.L. Al-Qadi (2009), "Combined Effect of Moving Wheel Loading and Three-Dimensional Contact Stresses on Perpetual Pavement Responses", *Transportation Research Record No. 2095*, TRB, Washington, D.C. pp. 53-61

# Adhesion Properties of Tar-Containing Antiskid Surface Layers on Runways in Airfield

Y. Xiao[1], A.A.A. Molenaar[1], M.F.C. van de Ven[1], S.P. Wu[2], M.R. Poot[1]

[1]Faculty of Civil Engineering & Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands; PH + 31 (0) 15 2781373; FAX + 31 (0) 15 2783443; email: Yue.Xiao@tudelft.nl
[1]Faculty of Civil Engineering & Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands; PH + 31 (0) 15 2784812; FAX + 31 (0) 15 273443; email: A.A.A.Molenaar@tudelft.nl
[1]Faculty of Civil Engineering & Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands; PH + 31 (0) 15 2782298; FAX + 31 (0)15 2783443; email: m.f.c.vandeven@tudelft.nl
[2]Key Laboratory of Silicate Materials Science and Engineering of the Ministry of Education, Wuhan University of Technology, Luoshi Road No. 122, Wuhan, China; PH + 86 (0) 27 87162595; FAX + 86 (0) 27 87162595; email: wusp@whut.edu.cn
[1]Faculty of Civil Engineering & Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands; PH + 31 (0) 15 2786536; FAX + 31 (0)15 2783443; email: m.r.poot @tudelft.nl

**Abstract:** Tar-containing antiskid surface layers are currently used on runways in airfields. They have good adhesion properties and excellent fuel resistance, which make them outstanding compared to other surface layers with bitumen as binder. However, tar-containing products have a high Polycyclic Aromatic Hydrocarbons (PAHs) content. They are toxic and not environmental-friendly. Tar-containing products are not allowed in the Netherlands since 20 years ago, with one temporary exception for antiskid surfaces on runways in airfields. After 2010, it will not be allowed to apply such antiskid layers anymore.

For this reason, alternatives to tar containing binders for antiskid runways are strongly required. Before looking into alternatives, the properties of tar-containing antiskid surface layer need to be evaluated and benchmarks should be developed.

This paper focuses on the mechanical properties of tar-containing antiskid surface layers. A special Leutner shear test and an advanced pull test were designed to test the shear and tensile strength in the interface between the tar containing antiskid layer and the asphalt mixture layer below. Test results are reported in this paper.

**Keywords:** Tar-containing; Antiskid surface layer; Leutner shear test; pull test; Dynamic Shear Rheometer

## 1. Introduction

Runway pavement surfaces are prepared and maintained to guarantee maximum friction for wheel braking. In order to keep high friction and minimize hydroplaning during heavy rain, antiskid runway surfaces are often used.

An antiskid runway surface is a high durable skid-resistant surface with a thickness of 3-5 mm. Generally a coal tar containing binder is used for the antiskid runway surface. Tar-containing antiskid runways possess excellent adhesion property and fuel resistance. However, tar is toxic because it has a high Polycyclic Aromatic Hydrocarbons (or PAHs) content. As a consequence tar was forbidden in every application except for airport runways in the Netherlands. Anyway, it is supposed to be not allowed anymore after 2010 [Mahler et al 2005; Van Metre 2005; Xiao 2010].

Therefore, there is an urgent need for finding alternatives to tar-containing antiskid layers. However one needs to characterize good performing tar-containing products first of all in order to develop a benchmark to which other products can be compared. The most important advantage of antiskid layers is their excellent adhesion characteristics. The strong adhesive strength can avoid raveling of the runway surface which may cause accidents. Good adhesion between the antiskid layer and the asphalt mixture gives good performance under high vehicle speed during taking off and landing.

Antiskid layers have to perform well under a wide range of environmental conditions. The outdoor temperature in the Netherlands varies normally between -5 to 25 °C. In this temperature range, the pavement surface could reach a temperature as high as 50 °C in the summer time, and drop to as low as -10 °C during a cold winter. Therefore, it is also necessary to analyze the temperature sensitivity of the adhesion properties.

Van Leest has reported about some tests that were performed on antiskid materials for runways in the past [van Leest 2005]. A simple pull test was used to measure the bonding strength between the antiskid layer and asphalt mixture layer in the field. Other test methods like the Torque Bond test, which was originally developed in Sweden for the in-situ assessment of bond conditions, has been adopted in the UK as part of the approval system for thin surfacing system. Furthermore the UTEP Pull-Off test was developed at the University of Texas at El Paso (UTEP). This test was widely used for evaluating the adhesion properties of interlayers [Tashman 2006; Deysarkar 2004]. But these bonding tests have no temperature and force loading speed control system.

The Leutner shear test was developed in Germany in the late 1970s as a simple method for evaluating the shear strength at the interface between two asphalt mixture layers [Collop 2003; Collop 2009; Xiao et al 2010]. However, *an* antiskid layer is a thin surface layer and does not fit in this standard shear test setup.

This paper includes two self designed test methods for analyzing the adhesion properties of antiskid layer mixtures. A special Leutner shear test and an advanced pull test were designed to test the shear and tensile strength of the interface between tar containing antiskid layer and the asphalt mixture layer below.

## 2. Materials and Test Methods

### 2.1 Materials

Cylindrical cores were collected from six airports, with antiskid layers of different ages in the Netherlands. These six airports are numbered from No.1 to No.6 in the research. No. 1 Airport has the newest antiskid layer while No.6 has the oldest. These cores contain a thin antiskid layer on the surface. These antiskid layers were constructed by the contractor Possehl Spezialbau GMBH. The aggregate for the antiskid layer is a special basalt 1.5 - 3.5 mm that comes from one certain place in Germany. The binder is a tar containing emulsion made in Germany.

Figure 1 shows Nano CT scan pictures of an antiskid layer from No. 4 Airport. These photos indicate that the antiskid layer has a large macrotexture and high void content.

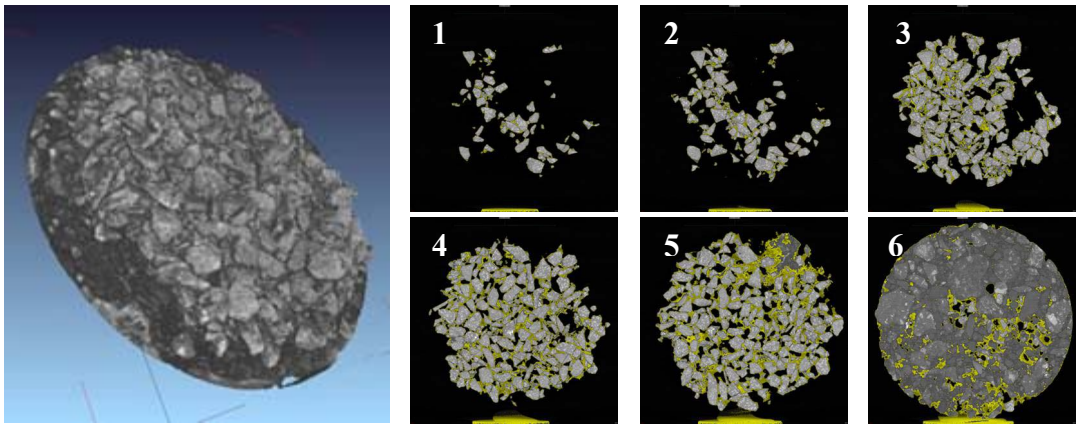



Figure 1 Nano CT scan pictures for antiskid layer

### 2.2 Leutner Shear Test

The Leutner shear test was developed in Germany in the late 1970s as a simple method of undertaking a direct shear test on the interface between two asphalt layers. The test is performed on 150 or 100 mm diameter cores with at least two layers taken either from a pavement or prepared in the laboratory. In this research, the antiskid layer has a thickness of approximately 5 mm. It is too thin to test the shear strength in the interface between antiskid layer and asphalt mixture layer by directly using the shear test setup. Therefore, *a* strong glue and *a* steel cylinder are used to modify this test. X6o glue is *a* 2-component fast curing glue, consisting of a liquid (component B) and a powder

(component A). It cures fast and can get high adhesion strength after cured at room temperature in dry conditions.

Figure 2 shows the preparation process of the shear test samples. First, the antiskid layer surface was carefully cleaned and dried. Then a steel cylinder with 100 mm diameter and 40 mm thickness was glued onto the antiskid surface with X60 glue. After the glue had reached full strength, samples were placed into a 20 °C temperature control cabinet for at least 3 hours before starting the test.

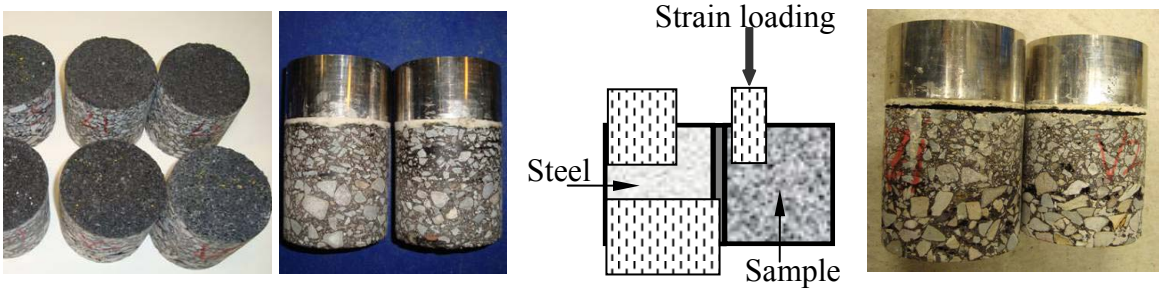



Figure 2 sample preparation and test setup for Leutner shear test

During the test, a constant vertical displacement of 50 mm/min is applied across the interface. The resulting shear force is measured. Test temperature is 20 °C. For each airport 2 samples were tested.

### 2.3 Pull Test

The direct tensile strength of the adhesive zone between the antiskid layer and the asphalt mixture layer was analyzed with a pull test.

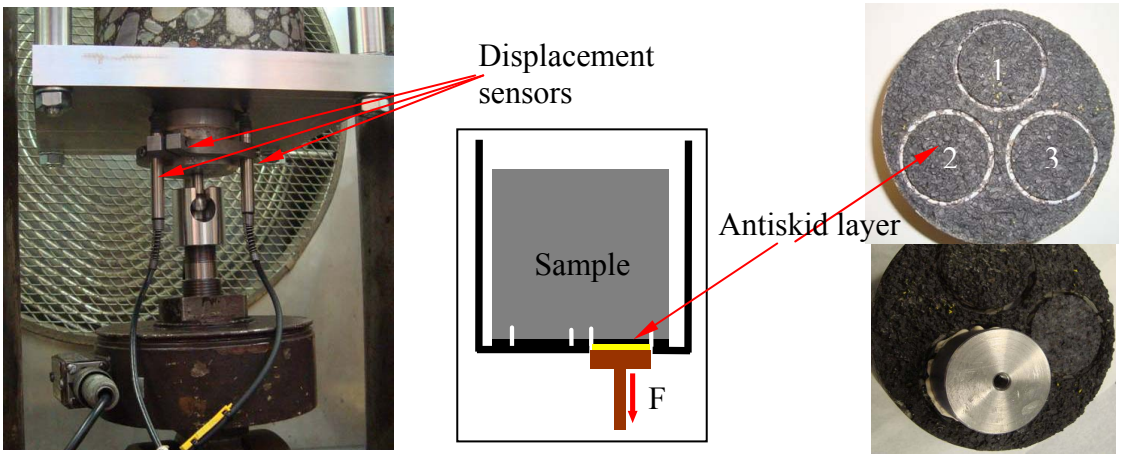



Figure 3 sample preparation and test setup for pull test

Cylindrical cuts with 50 mm diameter were cored on the samples to a depth of 10 mm so that the drill did surely pass the antiskid layer. A steel plate was then glued to the dry and clean surface with X60 glue. After the glue got its full strength, the antiskid layer was pulled off and the tensile force was measured. Figure 3 shows the sample preparation and setup for the pull test. Three pull tests can be done on the surface of one core.

Three test temperatures, 0 °C, 10 °C and 20 °C, with the same loading speed of 0.025 $N/mm^2s$ were evaluated. Three tests under different temperatures were done on one core. Three cores were used for parallel tests.

## 3. Test results

In this section, results from the Leutner shear test and the pull test will be discussed. Based on the sample preparation in this research, several fracture modes can be considered: A) in the asphalt mixture; B) in the antiskid layer mixture; C) at the interface; D) mix fracture in the asphalt mixture and at interface; E) mix fracture in the asphalt mixture, at interface and in the antiskid layer mixture; F) mix fracture at the interface and in the antiskid layer mixture. Due to the high enough adhesion strength of glue after curing, no debonding between the loading plate and the antiskid surface was occurred.

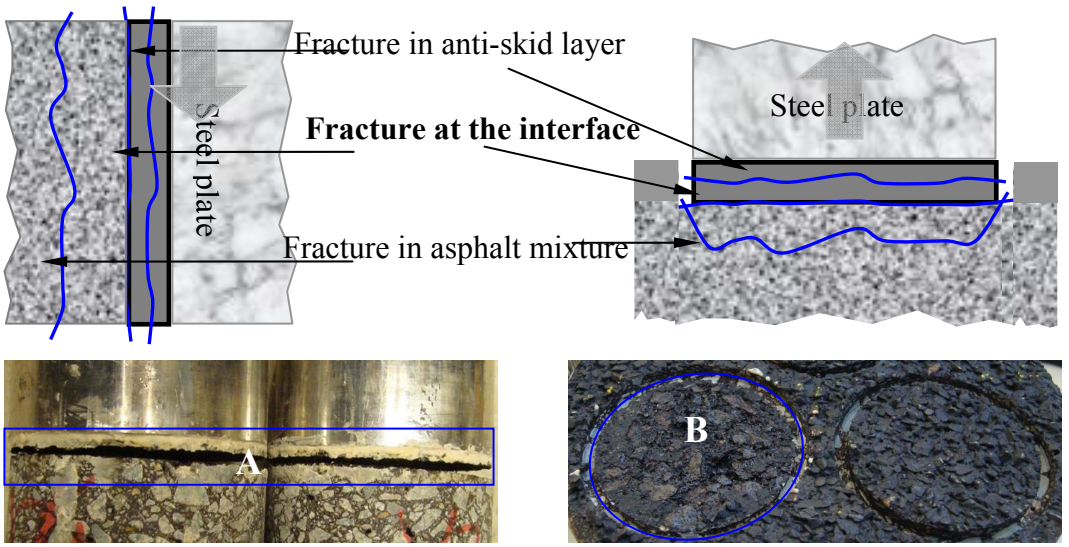



Figure 4 three possible fracture modes and practical fracture model during adhesion evaluations

As A and B in Figure 4 show, all Leutner shear tests and 78.3 % of the pull test samples failed according to mode C type fracture. This indicates that these self designed methods are workable methods for evaluating the adhesion properties between a thin layer and an asphalt mixture layer. No failure in asphalt mixture happened during all the fracture tests.

It shows that the strength at the interface is not higher than the strength of the asphalt mixture. This is because the asphalt mixture has larger size aggregates inside, which can generate an interlock structure. An interlock aggregate structure can make a great contribution to both shear and tensile strength.

### 3.1 Leutner Shear Test

Table 1 shows the test results of the Leutner shear test. No relation between the age and the shear strength can be found according to these results. The No. 6 Airport samples have the second highest strength although the airport has been in service for 20 years. The shear strength of No. 1 Airport samples, which have the newest antiskid layer, is the lowest.

Table 1 shear strength of the interface

| Airports and construction year of antiskid layers | Maximum tensile force/ Kn | | | Shear Strength/ MPa |
|---|---|---|---|---|
| | Test 1 | Test 2 | Average | |
| No. 1 airport / 2009 | 8.1428 | 9.3696 | 8.7562 | 1.159 |
| No. 2 airport / 2008 | 11.8479 | 10.798 | 11.32295 | 1.478 |
| No. 3 airport / 2007 | 11.9761 | 13.9416 | 12.95885 | 1.674 |
| No. 4 airport / 2005 | 9.5284 | 9.2537 | 9.39105 | 1.219 |
| No. 5 airport / 2004 | 16.0841 | 11.2374 | 13.66075 | 1.75 |
| No. 6 airport / 1989 | 12.7207 | 13.1968 | 12.95875 | 1.691 |

Figure 5 shows the shear strength and displacement as obtained from the Leutner shear test. Samples from No. 1 Airport have the lowest shear strength and the second highest displacement. Samples from No. 4 Airport have the highest displacement and *the second* lowest shear strength. The low shear strength of the No. 4 Airport samples is somewhat unexpected because one might expect that ageing would increase the strength as well as the stiffness. The low results of the No. 4 Airport might be caused by a different texture of the asphalt surface since the texture can *a*lso have a significant influence on the interface shear strength.

On some cores from No. 1 Airport, the Leutner shear test was also done for comparison reasons on the interface between asphalt mixture layers and through the asphalt mixture. Table 2 shows the results. These results indicate that the shear strength at the interface between antiskid layer and asphalt mixture layer is higher than the shear strength at the interface between asphalt mixture layers. It is lower than the shear strength in the asphalt mixture.

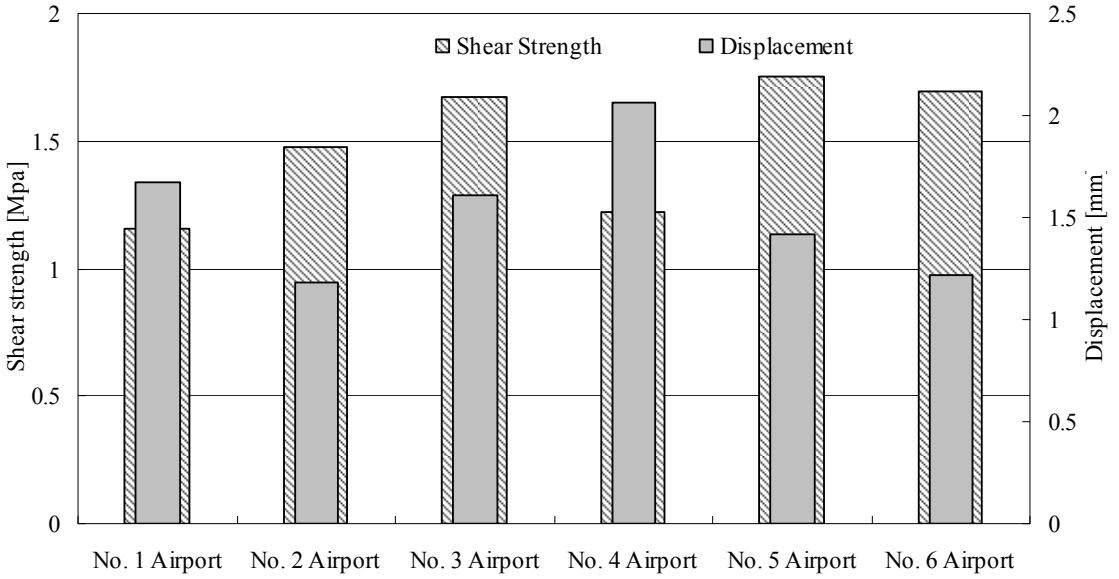



Figure 5 shear strength and displacement for shear test

Table 2 shear strength of the interface

| No. 1 Airport / 2009 | Shear Strength/ MPa | | | |
|---|---|---|---|---|
| | In mixture | Interface | In mixture | Interface |
| Test 1 | 2.416 | 1.196 | 2.004 | 1.120 |
| Test 2 | 2.446 | 0.813 | 1.683 | -- |
| Test 3 | 2.465 | -- | 2.065 | -- |
| Test 4 | 2.307 | -- | -- | -- |
| Average | 2.408 | 1.005 | 1.917 | 1.120 |

### 3.2 Pull Test

Figure 6 illustrates the *Direct Tensile Strength* (DTS) at three different temperatures on the some cores from No. 2 Airport and No. 4 Airport. The DTS *decreases* with increasing of test temperature. Figure 6 also shows that the variation between the three test results is limited which indicates that the repeatability of the DTS is quite good. This indicates that this advanced pull test works quite well for determining the tensile strength of the interface between a thin antiskid surface layer and asphalt mixture layer.

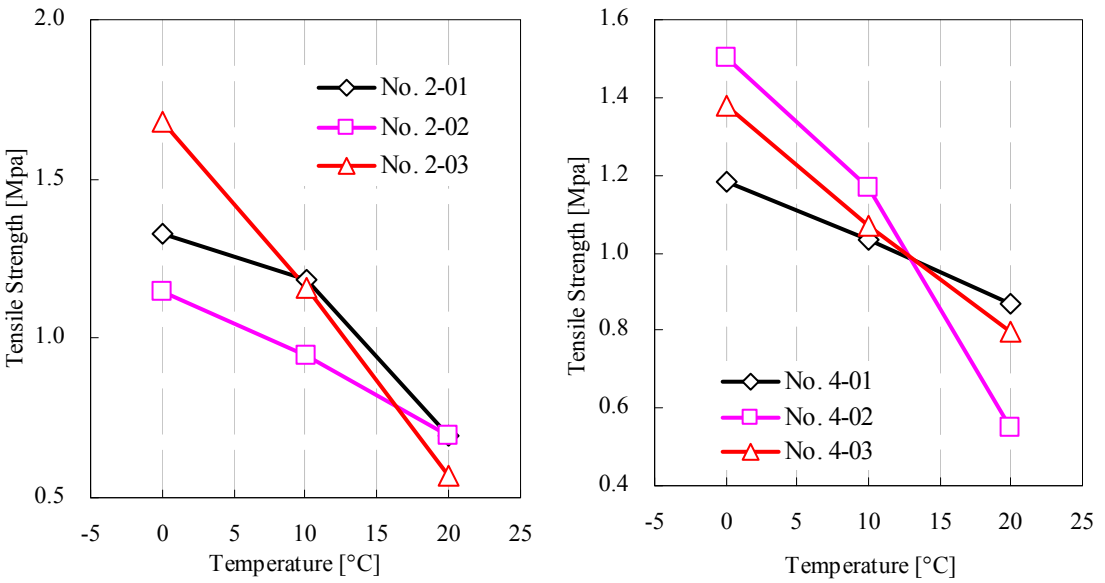



Figure 6 DTS at different temperatures on the same cores

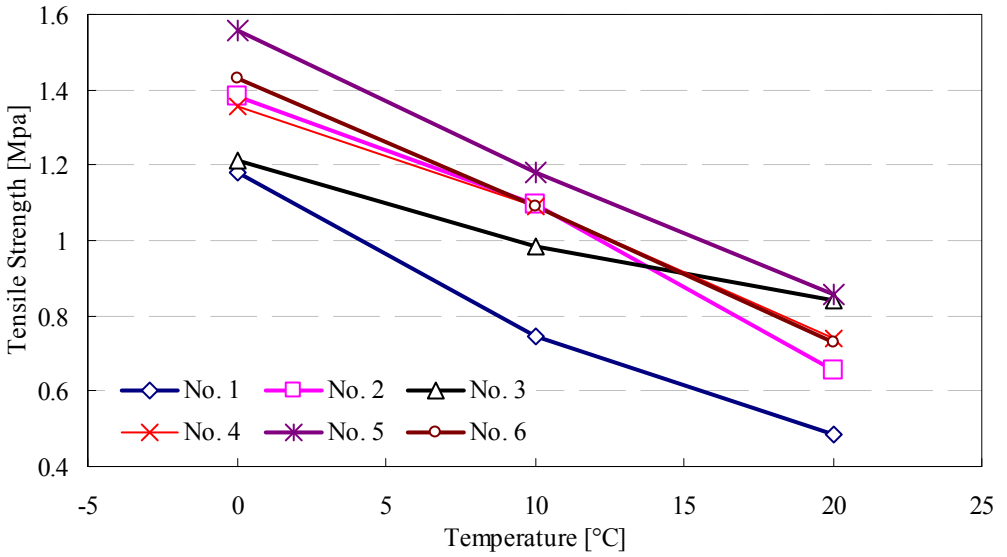



Figure 7 Average DTS at three evaluated temperatures

During the pull test, the samples from No. 1 Airport that were tested at 10 °C and 20 °C were all breaking in type B or type F fracture modes. This means that at 10 °C and 20 °C, the DTS of the No.1 Airport samples at the interface is higher than in the antiskid layer mixture. Type F fracture mode also happened on samples from No. 2 Airport at 10 °C.

Both in No. 2 and No. 4 Airport, there was one sample tested at 20 °C that showed type D fracture mode. This illustrates that the DTS in the interface can even be higher than the DTS of the asphalt mixture.

Figure 7 presents the average DTS at the interface between the antiskid layer and asphalt mixture layer at three evaluated temperatures. It illustrates the newest antiskid layer, antiskid layer in No. 1 Airport, has the lowest tensile strength. No. 5 and No. 6 Airports, the two oldest airports, have the highest tensile strength. The other three have nearly the same DTS. It is believed this is due to the influence of ageing. The tar-containing binder got harder and generates higher tensile strength with ageing. The surface characteristics of asphalt mixture layer may also have a significant influence on the interface tensile strength. It is remarkable to note however that the DTS results obtained on the No. 4 Airport samples are in line with the expectations while the Leutner shear test tesults were not.

## 4. Conclusions

Based on the experimental test results obtained on the antiskid layers, the following conclusions could be drawn.

1. The self designed pull test and shear test are workable *tools* for evaluating the direct shear strength and direct tensile strength between a thin antiskid layer and an asphalt mixture layer.

2. The shear strength at the interface between antiskid layer and asphalt mixture layer is higher than the shear strength at the interface between asphalt mixture layers. It is lower than shear strength in the asphalt mixture. They are between 1.2 and 1.7 MPa when tested at 20 °C and 50 mm/min.

3. Most of the pull tests resulted in failure at the interface. The average tensile strength at the interface at 20 °C, 10 °C and 0 °C are 0.7175 MPa, 1.0305 MPa and 1.3537 MPa, respectively.

4. The oldest airport has the highest shear strength and tensile strength. This indicates that ageing of the tar-containing binder has occured, which can increase its stiffness and strength.

## Acknowledgement

The research has been done under the supervision of CROW 2nd study group Alternatives for Tar-containing Antiskid on runways (ATA). The authors would like to thank the study group for allowing publication of the results. The scholarship from the China Scholarship Council is acknowledged. The test support of J. Moraal, J.W. Bientjes and D.C. Doedens are also appreciated.

**References**

Collop A.C., Thom N.H. and Sangiorgi C., (2003). "Assessment of bond condition using the Leutner shear test." *ICE J Transp 156 (TR4),* 211–217.

Collop A.C., Sutanto M.H., Airey G.D., Elliott R.C., (2009). "Shear bond strength between asphalt layers for laboratory prepared samples and field cores." *Constr Building Mater 23(6),* 2251 – 2258.

Deysarkar, I., (2004). "Test Set-Up to Determine Quality of Tack Coat." *Master's Thesis, Department of Civil Engineering, the University of Texas at El Paso.*

Mahler B.J., Van Metre, P.C., et al, (2005). "Parking lot sealcoat: an unrecognized source of urban polycyclic aromatic hydrocarbons." *Environ sci Technol, 39(15),* 5560-5566.

Tashman L., Nam K., and P. Tom, (2006). "Evaluation of the Influence of Tack Coat Construction Factors on the Bond Strength between Pavement Layers." *Washington State University, Report No.: WA-RD 645.1*

van Leest A.J. and Gaarkeuken G., (2005). The F.O.D. resistance of surface layers on airfields in the Netherlands; in situ and laboratory testing. *2005 European Airport Pavement Workshop.*

Van Metre P.C. and Mahler B.J., (2005). "Trends in hydrophobic organic contaminants in urban and reference lake sediments across the United States, 1970-2001." *Environ Sci Technol, 39(15),* 5567-5574.

Xiao Y., (2010). "Literature Review on Possible Alternatives to Tar for Antiskid Layers." *Delft University of Technology, Road and Railway engineering section, Report No. 7-10-185-1,* the Netherlands.

Xiao Y., M.F.C. van de Ven, et al, (2010). "Report on an Assessment Protocol for Tar-containing Antiskid Layers for Runways." *Delft University of Technology, Road and Railway engineering section, Report No. 7-10-185-2,* the Netherlands.

# Restoration of Data Collection System At Denver International Airport

Navneet Garg[1], David Brill[2], Chuck Teubert[3], Thomas Weinmann[4], Keith Johnson[5]

[1] Airport Technology R&D Branch, FAA's William J. Hughes Technical Center, Atlantic City Intl. Airport, NJ 08405; Phone: (609)485-4483; Email: navneet.garg@faa.gov.
[2] Airport Technology R&D Branch, FAA's William J. Hughes Technical Center, Atlantic City Intl. Airport, NJ 08405; Phone: (609)485-5198; Email: david.brill@faa.gov.
[3] SRA International Inc., 1201 New Road, Suite #242, Linwood, NJ 08221; Phone: (609)601-6800; Email: chuck.teubert@sra.com.
[4] Applied Geomechanics, 1390 Busch Parkway, Buffalo Grove, IL 60089; Phone: (847)850-5051; Email: Thomas.Weinmann@geomechanics.com.
[5] Planning & Development Division, 8500 Peña Boulevard, Denver, CO 80249; Phone: (303)342-2736; Email: keith.johnson@flydenver.com.

## *Abstract*

The Federal Aviation Administration initiated a runway instrumentation project in 1992 by installing 460 static and dynamic sensors in sixteen slabs in the take-off area of Runway 34R-16L at Denver International Airport. Static sensors such as embedded temperature and moisture gauges and dynamic sensors such as H-bar strain gauges, LVDTs, and position strain gauges were installed. The investigation was focused on measuring pavement response to aircraft loading and environmental conditions. Currently, the data collection system is severely damaged due to water infiltration and is not operational. Data was collected from 1995 through 1999 which included real-time pavement strains and deflections resulting from aircraft traffic. The slab strain and vertical deflection data was analyzed and compared to three-dimensional finite element model predictions. Overall, good agreement between measured and predicted strains and vertical deflections was observed. This information was used for, among other things, validating and calibrating the structural model in FAARFIELD.

An important parameter missing in FAARFIELD is the effect of curling on pavement thickness design. Severe curling can cause premature cracking and incur high maintenance costs. The primary objective of this project is to restore the data collection system to at least measure slab deflections and strains. The future emphasis will be on characterizing slab curling under varying environmental conditions and incorporating it into FAARFIELD. Efforts are underway to restore the data collection system. The entire project is expected to be completed by September 2010. Once resurrected, the instrumentation project will provide valuable slab curling data for updating structural models in FAARFIELD. This paper describes in detail the restoration of the Denver International Airport instrumentation project and the ways in which the FAA plans to make use of the collected data.

## *Introduction*

In mechanistic-empirical design procedures, pavement responses such as stresses and strains are related to pavement life through the use of transfer functions. The failure models used in the Federal Aviation Administration (FAA) pavement thickness design procedures (AC 150/5320-6E) are developed based on full-scale test results. Pavement instrumentation data helps in a better understanding of pavement system responses and can also be used for the validation/calibration of analytical response prediction models.

The instrumented section of runway at the Denver International Airport met the need to obtain response data for rigid pavement for existing and new landing gear configurations beyond the Boeing 747 such as the Boeing 777. In 1992, the Federal Aviation Administration (FAA) (with the support of the DIA airport authority, the City and County of Denver, and the FAA Denver Airport District Office) initiated a major research effort to study the in-situ response and performance of Portland Cement Concrete (PCC) pavements. The U.S. Army Corps of Engineers Waterways Experiment Station, under a contract with FAA, installed a comprehensive instrumentation system within the pavement structure in the take-off area of Runway 34R-16L. Instrumentation data was collected from 1995 to 1999. The details about the instrumentation database and data processing can be found in Dong (1997) and Lee (1997). Data collected during this time period included real-time pavement responses to actual aircraft traffic, environmental parameters and weather conditions. Pavement inspections, including falling-weight-deflectometer (FWD) tests, continue to be conducted periodically and added to the database. Approximately $3 million were spent on DIA instrumentation project till 2009. The collected data were used for comparing measured responses with existing theoretical models and for validating and calibrating the structural model in FAARFIELD (new FAA pavement thickness design software).

Currently, the data collection system is severely damaged due to water infilteration and is not operational. This paper describes the DIA instrumentation project, how the data collected so far has been used, need for data for future use, the ways in which the FAA plans to make use of the collected data, the current status of instrumentation and data collection system, and the plan for restoration of the DIA instrumentation project.

## *Denver International Airport (DIA) Pavement Instrumentation Project*

During the construction of Runway 34R-16L, 16 slabs in the take-off area were instrumented. Figure 1 shows the site layout. A total of 460 sensors were installed in the pavement structure. The pavement structure consists of 17 inches of PCC slab, on top of 8 inches of econocrete base, over 12 inches of lime-stabilized subgrade, resting on the natural subgrade. The instrumentation consists of two main categories of on-line sensors – “dynamic” and “static” sensors. The dynamic sensors are used to measure pavement response to aircraft traffic and the data are collected at frequencies between 100 and 1,000 Hertz. The static sensors measure pavement response to changes of environmental conditions such as variations of temperature and moisture. Data from static sensors are collected on an hourly

basis. In addition to on-line sensors, several off-line sensors have been installed. Data from these sensors are collected during field surveys.

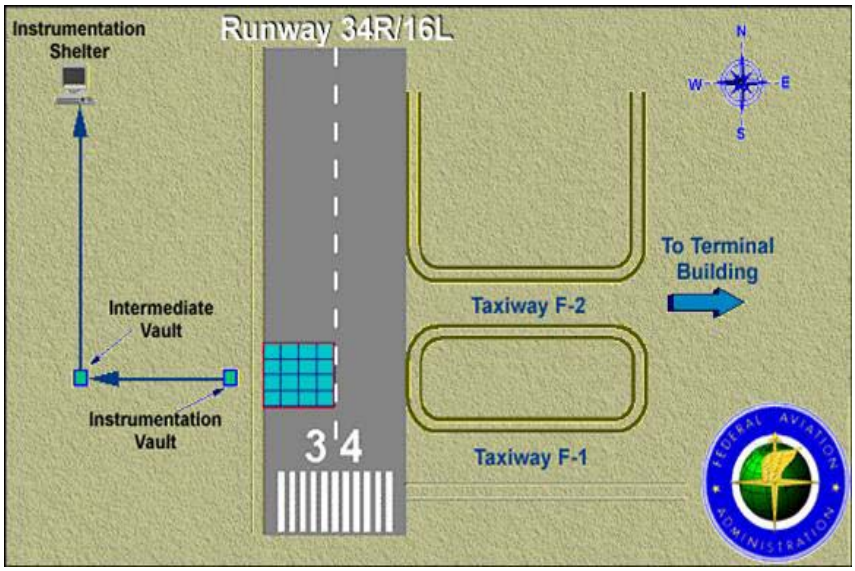



Figure 1. Layout of Instrumented Test Site at DIA

The sensors used in the project are listed below in Table 1.

Table 1. Responses Measured and Sensors Used at DIA

| **Response Type** | **Gage Type** |
|---|---|
| Strains | Carlson Strain Gages<br>Carlson Joint Strain Gages<br>H-Bar Strain Gages |
| Vertical Displacements | Linear Variable Displacement Transducers |
| Aircraft position, speed and acceleration | Position Strain Gages<br>Infrared Sensors |
| Environmental data in the pavement structure | Temperature Gages<br>TDR Gages<br>Resistivity Gages |
| Manual measurements of pavement surface elevations | Embedded Brass Pins |

An extensive data acquisition system was put in place to control these sensors and perform data collection. Remote access was established in 1994 enabling the FAA with on-line access to the data acquisition system. A video camera (Figure 2) was also installed to provide live images.

The collected data were processed and converted into engineering units for use by pavement researchers. The database was developed using Microsoft SQL Server7 and is managed by the FAA Airport Technology R&D Branch. The data can be accessed using internet and connecting to the DIA database server at the following link: http://www.airporttech.tc.faa.gov/denver/dbbrws/index.asp.

Figure 2. FAA Instrumentation Shelter with Video Camera at DIA

***DIA Database & FAARFIELD***

Data collected from the DIA Instrumented Pavement Project has been used by the FAA to develop and improve the FAARFIELD computer-based pavement thickness design procedure. Specific uses of the data to date fall into two general categories:

1. Field verification of the three-dimensional finite element (3D-FE) structural analysis model used in the rigid pavement design procedure.
2. Validation of design assumptions, such as joint load transfer, airplane wander, and the effectiveness of the slab-base interface for different gear positions.

The first use, field verification of the 3D-FE model, is discussed in detail in Brill (2000). Mean peak strain and deflection data collected over a 20-month period in 1996 and 1997 were compared to predictions of the static 3D-FE model for corresponding gear loads. A statistical analysis of the sensor data was performed to identify significant seasonal variations in the peak responses. It was found that while peak deflection values at slab edges were significantly affected by seasonal variations in temperature, there was relatively little variation in either the peak strain response or the deflection measured at the interior of the slab. It was also found that "significant bonding does develop at the center of the slab, notwithstanding the presence of a bond breaker." Overall, good agreement was found between the average peak responses recorded under aircraft loads and the corresponding 3D-FE predictions, which supported the FAA's decision to incorporate a 3D-FE model into the new rigid pavement design procedure.

Dynamic DIA responses were also analyzed to gain insight into various design assumptions used in the FAA model. The FAA design procedure for rigid pavements assumes a uniform 25% reduction in free edge stress at the slab edge to account for load transfer. Statistical analysis of in-situ load transfer efficiency

(LTE) using paired deflection sensors at the DIA site showed that (a) the load transfer developed in "dummy" (undoweled contraction) joints is strongly dependent on the season and may be very low (less than 10%) during winter months, and (b) neither the 3D-FE computed values of LTE nor the corresponding measured values from DIA sensors were strongly influenced by the loading aircraft type. Other analysis suggested that the FAA's assumption of a sliding surface between the slab and the base is conservative; i.e., the most realistic model is intermediate between fully unbonded and fully bonded.

Additional analysis of the DIA sensor data was performed under FAA sponsorship by the FAA Center of Excellence for Airport Technology (CEAT) at the University of Illinois Urbana-Champaign. Rufino et al. (2001) analyzed data from the embedded position sensors to make statistical statements about the lateral wander of D, 2D and 3D aircraft at the DIA site. The CEAT reasearchers found relatively good agreement between their results and a previous study by HoSang (1975). Rufino and Roesler (2006) analyzed deflection and strain data from DIA sensors to characterize the slab-base interaction for various temperature and loading conditions. Among other results, their study confirmed that high slab-base contact friction or "bonding action" develops under load at the slab interior. This result was significant for FAARFIELD development since it mitigated the need to separately consider an interior load case for 3D gears.

Future updates to the FAARFIELD rigid pavement design procedure are expected to consider multiple-gear loads, as well as the effects of slab curling, slab size and joint stiffness on the computed design stress. FEAFAA, the FAA's 3D-FE analysis program, has an improved modeling capability that considers the shape of the curled slabs, as well as arbitrary placement of multiple gear loads relative to slabs. However, reliable field data will be essential for validating these complex models. Using the restored DIA data collection system, it is planned to match multiple dynamic responses from individual traffic events to FEAFAA model predictions, using inputs such as estimated slab curling and joint stiffness as model parameters. The goal is to develop a realistic range for these parameters that can be used in the next generation of FAARFIELD rigid pavement designs. Responses of interest are not limited to those directly under the wheel load (as in the previous studies) but also include responses at a distance from the wheel center that may contribute to top-down cracking of the slabs.

### *Current Status of DIA Project*

The data collection system is severely damaged due to water infiltration over the years and is not operational. Figures 3 to 5 show the condition of vault as documented during site visit in November 2009. The work performed on-site was in a confined space and safety of workers is of utmost importance. Working with DIA Safety Coordinator, a Hazard Analysis was performed and a Safety Plan was developed and included a Hazard Reduction Plan. Prior to each evening's work, the Safety Plan was reviewed with staff. As part of the safety plan, ventilation and air monitoring of the vault was continuous and working personnel were prevented from working more than 30 minutes continuously in the vault. Figure 3 shows the FAA instrumentation vault and preparation of vault for entry and inspection. Figures 4 and 5 show the vault interiors.

FAA Instrumentation Vault at DIA

Pumping Accumulated Water and Preparing Vault for Inspection

Figure 3. FAA Instrumentation Vault at DIA


Figure 4. Condition of FAA Instrumentation Vault at DIA

Once the water was pumped out from the vault, the vault has a strong sulfurous odor; all the contents were covered in the hard, black residue. Hydrogen sulfide and carbon monoxide levels were high while oxygen was low. Inside the vault, most of the wires appeared to be docked inside metal instrumentation cabinets. Most of the cabinet doors were closed.

Figure 5. Condition of Data Acquisition System at DIA

### *Plan for Restoration of DIA Instrumentation Project*

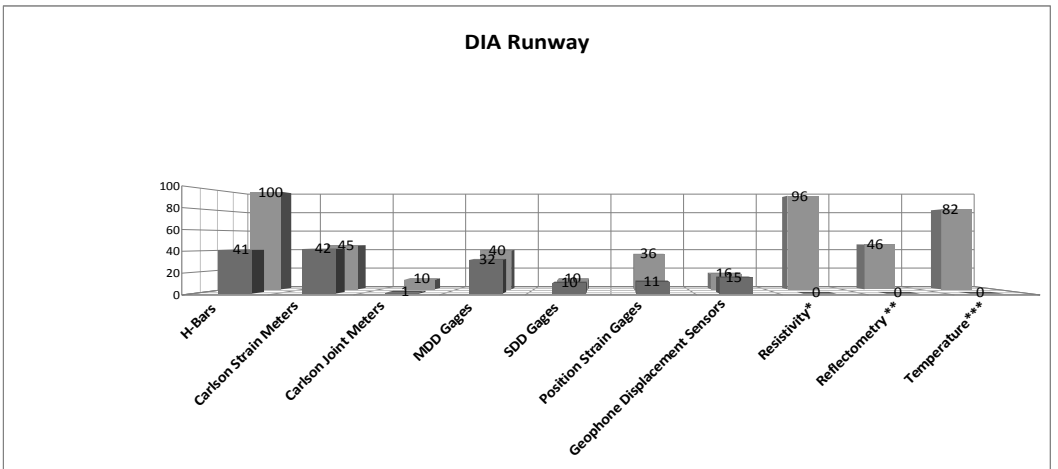
The goal of this project is to restore the data collection effort at DIA. To meet this goal, the objectives to be accomplished are to

- Evaluate the status of existing embedded pavement sensors to determine their operational capabilities; and
- If appropriate, redesign the data acquisition system and move from the underground vault to an above ground data collection facility.

The plan for restoration of DIA instrumentation project is under development. The plan will go forward only if a justifiable number of sensors are found in working condition and it is determined that meaningful data can be collected from these sensors. It consists of several steps which are described in this section.

Evaluation of Existing Sensors:
The first step, after the vault has been prepared for entry, was to document and label the sensors cables in the existing system. The field evaluation of the current status of sensors has been completed. The results are shown in Figure 6. The classification of "working gages" is based on current reading at a point in time. Information used in this determination includes actual reading output, comparison of current reading with past reading (if available), and current reading compared with expected range of operation output. This classification does not guarantee performance or calibration of response for these sensors.

System Relocation:
The initial phase of this task will be to coordinate with DIA Operations and move all sensor wiring from the vault and route to a location about 250-feet (or more) from the runway centerline, where a new data acquisition system will be installed.




| | H-Bars | Carlson Strain Meters | Carlson Joint Meters | MDD Gages | SDD Gages | Position Strain Gages | Geophone Displacement Sensors | Resistivity* | Reflectometry ** | Temperature *** |
|---|---|---|---|---|---|---|---|---|---|---|
| Working Gages | 41 | 42 | 1 | 32 | 10 | 11 | 15 | 0 | 0 | 0 |
| Total Gages | 100 | 45 | 10 | 40 | 10 | 36 | 16 | 96 | 46 | 82 |

Figure 6. Current Sensor Status at DIA Instrumentation Project

Figure 7 shows the existing vault location and the proposed vault location.

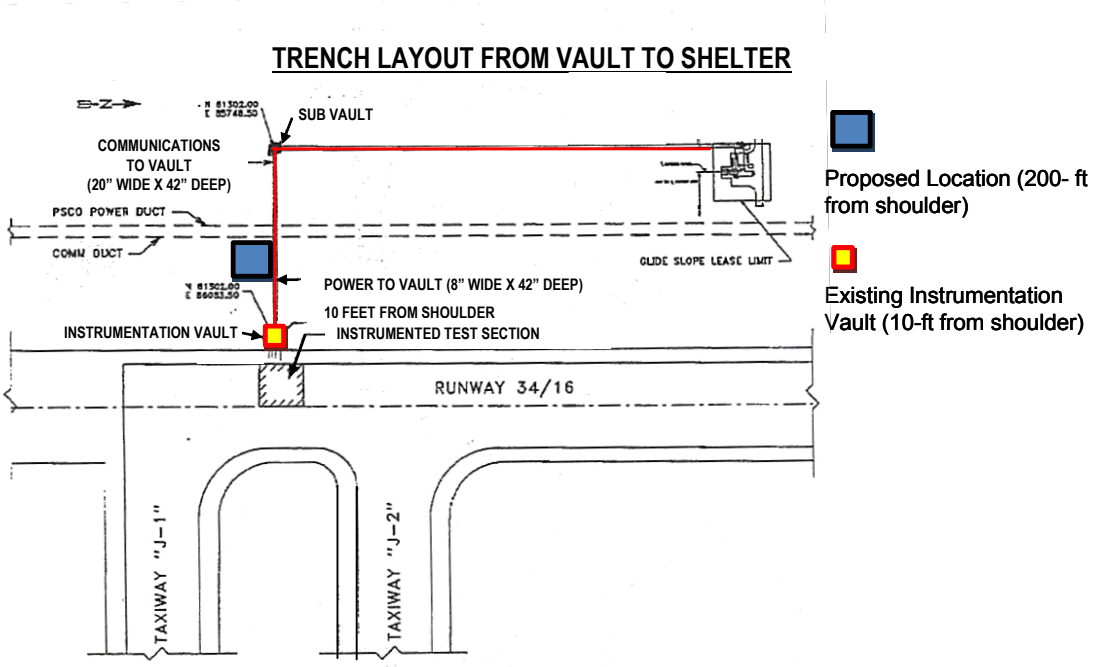



Figure 7. Vault Location (existing and proposed) at DIA Instrumentation Project

The new system will be relocated beyond the limits for height constraints of structures for DIA (or allow for minimal height of system cabinet). This will require adding sensor lead wire lengths to existing working sensors perpendicular to runway approximately 250 feet. The intent would be to excavate a trench where the existing power to the vault was provided. At the new location, 6 feet by 6 feet concrete pad would be placed with a new power line. This location would house the new data acquisition system in an enclosure no higher than 18 inches above the concrete pad.

In the preliminary stages, a simple data acquisition system will be designed to read all the working gages (see Figure 6) statically (1 Hz). Once it has be ascertained that significant number of sensors are working properly (by analyzing the statically collected data) and can provide valuable data which could be used in the development of new improved structural models for FAARFIELD, different options will be studied to trigger data collection from sensors. Some of the options for triggering data collection under consideration are:

1. software programmable time intervals;
2. remote trigger based on camera view of real-time activity; and
3. automated trigger based on airplane position.

This is an ongoing project and information will be updated when available.

***Lessons Learned***

As described earlier, the data collection system was housed in an underground vault adjacent to the runway. Aside from the instrumentation control equipment the vault also housed two standard PC computers, mainly for data storage and a

modem that could be remotely activated via the ISDN telephone line that it was connected to. This vault was "sealed" with a cover that required tools to remove. The vault was not environmentally controlled or environmentally monitored. From the beginning it became apparent that because there was no local support, the most minor problems such as a short time power loss could necessitate a system reset that could not be determined remotely. The only indication was "no contact" with the system and with no other indication of what might be wrong, personnel from the FAA Technical Center in Atlantic City had to travel with or acquire troubleshooting and repair equipment after they arrived. Since access to the vault was generally only available from 10 pm to 6 am, a minimum of two trips to the site had to be planned, allowing for troubleshooting one night, obtaining parts/equipment the next day and effecting the repair that evening. Or, in a simple case of "rebooting" the system, the second day would be "wasted" or require costly travel itinerary changes. As different failures occurred, more redundancy and backups were introduced into the system in an attempt to make it more "robust" but some other problem would eventually appear. The majority of the problems encountered were caused by poor environmental conditions in the vault that included flooding and caustic air quality or the destruction of above ground triggering sensors by grass cutting or maintenance personnel.

Therefore, future instrumentation projects have always included in site planning:

- Protecting systems from the environment and allowing easy monitoring;
- Providing redundant (backup) power and communications;
- Establishing a local point of contact that is aware of all of the different components of system(s), has the ability to physically inspect/monitor them and can provide at least minimum; troubleshooting and repair capabilities;
- Scheduling preventative maintenance and local contact coordination meetings.

***Summary***

This paper summarizes FAA's research efforts for the restoration of DIA instrumentation project. The main objective of FAA's project on Field Testing and Instrumentation is better understanding of long-term pavement behavior in the field under varied climatic and operating conditions, and improved characterization of paving materials. The DIA instrumentation project is one of the three FAA field instrumentation projects (others being Atlanta's Hartsfield-Jackson International Airport, and New York's LaGuardia International Airport project). The analyses of data collected from these projects have shown that the instrumentation of an in-service airport pavement allows not only the evaluation of existing structural models to predict pavement responses, but also studying the

1. Changes in pavement responses over time;
2. Effects of aircraft type and gear; and
3. Effects of environment.

The data from field instrumentation projects coupled with NAPTF data will help in designing longer lasting and better performing airport pavements. Improved pavement design and evaluation tools will conserve airport development funds

and reduce the downtime of airfield pavements for construction and maintenance activities.

***Acknowledgement***
The work described in this paper was supported by the FAA Airport Technology Research and Development Branch, AAR-410, United States of America, Dr. Satish K. Agrawal, Manager. Acknowledgements are also due to the Denver International Airport staff (Michael H. Steffens, Alvin Lorenzo, Robert W. Meyers, Ricky Rogers, Shawn Clark, and Jacqueline Meyer) for their continued support in all matters during this project. The contents of the paper reflect the views of the authors, who are responsible for the facts and accuracy of the data presented within. The contents do not necessarily reflect the official views and policies of the FAA, USA. The paper does not constitute a standard, specification, or regulation.

***References***
1. Brill, David R., *Field Verification of a 3D Finite Element Rigid Airport Pavement Model*. Report DOT/FAA/AR-00/33, Federal Aviation Administration, Office of Airport Safety and Standards, Washington, DC, July 2000.
2. Dong, M., Hayhoe, G.F., and Fang, Y., *Runway Instrumentation at Denver International Airport: Dynamic Sensor Data Processing*. Proceedings of the ASCE Airfield Pavement Conference, August 17-20, 1997, U.S.A.
3. HoSang, V. A., *Field Survey and Analysis of Aircraft Distribution on Airport Pavements*. Report No. FAA-RD-74-36, Federal Aviation Administration, Washington, D. C., 1975.
4. Lee, X., Hovan, M., King, R., Dong, M., and Hayhoe, G.F., *Runway Instrumentation at Denver International Airport – Development of Database*. Proceedings of the ASCE Airfield Pavement Conference, August 17-20, 1997, U.S.A.
5. Rufino, D. J.R. Roesler, E. Tutumluer, and E.J. Barenberg. *Wander Patterns for Commercial Aircraft at Denver International Airport.* Proceedings of the 2001 Airfield Specialty Conference, 27th ASCE International Air Transportation Conference, Chicago, Illinois August 2001.
6. Rufino, D. and J. R. Roesler., *Effects of Slab-Base Interaction of Measured Concrete Pavement Responses*, ASCE Journal of Transportation Engineering, Vol. 132, No. 5, pp. 425-434.

# Flowable Fibrous Concrete for Thin Concrete Inlays

A.C. Bordelon[1] and J.R. Roesler[2]

[1]Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, 205 N. Matthews Ave, Urbana, IL 61801; PH (217) 333-7311; FAX (217) 333-1924; email: bordelon@illinois.edu
[2]Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, 205 N. Matthews Ave, Urbana, IL 61801; PH (217) 265-0218; FAX (217) 333-1924; email: jroesler@illinois.edu

## Abstract

A thin concrete pavement wearing surface has been developed using a flowable fibrous concrete (FFC) as a 50 mm (2 in.) inlay on a milled asphalt concrete pavement. The flowable concrete mixture incorporating a hybrid of synthetic fibers was optimized for rapid placement and consolidation. An objective of the wearing surface was to construct reasonable slab sizes and crack widths while ensuring economic feasibility. Laboratory testing demonstrated the FFC has a slump flow diameter value of 400 mm (15.5 in.) and post-cracking residual strength ratio over 47 percent with 0.5 percent volume fraction of synthetic macrofibers. A full-scale demonstration project was cast to evaluate constructability and concrete material performance including placement issues, crack spacing and width development, and interface bonding conditions. Different slab sizes were explored from 1.2 to 3.3 m (4 to 11 ft) with the longer slabs having the earliest and widest cracking up to 1.25 mm (0.05 in.). The joints cracked before day one reside as the largest crack widths measured at later ages. In-situ bonding tests confirmed a good bond between the asphalt and concrete, except in locations where debris from the asphalt layer was not adequately removed.

## Introduction

Asphalt concrete surfaces are historically the standard choice for thin overlays of existing pavement structures due to their rapid placement, ability to construct in a thin layer, and allowance for opening to traffic shortly after placement. Current concrete technology allows for the placement of thin concrete sections with minimal consolidation and with reduced construction time. The concrete material constituents and proportions must be chosen, and their effects on the concrete overlay performance must be studied, to ensure

economically viability of this new concrete inlay technology to achieve 10 to 15 year service life.

Several recent projects and studies have experimented with 38 to 50 mm thick concrete overlays. Thin steel (up to 0.5 percent volume fraction) fiber-reinforced concrete overlays have been studied for bridge decks (Carlswärd 2006), or with strain-hardening concrete material composed of 2 percent of a poly-vinyl alcohol fiber-reinforcement (Zhang and Li 2002; Kunieda and Rokugo 2006). A recent experiment has also been completed to determine the feasibility of thin concrete pavement overlays using self-consolidating concrete (Riley 2005). Self-compacting concrete (SCC) technology has been used in structural concrete to rapidly construct walls and beams, and minimize the amount of vibration necessary for proper consolidation. Analogously, by using a more workable or flowable concrete mixture to promote rapid concrete pavement inlay construction, the likelihood of voids from poor compaction/consolidation efforts are reduced. Higher volume fiber reinforcement is one key to good performance of thinner concrete sections (Zhang and Li 2002; Ward and Li 1990), which in turn requires a flowable mixture to counteract the loss in workability during placement.

This preliminary research has developed a flowable concrete mixture utilizing a hybrid of fiber-reinforcements. The mixture is designed to be placed rapidly in a thin layer and provide a high cracking resistance and residual load capacity for the thin inlay system. A full-scale thin overlay project was created to demonstrate and observe the placement of this material and to monitor the cracking and bonding performance of the hardened concrete pavement.

**Concrete Mixture Characteristics**

A flowable fibrous concrete (FFC) mixture was developed to facilitate construction of a concrete pavement inlay for low to moderate-speed roads and parking lots. This FFC technology integrates the principles of SCC mixture design (necessary for quick placement and proper consolidation of the thin pavement with a higher fiber content) with the use of hybrid fiber-reinforcement (necessary for fracture toughness and to achieve the service life).

The benefits of fiber-reinforced concrete (FRC) for thin concrete surfaces are to increase the material toughness and load carrying capacity of slabs as well as minimize spalling and reduce the rate of crack growth. The FFC mixture used in the field project contained 0.48 percent volume fraction of 40 mm-long polymeric structural macrofibers combined with 0.06 percent of 6 mm-long polypropylene microfibers (commonly used for plastic shrinkage mitigation). The addition of the short synthetic fibers in the FFC is for increasing the yield stress, viscosity, and segregation resistance of the mixture (Banfill et al. 2006). The cementitious materials content was higher than typical paving mixtures and used in combination with a superplasticizer to improve the workability and flow becasue of the use of fibers. Because the concrete was placed as a 50 mm inlay, a smaller maximum size aggregate (9.5 mm) was used. Table 1 lists the mixture design constituents and proportions.

**Table 1. Field Concrete Mixtures**

| Material | Unit | Truck Batches | Lab Mixture |
|---|---|---|---|
| Sand | kg/m$^3$ | 890 ± 5 | 857 |
| Limestone (9.5 mm NMAS) | kg/m$^3$ | 1068 ± 2 | 1052 |
| Type I Cement | kg/m$^3$ | 292 ± 1 | 292 |
| Type C Fly Ash | kg/m$^3$ | 128 ± 6 | 125 |
| HRWR | ml/m$^3$ | 2.8 ± 0.1 | 1.3 |
| AEA | ml/m$^3$ | 0.10 ± 0.01 | 0.10 |
| Water | kg/m$^3$ | 168 | 167 |
| Fiber 1 (40 mm long) | kg/m$^3$ | 4.4 | 4.4 |
| Fiber 2 (6 mm long) | kg/m$^3$ | 0.6 ± 0.1 | 0.5 |
| Estimated w/cm | - | 0.40 ± 0.01 | 0.40 |
| w/c | - | 0.58 ± 0.002 | 0.57 |

NMAS = nominal maximum aggregate size
AEA = air- entraining admixture
HRWR = high range water reducer

The average slump flow, unit weight, and air content were 394 mm (15.5 in.), 2236 kg/m3 (140 lb/cf), and 6.8 percent air, respectively, for the FFC mixture used in the field project. The increased workability of the SCC-like mixture reduces the required compaction energy relative to typical paving mixtures. The FFC mixture is unlike conventional SCC mixtures since the FFC is cohesive enough to hold the longitudinal grade and transverse slope of a roadway. The averaged hardened strength (compressive, split-tensile, and flexural) properties are shown in Table 2. The FFC was designed to have a high post-cracking toughness performance, as illustrated by its measured residual strength ratio $R_{150}^{150}$ (ASTM Standard C1609 2007) of 47.6 percent at 7 days. Recommendations for fiber reinforcement in a similar thin concrete overlay application of ultra-thin whitetopping (UTW) require a 20 percent residual strength ratio, which is significantly less than the FFC mixture (Roesler et al. 2008).

Fracture properties are necessary inputs for finite element-based fracture models that are capable of characterizing the load capacity of concrete slabs (Gaedicke 2009). The critical mode I stress intensity factor ($K_{Ic}$), critical crack tip opening displacement ($CTOD_c$), initial fracture energy ($G_f$), and total fracture energy ($G_F$) of the FFC were calculated from the test results of a single-edge notched beam (SEN[B]). The initial fracture properties are comparable to other FRC mixtures used for pavement design, while the total fracture energy of FFC is significantly greater than lower fiber content FRC mixtures (Roesler et al. 2008). Figure 1 shows the effect of specimen size on the flexural stress versus normalized deflection for the FFC material. For the same FFC material, as the height of the specimen decreases, the residual capacity increases due to better fiber alignment. The influence of specimen casting is verified by the testing of a 7.5 cm beam extracted from a 15 cm beam (see illustration inset in Figure 1), which produced similar residual stresses to the 15 cm beam, as seen in Figure 1. Another FRC material factor is the decrease in the average residual strength and fracture energy values with respect to the concrete age, seen in Table 2. This reduction is related to the bond strength between the fiber and cement matrix and the fiber alignment. As the bond

strength increases, more fibers tend to rupture, especially when fibers are not orientated perpendicular to the crack plane.

**Table 2. Average Hardened Properties of the FFC**

| **Cylinder Properties (10 cm diameter x 20 cm)** | | | | **Flexural Properties (15 x 15 cm beams)** | | |
|---|---|---|---|---|---|---|
| Age | $f'_c$ | $f'_{sp}$ | $E$ | $MOR$ | $f^{150}_{150}$ | $R^{150}_{150}$ |
| (days) | (MPa) | (MPa) | (GPa) | (MPa) | (MPa) | (%) |
| 7 | 35.8 | 3.49 | - | 3.78 | 1.80 | 47.6% |
| 28 | 41.5 | 4.25 | 31.7 | 5.02 | 1.41 | 28.0% |

| **Flexural Properties (7.5 x 7.5 cm)** | | | | | |
|---|---|---|---|---|---|
| Age | $MOR$ | $f^{75}_{150}$ | $R^{75}_{150}$ | $f^{75}_{75}$ | $R^{75}_{75}$ |
| (days) | (MPa) | (MPa) | (%) | (MPa) | (%) |
| 7 | 4.49 | 2.61 | 58.1% | 2.19 | 48.7% |
| 28 | 4.77 | 2.56 | 53.7% | 2.12 | 45.0% |

| **Fracture Properties (15 cm deep SEN(B))** | | | | | |
|---|---|---|---|---|---|
| Age | $P_c$ | $K_{Ic}$ | $CTOD_c$ | $G_f$ | $G_F$ |
| (days) | (kN) | (MPa m$^{1/2}$) | (mm) | (N/m) | (N/m) |
| 7 | 4.24 | 1.24 | 0.026 | 63.2 | 3691 |
| 28 | 4.21 | 1.26 | 0.019 | 52.8 | 3175 |

$f'_c$ = compressive strength
$f'_{sp}$ = split-tensile strength
$E$ = elastic modulus
$MOR$ = modulus of rupture or flexural strength
$f^d_X$ = residual stress of for beam depth $d$ (in mm) at deflection of span/$x$
$R^d_X$ = residual stress ratio of for beam depth $d$ (in mm) at deflection of span/$x$
$P_c$ = peak load

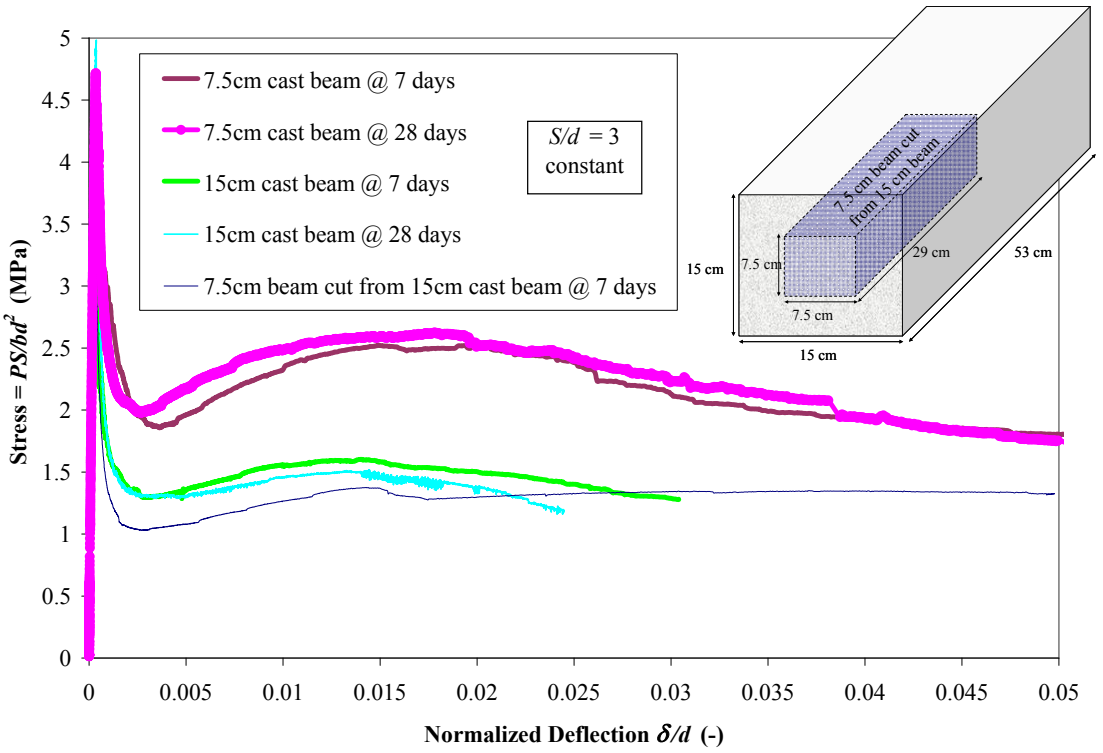



**Figure 1. Plot of four-point bending flexural stress versus a normalized deflection (midspan deflection/beam depth) for the FFC mixture at different ages and for various cast beam sizes.**

## Full-Scale Test Section Construction

After a satisfactory FFC mixture was developed in the laboratory that met workability, strength, and toughness requirements, a full-scale, thin concrete inlay project was constructed to investigate the construction aspects of FFC and to gather cracking and bond condition data. Previous studies (Carlswärd 2006) have shown challenges assessing the bonding, crack spacing, and crack width of thin overlays from laboratory specimens. This field project provided more realistic stresses and allowed for the evaluation of the de-bonding potential between the concrete inlay and the underlying asphalt substrate layer. Figure 2 illustrates the proposed joint layout of the three full-scale test sections, all constructed at a thickness of 5cm; the only variable between the three sections was the slab size (1.15 x 1.12 m, 1.73 x 1.68 m, or 1.72 x 3.35 m).

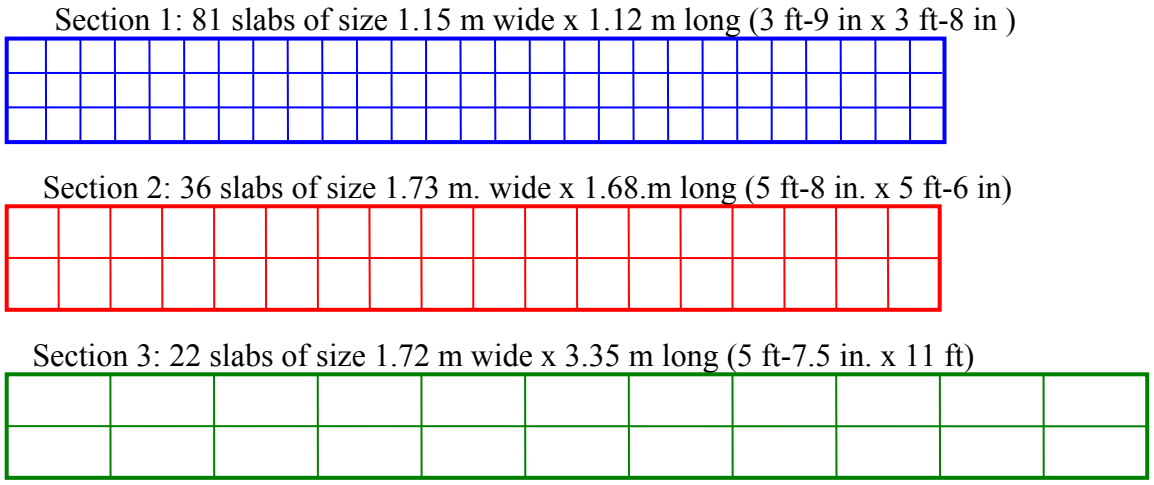



**Figure 2. Top view of the full-scale section configurations.**

The service life of a thin bonded concrete overlay or inlay is directly related to the concrete-asphalt interface bond. If the bonded concrete overlay is designed and constructed properly such that bonding exists, the tensile stresses within the concrete inlay are reduced such that the concrete functions as a wearing surface and improves the overall structural capacity of the pavement section. Researchers have suggested a variety of parameters to enhance bond, reporting some inconsistency on whether milling or grinding is required, as long as the underlying surface is clean prior to casting (Silfwerbrand and Paulsson 1998; Harrington 2008). At the site location, an existing aged and oxidized hot-mixed asphalt (HMA) was milled 50 mm and roughly 3.3 m wide. A 50 mm lip of the remaining HMA on either side of the lane acted as the side (vertical) formwork for containing the FFC, similar to a curb or gutter on a city street. Prior to a casting, the milled HMA surface was cleaned using high pressure water to remove dust and loose asphalt pieces and was allowed to dry overnight.

During the FFC casting, the material was spread efficiently using the chute from the ready mixed concrete truck. When necessary, the concrete was further moved in place with shovels and rakes, as shown in Figure 3. A vibratory screed was used to strike-off and externally consolidate the concrete surface. Internal vibration was not necessary on this test section. Small fiber balls did occur occasionally with a few truck batches as a result of inadequate shearing during mixing; visible clumps were removed by hand prior to finishing. A bull-float was used for finishing followed by the application of a curing compound. Roughly four hours after casting, an early-entry saw was used to cut joints, according to the plan in Figure 2, at a depth of 13 mm (0.5 in.) (approximately 1/4 of the slab thickness).

**Figure 3. Photographs from construction of the FFC inlay.**

## Inlay Cracking and Interface Bonding

The activation of joints was monitored at ages of 1, 3, 8, and 20 days by crack frequency and crack widths. The transverse contraction joints exhibited cracking as summarized in Figure 4. The transverse joints in the largest slabs (section 3) exhibited the most percentage of joints cracked after one day, but joints within section 1 closely matched the crack frequency in section 3 at later ages. By 20 days, all of the transverse joints in all sections had cracked. There was no visible, random cracking within the slabs.

A plot of the crack width versus age for each section can be seen in Figure 5. As expected, the longer the slab length, the greater the measured crack width. Cracks measured at one day were investigated to determine whether any link between early age crack development and the crack widths at later ages exists, as previously discovered on UTW projects in Illinois (Roesler et al. 2008). As seen in Figure 5, the data confirms that cracks appearing by day one have a larger average crack width than the average crack width of all joints at day 20. This reinforces the desire to have contraction joints crack as early as possible as this would likely reduce the average crack widths at each joint at all ages and distribute the movement of each slab to avoid the potential for de-bonding.

The interface bonding condition was investigated on the field project after the concrete had gained sufficient strength. Initially, a chain drag used on the surface of each slab indicated some debonding regions occurred in part of section 1 near the joints. Section 1 contained a stone matrix asphalt surface, which was not completely removed in areas during milling, and therefore left the potential for debonding/sliding between the concrete and underlying asphalt. For the remaining sections, an in-situ torsional test (Tashman et al. 2006; Leng et al. 2008) was selected to estimate the bond strength between the new concrete and asphalt layers at the center of the slab. Although the torsional test was found to be highly variable, several of the tests exceeded the range on the torque wrench at 500 N-m. Typical HMA surface layers on concrete have been reported to produce torque values around 400 N-m.

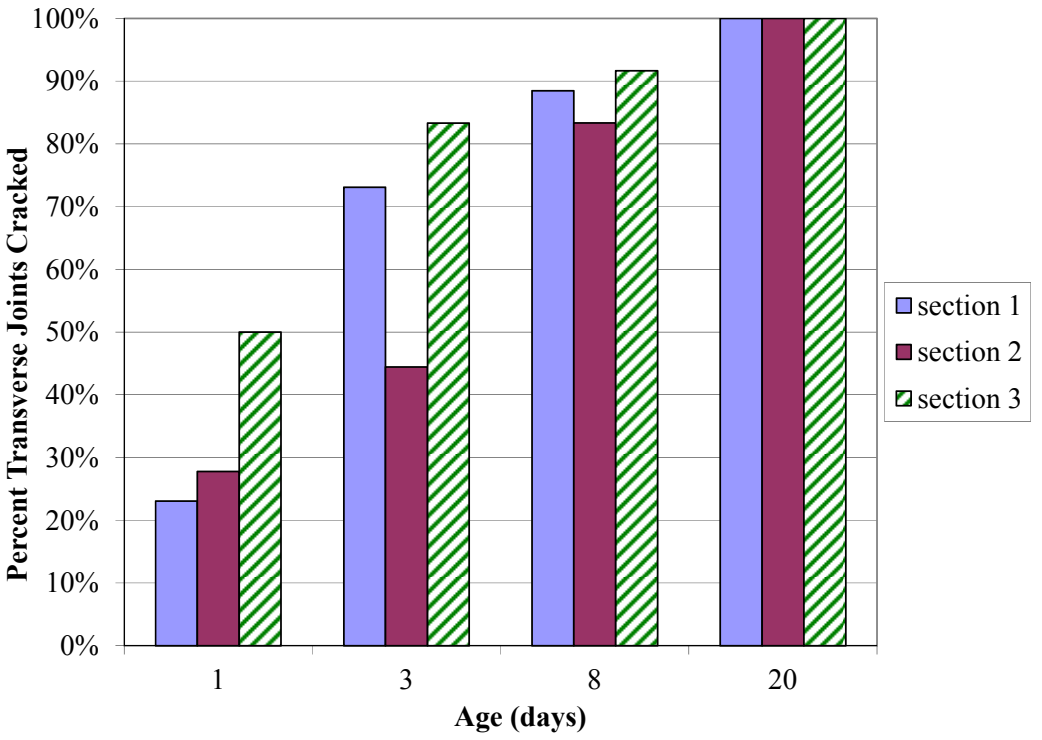



**Figure 4. Percent of transverse joints cracked at different concrete ages.**

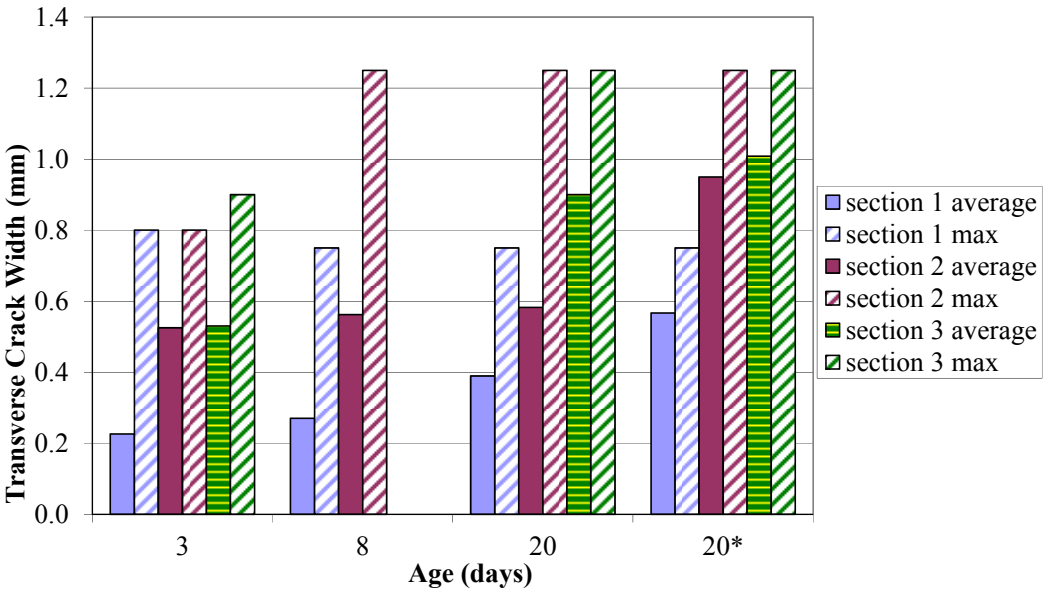



**Figure 5. Average and maximum transverse crack widths at different concrete ages (20* denotes the average crack widths calculated using only joints cracked at 1 day).**

## Conclusions

Concrete can be designed for thin, rapid pavement inlay construction and has the potential to be economically competitive in service life and performance to conventional asphalt overlays. A flowable fibrous concrete mixture was developed using hybrid fiber-reinforcement at higher volume fraction than other FRC mixtures commonly used in pavements. With the aid of a slightly higher cementitious content, superplasticizer, and smaller maximum aggregate size, the mixture was workable enough to be easily constructed as a thin inlay with only external vibration. The FFC mixture was tested in the laboratory and found to provide superior residual capacity and fracture toughness properties far above the requirements for UTW design.

The initial performance was measured from full-scale inlay sections of the FFC mixture at 5 cm thick on a milled asphalt pavement. Different slab sizes ranging from 1.2 to 3.3 m were constructed in the full-scale sections. The longest slab sizes produced the earliest and widest cracking at the joints, with the joints cracking before day one also residing as the largest crack widths measured at later ages. The sections predominantly exhibited good interfacial bonding between the asphalt and concrete. De-bonding was only found in locations where debris remained at the interface from insufficient milling and cleaning.

## References

ASTM Standard C1609. (2007). Standard Test Method for Flexural Performance of Fiber-Reinforced Concrete (Using Beam With Third-Point Loading). ASTM International, West Conshohocken, PA.

Banfill, P. F. G., Starrs, G., Derruau, G., McCarter, W. J., and Chrisp, T. M. (2006). "Rheology of low carbon fibre content reinforced cement mortar." Cement and Concrete Composites, 28(9), 773-780.

Carlswärd, J. (2006). "Shrinkage cracking of steel fibre reinforced self compacting concrete overlays: test methods and theoretical modelling." PhD thesis, Lulea University of Technology, Lulea, Sweden.

Gaedicke, C. (2009). "Fracture-based Method to Determine the Flexural Load Capacity of Concrete Slabs." PhD thesis, University of Illinois at Urbana-Champaign, Urbana, IL.

Harrington, D. (2008). Guide to Concrete Overlays. American Concrete Pavement Association, Skokie, IL.

Kunieda, M., and Rokugo, K. (2006). "Recent Progress on HPFRCC in Japan: Required Performance and Applications." Journal of Advanced Concrete Technology, 4(1), 19-33.

Leng, Z., Ozer, H., Al-Qadi, I. L., and Carpenter, S. H. (2008). "Interface bonding between hot-mix asphalt and various portland cement concrete surfaces laboratory assessment." Transportation Research Record, (2057), 46-53.

Riley, R. (2005). "SCC Quick Base and Patching Material 2005." Illinois Chapter of the American Concrete Pavement Association Presentation.

Roesler, J., Bordelon, A., Ioannides, A., Beyer, M., and Wang, D. (2008). "Design and Concrete Material Requirements for Ultra-Thin Whitetopping." Rep. No. FHWA-ICT-08-016, Research Report for the Illinois Center of Transportation R27-3A, Urbana, IL.

Silfwerbrand, J., and Paulsson, J. (1998). "Better bonding of bridge deck overlays." Concrete International, 20(10), 56-61.

Tashman, L., Nam, K., and Papagiannakis, T. (2006). "Evaluation of the Influence of Tack Coat Construction Factors on the Bond Strength between Pavement Layers." Rep. No. WA-RD 645.1, Washington State University, Washington Center for Asphalt Technology, Pullman, WA.

Ward, R. J., and Li, V. C. (1990). "Dependence of Flexural Behavior of Fiber Reinforced Mortar on Material Fracture Resistance and Beam Size." ACI Materials Journal, 87(6), 627-637.

Zhang, J., and Li, V. C. (2002). "Monotonic and fatigue performance in bending of fiber-reinforced engineered cementitous composite in overlay system." Cement and Concrete Research, 32(3), 415-423.

# Estimation of mechanical characteristics of cement treated demolition waste

Dongxing Xuan[1] André A.A. Molenaar[2] Lambert J.M. Houben[3] Zhonghe Shui[4]

[1]MSc, Ph.D student, Section Road and Railway Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands; PH + 31 152784674; FAX + 31 152783443; e-mail: d.xuan@tudelft.nl
[2]Professor, Section Road and Railway Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands; PH + 31 152784812; FAX + 31 152783443; e-mail: A.A.A.Molenaar@tudelft.nl
[3]Associate Professor, Section Road and Railway Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands; PH +31 152784917; FAX +31 152783443; e-mail: L.J.M.Houben@tudelft.nl
[4]Professor, School of Materials Science and Engineering, Wuhan University of Technology, Luoshi Road No. 122, Wuhan, P.R. China; PH + 86 027 87210782; FAX + 86 027 87210782; e-mail: shuizh@whut.edu.cn

## ABSTRACT

Cement treated granular materials have been applied for pavement bases and subbases. The availability of prediction models that allow their mechanical properties to be estimated from material parameters like composition and degree of compaction would be extremely useful because it would overcome the need for extensive testing. This paper presents an experimental study on establishing models to estimate the mechanical properties of cement treated mix granulates with recycled crushed masonry and concrete aggregates ($CTM_iG_r$). Mechanical properties include compressive strength, indirect tensile strength and elastic modulus. Variables influencing the strength and the elastic modulus, such as the ratio of crushed masonry to crushed concrete, the cement content, the water content and the degree of compaction were considered during the mixture preparation of $CTM_iG_r$. The experimental results show that the compressive strength, the indirect tensile strength and the elastic modulus of $CTM_iG_r$ after 28 days of curing may be accurately estimated on basis of the content of recycled masonry, the ratio of cement to water and the dry density.

**KEY WORDS:** cement treated road base; mechanical properties; material parameters; demolition waste; mix granulates

## 1. INTRODUCTION

Cement treated granular materials (CTGM), which are a family of compacted mixtures with aggregates, cement and water, have been widely applied for road bases and subbases. Since 1915, when a pavement in Sarasota, Florida, was constructed and compacted by using a mixture of shells, sand and Portland cement, more and more granular materials are treated with cement to obtain the required performance (Terrel et al 1979). In practice, the application of a cement treatment is mainly because of the following reasons (TRH 13 1986):

- modify the workability of engineering materials,
- increase the strength of the mixture,
- improve the durability, and
- enhance the load spreading capacity (elastic modulus)

CTGM as a road (sub-)base material is generally produced by using coarse natural or crushed aggregates (Bell 1993). In order to protect natural resources and reduce the environmental pollution from dumping of Construction and Demolition waste (C&D), some recycled aggregates, such as crushed concrete and crushed masonry, have been considered for use in road bases. In the Netherlands, unbound mix granulates (mixtures of crushed concrete and crushed masonry) have been successfully and extensively used as road base materials since the 1970s. Currently over 80% of the materials used for road bases are mix granulates in the Netherlands (Van Niekerk 2002; Molenaar 2005). Also in China, the feasible reuse of recycled concrete aggregates and crushed clay brick as unbound road sub-base has received attention recently (Poon and Chan 2006). Although the recycling aggregates from C&D may be reused in the road base, the performance of cement treated demolition waste as a base course is not effectively investigated yet. Therefore, it is considered valuable and meaningful to explore the mechanical properties of the marginal-quality C&D waste as cement treated base or sub-base in pavement structures.

Moreover, the current mixture design methods of CTGM are based on the optimum moisture content/maximum density and minimum compressive strength. But, it has to be accepted that the optimum moisture content and maximum density can not be found or are not obvious for some materials in practice (Molenaar 2005). This certainly influences the mixture design of CTGM. In addition, the traditional mixture design of CTGM is strongly determined by the minimum compressive strength, which needs to be determined by experimental tests (Kennedy 1983; Molenaar 2007). If the minimum compressive strength can not be obtained, the mixture design needs to be done again. It is obvious that models that allow the mechanical properties of CTGM to be estimated from material parameters like composition and degree of compaction are extremely useful. Such models would allow estimations to be made of the composition and the degree of compaction that are required to obtain certain characteristics. This would significantly limit the amount of testing to be done. Only few tests to verify the proposed design are then necessary.

On basis of the considerations mentioned above, this paper intends to explore the mechanical behavior of cement treated mix granulates with recycled crushed masonry and concrete aggregates ($CTM_iG_r$). Models to predict the mechanical properties of $CTM_iG_r$ will be developed on basis of its mixture composition and degree of compaction.

## 2. EXPERIMENTAL PROGRAM

### 2.1 MATERIALS

Raw materials used in this research, recycled concrete aggregates (RCA) and recycled masonry aggregates (RMA), were sourced from two Dutch companies. Figure 1 shows the original concrete and masonry rubble. There are less than 15% of other types of stony materials for the RMA fraction on the sieve 8 mm.

**Figure 1. Original concrete rubble (left) and masonry rubble (right).**

The crushed RCA and RMA were individually sieved in the following 6 fractions: 31.5-22.4 mm, 22.4-16.0 mm, 16.0-8.0 mm, 8.0-5.6 mm, 5.6-2.0 mm, <2.0 mm. The target grading curve of mix granulates was designed by Equation 1:

$$P = (100 - F) \cdot \frac{d^n - 0.063^n}{D^n - 0.063^n} + F \quad (1)$$

Where, P = percentage passing sieve size d [mm]
D = maximum particle size [$d_{max}$] (31.5 mm in this study)
F = Filler content (F=2.24, close to the fines content (< 0.063 mm) in crushed concrete aggregates)
n = a parameter describing the shape of the grading curve (n=0.45 in this study)

In the studied $CTM_iG_r$, the following ratios of masonry to concrete by mass were applied 100%:0%, 65%:35%, 35%:65%, 0%:100%, respectively. Figure 2 shows the cross sections of $CTM_iG_r$ with different masonry contents.

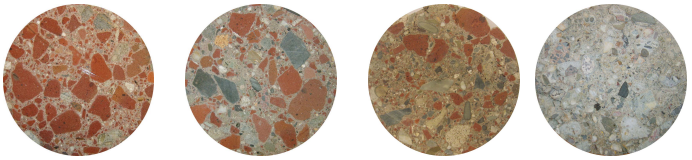


**Figure 2. Cross sections of CTM $_iG_r$ (masonry content decreases from left to right).**

In addition to the recycled aggregates, EN 42.5 Portland cement and tap water were used to prepare the test specimens.

## 2.2 MIXTURE DESIGN

Cement content and degree of compaction were selected as variables to prepare the $CTM_iG_r$ mixtures. The mixture composite design was determined by using the central composite design method (Muraya 2007). Five levels were used for each variable. The cement content is based on the ratio of cement mass to the total mass of the aggregates and varied from 2.5% to 5.5%. The water content is determined by the One Point Proctor test, Annex B of EN 13286-2. The degree of compaction refers to the One Point Proctor density and varied from 97% to 105%. Figure 3 shows the mixture design method with two variables and five levels.

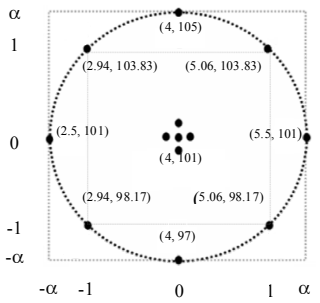



**Figure 3. Coded test conditions in the central composite design.**

## 2.3 MIXTURE PREPARATION

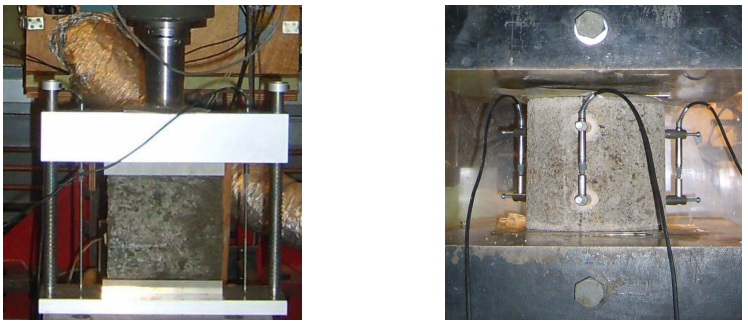
In the laboratory the $CTM_iG_r$ mixture was firstly mixed by using a laboratory mixer. The fresh mixture obtained was compacted in three layers in the mould of Φ150×150 mm by using a Bosch vibrating hammer. The detailed compaction procedure is as follows:

- pour a specific mass of the fresh mixture in the mould of Φ150×150 mm, which was determined by the degree of compaction;
- distribute the material evenly to avoid the segregation;
- poke the fresh mixture a couple of times near the edge of the cylinder mould by using a screwdriver;
- pre-compact the mixture by using the hammer;
- compact each layer by the vibrating hammer (compacting time depends on the required degree of compaction);
- make the surface of the compacted layer rough by a screwdriver to ensure a good adhesion with the next layer;
- continue to finish three layers.

After 24-hours curing at room temperature, all specimens were demolded and subjected to a fog-room curing at 20ºC. After a curing time of 28 days, the specimens were ready for testing.

## 2.4 MEASUREMENT OF COMPRESSIVE STRENGTH AND INDIRECT TENSILE STRENGTH

The unconfined compressive strength (UCS) and indirect tensile strength (ITS) of the mixtures have been determined. The tests were done using a MTS actuator of 245 kN and 150 kN respectively in the displacement control mode. Displacements were controlled by linear variable differential transformers (LVDTs) along the axial deformation of the specimen. For UCS, a friction reduction system was used to obtain uniform radial deformations over the height of the specimen (Erkens 2002). The strain rate for UCS is $10^{-5}$/second. The displacement rate chosen for ITS was 0.2 mm/second. The data of the force and the deformation are automatically recorded by a MP3 or Labview program. Figure 4 shows the experimental set-ups in the laboratory.



(a) Indirect tension test (b) Unconfined compression test

**Figure 4. Mechanical tests in the laboratory.**

## 3. RESULTS AND DISCUSSION

### 3.1 MOISTURE-DRY DENSITY RELATIONSHIP

In accordance with the EN 13286-2 standard, $CTM_iG_r$ with a cement content of 4% by mass was compacted in order to produce the conventional moisture-dry density curve. Figure 5 shows the moisture content-dry density curve of $CTM_iG_r$. It can be seen that when the masonry content decreases from 100% to 0%, the dry density of $CTM_iG_r$ becomes higher. It means that the replacement of RMA by RCA results in an increase of the dry density of $CTM_iG_r$. This is attributed to the bigger density of RCA than that of RMA. Also note that in some cases, the maximum dry density of $CTM_iG_r$ may be found, but the tendency is not obvious or the change of the dry density with the moisture content is very small as is shown by means of the two dashed lines. This proves that the

optimum moisture content and maximum density can not be found or is not obvious for some materials in practice (Van Niekerk 2002; Molenaar 2005).

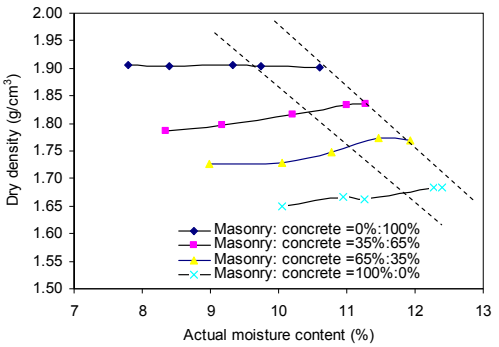



**Figure 5. Moisture-dry density relationship of $CTM_iG_r$.**

Meanwhile, it is found that when the masonry content increases, more water is needed to obtain a good workability. In this study, the cumulative water absorption of all fractions of masonry aggregates is approximately 10%. It was also observed that $CTM_iG_r$ with RMA could not keep more water in the fresh mixture under the Proctor compaction when the water content is close to the optimum level. This is due to the porous nature of recycled aggregates, especially masonry. When adding more than 10%water, the free water in the pore system (internal or in-between particles) can not be easily kept. As a result, the free water will flow away, which results in the loss of cement paste in the mixture as well. Because of these reasons, the reference density and moisture content in this study were determined by means of the standard One Point Proctor test according to Annex B of the European Standard EN 13286-2.

Table 1 lists the actual moisture content and dry density of $CTM_iG_r$ after the One-Point-Proctor compaction. The moisture content of $CTM_iG_r$ is over 9 % for a good workability. This value is higher than the normal optimum moisture content of cement treated natural aggregates (approximately between 5% and 8%) (Sherwood 1995).

**Table 1. Actual moisture content and dry density of $CTM_iG_r$ after the One-Point-Proctor compaction.**

| Ratio of RMA to RCA | Actual water (%) | Dry density ($g/cm^3$) | Appearance of fresh $CTM_iG_r$ |
|---|---|---|---|
| 100% : 0% | 11.81 | 1.662 | A little shinny; less bleeding |
| 65% : 35% | 10.94 | 1.754 | |
| 35% : 65% | 10.44 | 1.834 | |
| 0% : 100% | 9.54 | 1.907 | |

## 3.2 UNCONFINED COMPRESSIVE STRENGTH (UCS)

The UCS is a design index widely used to evaluate the performance of cement treated materials. Figure 6 shows the UCS of $CTM_iG_r$ in relation to the cement/water ratio, the degree of compaction and the masonry content. The dry density is related to the degree of compaction according to the mixture design in Figure 3. One will observe that the UCS linearly increases with the ratio of cement (C)/water (W) and exponentially increases with the dry density or the degree of compaction (D). And, with the increase of the masonry content (M), the slopes of these curves increase as shown in Figure 6.

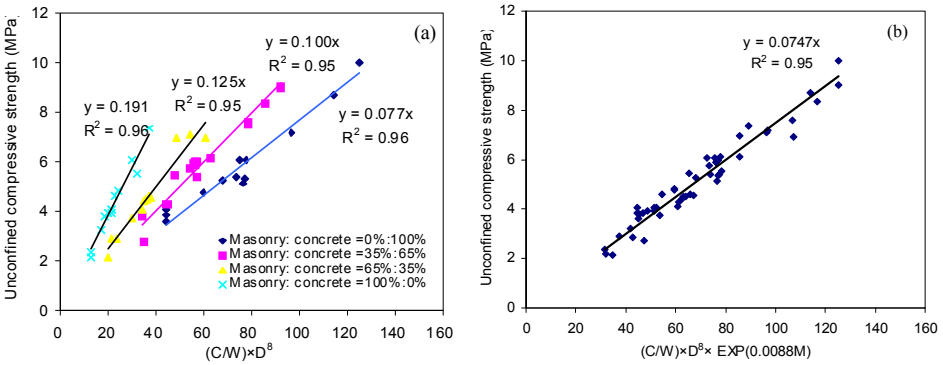



**Figure 6. UCS of $CTM_iG_r$ at 28 days.**

Based on the experimental data, a model to estimate the compressive strength of $CTM_iG_r$ has been developed:

$$f_c = 0.0747 \cdot \frac{C}{W} \cdot D^8 \cdot e^{0.0088 \cdot M} \quad (R^2 = 0.95,\ S=0.41\ \text{MPa}) \qquad (2)$$

Where, D is dry density, g/cm$^3$; C is the cement content by mass of aggregates, %; W is the water content by mass of aggregates, %; M is the masonry content by mass of the total aggregates, %. $R^2$ is the coefficient of determination; S is the standard deviation.

It is a well-known fact that the cement content and the degree of compaction of cement treated materials play important roles to improve the cohesiveness of cement-stabilized materials and their mechanical properties (Terrel et al 1979; TRH 13 1986; Williams 1986). It is more economic and efficient to achieve a high strength by good compaction rather than by trying to increase the cement content. The fact that a high density during compaction is the best way to ensure long-term durability is also recognized by others (Sherwood 1995). This study shows that the masonry content is another important factor to influence the mechanical performance of $CTM_iG_r$.

### 3.3 ELASTIC MODULUS (E)

The tangent elastic modulus is determined by the linear part of the stress-strain curve at the beginning of the compression test. Figure 7 shows the elastic modulus of $CTM_iG_r$ related to material parameters. It is found that an estimation model for the elastic

modulus of $CTM_iG_r$ may also be established with the C/W ratio, the dry density and the masonry content.

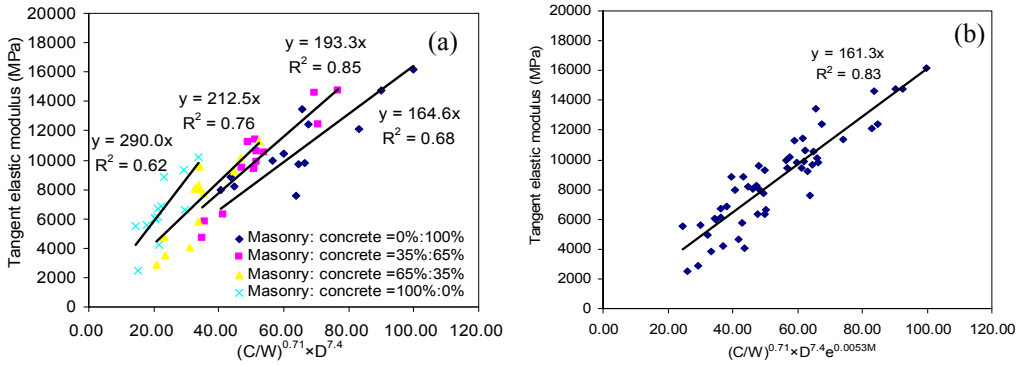



**Figure 7. Tangent elastic modulus of $CTM_iG_r$ at 28 days.**

In previous researches, most of the prediction models for the elastic modulus are not related to material parameters, but to the compressive strength (Xuan 2009). However, as both strength and elastic modulus are controlled by the nature of the $CTM_iG_r$ structure, an estimation model for the elastic modulus of $CTM_iG_r$ may be established. In this study, it is given as:

$$E_{\tan gent} = 161.3 \cdot (\frac{C}{W})^{0.71} \cdot D^{7.4} \cdot e^{0.0053 \cdot M} \quad (R^2 = 0.83, S=1317 \text{ MPa}) \quad (3)$$

### 3.4 INDIRECT TENSILE STRENGTH (ITS)

Figure 8 shows the influence of the C/W ratio, the dry density and the masonry content on the ITS of $CTM_iG_r$. A similar prediction model like that for the UCS (equation 2) can be derived. The ITS of $CTM_iG_r$ is also proportional to the increase of the C/W ratio, the dry density and the masonry content. The dry density is the most important factor for the ITS.

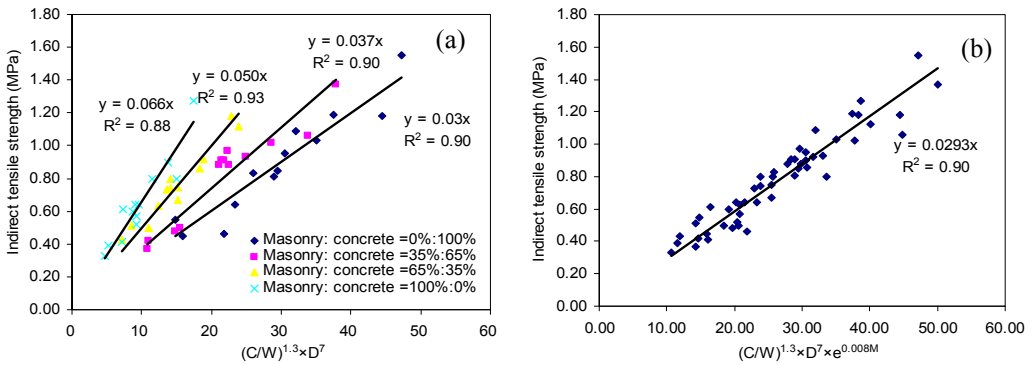



**Figure 8. ITS of $CTM_iG_r$ at 28 days.**

So far the prediction models for the ITS of cement treated materials are mostly focused on the linear relationship with UCS (Williams 1986; Xuan 2009). Based on the experimental results in this research, the following estimation model for the ITS ($f_{it}$) of $CTM_iG_r$ could be derived:

$$f_{it} = 0.0293 \cdot (\frac{C}{W})^{1.3} \cdot D^7 \cdot e^{0.008 \cdot M} \quad (R^2 = 0.90, S=0.09 \text{ MPa}) \qquad (4)$$

## 4. CONCLUSIONS

This study was conducted to explore the mechanical properties of cement treated mix granulates with recycled crushed masonry and concrete aggregates from C&D waste. It has been found that the ratio of cement to water and dry density influence the mechanical properties of $CTM_iG_r$. The masonry content in $CTM_iG_r$ is another unique factor to determine the mechanical properties of $CTM_iG_r$.

As both strength and elastic modulus are controlled by the nature of the $CTM_iG_r$ structure, a prediction model for them may be put forward. On basis of the experimental results, this general prediction model and its parameters for strength and modulus at 28 days are listed in table 2.

**Table 2. A model and its parameters for the strength and modulus of $CTM_iG_r$.**

| Model and Parameters | $A \cdot (\frac{C}{W})^{n1} \cdot D^{n2} \cdot e^{K \cdot M}$ | | | | | |
|---|---|---|---|---|---|---|
| | A | n1 | n2 | K | $R^2$ | S |
| $f_c$ | 0.0747 | 1 | 8 | 0.0088 | 0.95 | 0.41 |
| $f_{it}$ | 0.0293 | 1.3 | 7 | 0.0080 | 0.90 | 0.09 |
| $E$ | 161.3 | 0.71 | 7.4 | 0.0053 | 0.83 | 1317 |

## ACKNOWLEDGEMENTS

The authors would like to thank Mr. M.R. Poot, Mr. J.W. Bientjes and Mr. D.C. Doedens (lab staff in Road and Railway Engineering of Delft University of Technology) for their assistance during this research.

## REFERENCES

Bell F.G. (1993). "Engineering treatment of soils." *E&FN Spon*, London.

Erkens S.M.J.G. (2002). "Asphalt Concrete Response (ACRe)." *Ph.D Thesis, Delft University of Technology,* the Netherlands.

Kennedy J. (1983). "Cement-bound materials for sub-bases and roadbases." Cement *and Concrete Association Publication 46.027*, Slough.

Molenaar A.A.A. (2005). "Cohesive and non-cohesive soils and unbound granular materials for bases and sub-bases in roads." *Lecture notes CT 4850, Delft University of Technology*, the Netherlands.

Molenaar A.A.A. (2007). "Design of flexible pavements." *Lecture Report, Delft University of Technology*, the Netherlands.

Muraya P.M. (2007). "Permanent deformation of asphalt mixes." *Ph.D Thesis, Delft University of Technology*, the Netherlands.

Niekerk A.A.Van. (2002). "Mechanical behavior and performance of granular bases and sub-Bases in pavements." *Ph.D Thesis, Delft University of Technology*, the Netherlands.

Poon C.S. and Chan D. (2006). "Feasible use of recycled concrete aggregates and crushed clay brick as unbound road sub-base." *Constr Building Mater 20 (8),* 578-585.

Sherwood P.T. (1995). "Soil Stabilization with Cement and Lime." *State of the ART Review, HMSO press*, London.

Terrel R.L., Epps J.A., Barenberg E.J., Mitchell J.K., Thompson M.R. (1979). "Soil stabilization in pavement structures, a User's Manual-Volume 1: Pavement design and construction considerations, FHWA-IP-80-2." *Federal Highway Administration, Department of Transportation,* Washington D.C.

Terrel R.L., Epps J.A., Barenberg E.J., Mitchell J.K., Thompson M.R. (1979). "Soil stabilization in pavement structures, a User's Manual-Volume 2: Mixture design considerations, FHWA-IP-80-2." *Federal Highway Administration, Department of Transportation,* Washington D.C.

TRH 13. (1986). "Cementitious stabilizers in road construction." *National Institute for Transport and Road Research, Committee of State Road Authorities,* Pretoria, South Africa.

Williams R.I.T. (1986). "Cement-treated pavement: materials, design and construction." *Elsevier Applied Science Publishers*, England.

Xuan, D.X. (2009). "Literature review of research project: Structural properties of cement treated materials." *Report 7-09-217-1, Section Road and Railway Engineering, Delft University of Technology*, the Netherlands.

# PERFORMANCE OF PRECAST CONCRETE PAVEMENTS

S. Tayabji[1], N. Buch[2] and D. Ye[1]

[1]Fugro Consultants, Inc., Columbia, Maryland, USA
(phone: 410-997-9020; email: stayabji@aol.com)
[2]Michigan State University, Lansing, Michigan

## ABSTRACT

Precast pavement technology is a new and innovative construction method that can be used to meet the need for rapid pavement repair and construction. Precast pavement systems are fabricated or assembled off-site, transported to the project site, and installed on a prepared foundation (existing pavement or re-graded foundation). The system components require minimal field curing time to achieve strength before opening to traffic. These systems are primarily used for rapid repair, rehabilitation, and reconstruction of asphalt and portland cement concrete (PCC) pavements in high-volume-traffic roadways. The precast technology can be used for intermittent repairs or full-scale, continuous rehabilitation. As part of the US Strategic Highway Research Program 2 (SHRP 2), a study (Project R05) is underway to develop tools for the design, construction, installation, maintenance, and evaluation of precast concrete pavements. As part of this study, testing was conducted to obtain field performance data from selected precast concrete pavement projects constructed throughout the US. This paper summarizes the field test data collected from intermittent repair projects as well as from continuous application projects and presents the findings of the data evaluation.

## INTRODUCTION

Pavement rehabilitation and reconstruction, major activities for all U.S. highway agencies, have significant impact on agency resources and traffic disruptions because of extensive and extended lane closures. Traffic volumes on the primary highway system, especially in urban areas, have increased tremendously over the last 20 years, leading in many instances to an earlier-than-expected need to rehabilitate and reconstruct highway pavements. Pavement rehabilitation in urban areas is resulting in serious challenges for highway agencies because of construction-related traffic congestion and safety issues. A promising rehabilitation strategy is the effective use of precast concrete pavement systems, which provide for accelerated repair and rehabilitation of pavements and also result in durable, longer-lasting pavements. Precast concrete pavement systems are systems that are essentially fabricated or assembled off-site, transported to the project site and installed on a prepared foundation (existing pavement or re-graded foundation). These systems do not require field curing for the precast concrete panels and require only minimal time for system components to achieve strength before opening to traffic.

## SHRP 2 Modular Pavement Study

The "Renewal" focus area of the US Strategic Highway Research Program 2 (SHRP 2) emphasizes the need to complete highway pavement projects rapidly, with minimal disruption to the users and local communities, and to produce pavements that are long-lasting. A goal of this focus area includes applying new methods and materials to preserve, rehabilitate, and reconstruct roadways. The effective use of precast concrete pavement technologies for rapid repair, rehabilitation, and reconstruction of pavements addresses this goal. One of the projects funded under SHRP 2 is Project R05, Modular Pavement Technology. The objective of this project is to develop tools for use by highway agencies to design, construct, install, maintain, and evaluate modular pavement systems.

The Phase I effort under Project R05 identified a serious lack of field performance data from installed precast concrete pavement systems. Because of this serious gap in available performance data, the project team contacted several highway agencies in the US to request support with the performance data collection effort, as part of Phase II effort under Project R05. This paper summarizes the field test data collected to date from several precast concrete pavement projects and presents the findings of the data evaluation.

## BACKGROUND

The precast concrete pavement technology is gaining wider acceptance by North American highway agencies and contractors and precasters are beginning to seriously explore business opportunities related to precast concrete pavement applications. The precast concrete pavement technology is generally based on sound technical/engineering considerations and field installation processes appear to be workable given the severe working conditions for many of these projects. The application of precast concrete pavement technology can be classified as follows:

1. Intermittent repairs of concrete pavements - isolated full-depth repairs or isolated full slab replacement using precast concrete slab panels
2. Continuous concrete paving - full-scale project level rehabilitation (resurfacing) or reconstruction of asphalt and concrete pavements is performed using precast concrete panels.

### Precast Pavement Systems

Several recently developed techniques are available in the US, as follows:

1. Precast prestressed concrete pavement (PPCP) developed at the University of Texas
2. Jointed precast concrete pavement, proprietary and generic systems:
    a. Fort Miller Super-Slab system (proprietary)

b. Kwik Slab system (proprietary)
c. Roman Stone system (proprietary)
d. Michigan system (generic)
e. Illinois Tollway system (generic)
f. La Guardia International Airport system (generic)

Discussion of the various systems and techniques is given elsewhere (Tayabji et al., 2009; Hall and Tayabji, 2008; Merritt and Tayabji, 2009).

**Precast Pavement Use in the US**

Since about 2000, many highway agencies in North America have expressed interest in considering use of precast concrete for intermittent repair or continuous applications in heavily trafficked urban areas where extended lane closures are difficult. The following U.S. and Canadian highway agencies have accepted the use of precast pavement for production work:

1. Caltrans
2. Illinois Tollway Authority
3. Iowa DOT (as an alternate for bridge approach slabs)
4. Ministry of Transport, Ontario
5. Ministry of Transport, Quebec
6. New Jersey DOT
7. New York State DOT
8. New York State Thruway Authority

The following U.S. agencies have investigated or are investigating use of precast pavement:

1. Colorado DOT
2. Delaware DOT
3. Florida DOT (demonstration project planned for construction, 2010)
4. Hawaiian Agencies
5. Indiana DOT
6. Michigan DOT
7. Minnesota DOT
8. Missouri DOT
9. Texas DOT
10. Virginia DOT
11. Airport Authorities
    a. Port Authority of New York and New Jersey
    b. Metropolitan Washington Airport Authority
12. US Air Force

In addition to the North American initiatives, the Netherlands, France, Russia, and Japan are actively investigating or are using the precast concrete pavement

technologies. Also, a 32 plus km (20 plus mile) section of a tollway in Indonesia was constructed recently using the PPCP system.

## SHRP 2 FIELD TESTING PROGRAM

The following precast pavement projects were tested, between December 2009 and May 2010:

1. PPCP projects (all based on the Texas system)
    a. Georgetown Frontage Road, Texas
    b. I-57, Missouri
    c. Route 896, Delaware
    d. I-66, Virginia
2. Jointed continuous projects (all Fort Miller Super Slab system)
    a. Tappan Zee Toll Plaza, New York
    b. TH 62, Minnesota
    c. I-66, Virginia
3. Jointed repair projects
    a. I-295, New Jersey (Fort Miller Super Slab system)
    b. I-280, New Jersey (Fort Miller Super Slab system)
    c. I-675, Michigan (Michigan system)
    d. Route 27, New York (Roman Stone system)

### Test Program Overview

The following data were planned to be collected at each project to provide an assessment of the structural and functional performance of the installed/in-service precast pavement systems:

1. Condition Data (Visual condition survey)
2. Ride – using a high speed Profiler to determine the section IRI values.
3. Joint Elevation Difference (Faulting)
4. Joint Width Measurement
5. Deflection testing
    a. Mid-panel (basin) testing
    b. Testing at joints (wheel path) for LTE and void detection

### Deflection Testing Protocols

For jointed precast concrete pavement system:

Testing was conducted at the outer wheel path (OWP) locations at transverse joints and at the midslab location of the precast panel and control slabs of the existing pavement, as shown in Figure 1. The testing sequence was as follows:

1. Intermittent Repair Projects – The sequence of testing is as follows:
    a. Precast panel test, as shown in Figure 1, locations 1 to 5
    b. Test at the first existing slab panel that is not adjacent to a precast panel, using the test locations 6 to 8 only as shown in Figure 1.
2. Continuous Application Projects – For continuous application projects, the testing was conducted for as many precast panels as possible within the allowable traffic closure period. The sequence of testing was locations 1 to 3 for each panel.

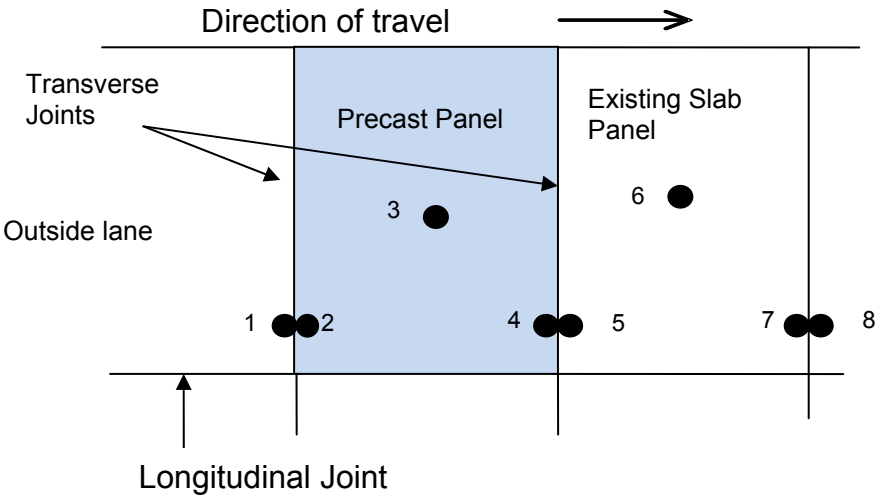



Figure 1: FWD test locations

The target load levels used were: seating load of 40 kN (9,000 lb), 26.7 kN (6,000 lb), 40 kN (9,000 lb), 53.4 kN (12,000 lb), and 66.7 kN (15,000 lb). The LTE test was conducted with deflection sensors located approximately 152 mm (6 in.) from the center of the joint on each side and the load located tangential to the joint on the loaded side of the joint. The interior (basin) load test was conducted at an approximate midslab location for Test Locations 3 and 6. The test loads were the same as shown in Table 1. Deflection sensors were spaced at 305 mm (12 in.) and a minimum of six sensors were used, including the load plate sensor. The total number of precast panels that were tested at a site was governed by the site access conditions. The plan was to test at least 20 precast panels and at least 10 existing slab panels.

For precast prestressed concrete pavement (PPCP) system:

Testing was typically conducted at the outer wheel path (OWP) locations at transverse expansion joints of the outside lane and at several midslab locations of the post-tensioned segments, as shown in Figure 2. At the Missouri and the Delaware projects, similar testing was also conducted at several conventional pavement slabs.

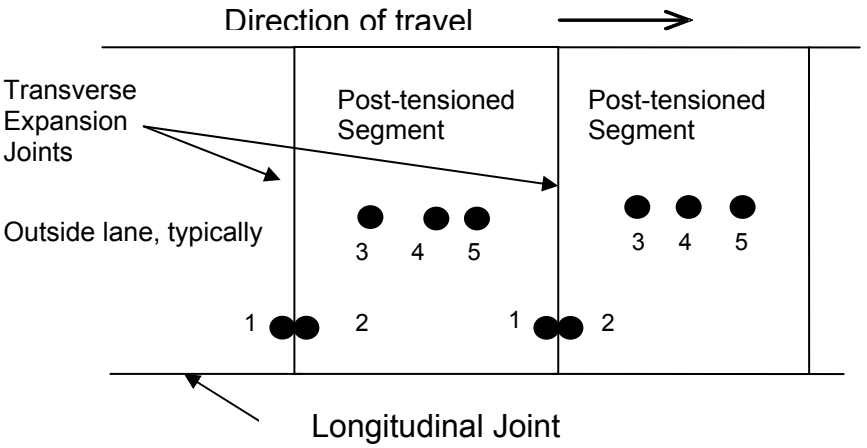



Figure 2 – FWD test locations

The target load levels were the same as for the jointed precast pavement testing. The LTE testing was performed similar to the procedure used for the jointed precast pavements. The interior (basin) tests were conducted at Test Locations 3, 4, and 5. Deflection sensors were spaced at 305 mm (12 in.) and a minimum of six sensors were used, including the load plate sensor.

## SUMMARY OF FIELD TESTING

Highlights of the field testing are presented in this section for representative projects only. Because of site access, field testing was typically limited to about 3 to 4 hours, daytime or nighttime. The respective highway agencies provided traffic control. In addition, the Michigan DOT and the Missouri DOT performed the deflection testing. In all cases, deflection testing was performed using a Dynatest FWD. Several agencies also provided ride data.

### Georgetown, Texas PPCP Project

This is the oldest PPCP project, constructed during November 2001. Pavement details are as follows:

- Panel thickness: 203 mm (8 in.)
- Panel dimensions: length – 3 m (10 ft); Width – 5, 6, and 11 m (16, 20, and 36 ft) (two 3.7 m [12 ft] lanes and 1.2 and 2.4 m [4 and 8 ft] shoulders)
- Base: asphalt treated base
- Panel/base interface: polyethylene sheet
- No. of panels post-tensioned together, comprising a segment: 25 typical, 32 (first segment, and 22 (last segment)
- Post-tensioning method: from mid-segment location
- Total project length: 701 m (2,300 ft) (both sides of a bridge)

- Total no. of expansion joints: 11 (including one at the beginning and one at end)
- Traffic level: Light with few trucks per day; two-way traffic

The project was tested at night during December 2009 and during daytime during January 2010. Overall condition of the project is good. There was tight longitudinal cracking in a few 11 m (36 ft) wide panels. The expansion joint seals were not in good condition. The seals were torn and the expansion joints were filled with debris. The joint width at the time of the daytime testing ranged from 20 to 79 mm (0.8 to 3.1 in). The temperature during the daytime testing was about 13 C (55 F). The pavement condition is shown in Figure 3.

Figure 3: Georgetown PPCP project views

The joint deflections ranged from about 0.102 to 0.508 mm (4 to 20 mils) for the 40.0 kN (9,000 lbf) load level. Basin testing resulted in maximum deflections ranging from about 0.051 to 0.127 mm (2 to 5 mils) for the 40.0 kN (9,000 lbf) load level. The daytime and the night-time testing resulted in similar results. The deflection testing indicates that the LTE at the expansion joints is poor, ranging from about 3 to about 40%.

**I-57, Missouri PPCP Project**

This PPCP project was constructed during December 2005. Pavement details are as follows:

- Panel thickness: variable – 143 mm (5.625 in.) (shoulder edge) to 276 mm (10.875 in.) (at centerline) due to cross-slope
- Panel dimensions: length – 3 m (10 ft); Width – 12 m (38 ft) (two 3.7 m [12 ft] lanes and 1.2 and 3 m [4 and 10 ft] shoulders)

- Base: permeable asphalt treated base
- Panel/base interface: polyethylene sheet
- No. of panels post-tensioned together, comprising a segment: 25
- Post-tensioning method: from expansion joints
- Total project length: 302 m (1,000 ft)
- Total no. of expansion joints: 5 (including one at the beginning and one at end)
- Traffic level: Moderate with high level of trucks per day

The project was tested during daytime during March 2010. Overall condition of the project is good. The expansion joint had been reconstructed during October 2009. The sealant material was in good condition. Many panels exhibited transverse and/or longitudinal cracking. The transverse cracking was generally located within the two driving lanes and did not continue all the way to the edges of the shoulders. The joint width at the time of the daytime testing was about 51 mm (2 in.) and the temperature during testing was about 16 C (60 F). The pavement condition is shown in Figure 4. Spalling was noticed at panel joints at few locations.

Figure 4: I-57 PPCP views (right photo shows reconstructed expansion joint)

The deflection data for the outer wheel path test indicated LTE at expansion joints of about 60%, compared to LTE of 80 to 85% at transverse joints of the adjacent 305 mm (12 in.) thick jointed plan concrete pavement section. The LTE at the expansion joints is considered fair. The joint deflections at the expansion joints ranged from about 0.178 to 0.203 mm (7 to 8 mils) for the 40.0 kN (9,000 lbf) load level. Joint deflections for the 26.69 kN (6,000 lbf) load at the PPCP expansion joints, the PPCP interior keyway joints (between individual 3 m (10 ft) long panels), and at transverse joints of the adjacent JPCP section are shown in Figure 6. The deflections at the PPCP expansion joints are about two times the deflections at the transverse joints of the adjacent JPCP section. Also, the deflection at an interior joint in PPCP segment PPCP 2 was very high as shown in Figure 5. There is a transverse crack in one of the panels adjacent to that joint. The LTE at this interior joint was about 30%, compared to LTE of about 90 to 95% for the other PPCP interior joints tested. The deflection data indicate that there may not be adequate residual prestress at the mid-segment

location in that segment. Basin testing resulted in maximum deflections ranging from about 0.076 to 0.127 mm (3 to 5 mils) for the 40.0 kN (9,000 lbf) load level.

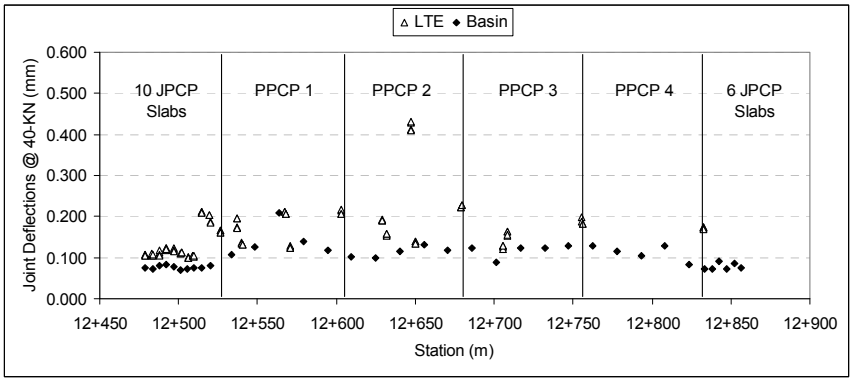



Figure 5: Average deflections at transverse joints (40.0 kN 9,000 lb] load level)

**Tappan Zee Toll Plaza, New York State Jointed Continuous Project**

This is the oldest jointed precast concrete project, constructed during October 2001 and July 2002. Pavement details are as follows:

- Precast pavement system: Fort Miller's Super Slab system
- Panel thickness: 254 mm (10 in.)
- Panel dimensions: length – 5.5 m (18 ft); Width – 3 m (10 ft) (toll plaza drive lanes; 12 lanes)
- Number of panels installed: 1,071
- Base: existing granular base (top 51 mm [2 in.] removed) with 38 mm (1.5 in.) leveling stone dust
- Joints: Doweled transverse joints; longitudinal joints tied
- Total project area: over 33 445 $m^2$ (40,000 $y^2$) (both sides of the toll booths)
- Traffic level: Heavy commuter traffic (New York city area) with large number of trucks per day (eastbound through toll plaza - 72,000 vpd)

The project was tested during the daytime during May 2010. Overall condition of the project is very good. The joint seals are in good condition. Some very tight transverse cracking was noted on a few panels but is not considered to be of concern because of the steel reinforcement in each panel. The deflection testing was conducted along two toll plaza lanes – Lanes 9 and 10. A total of 20 panels were tested. Test data for the outer wheel path test indicated LTE at transverse joints of about 85 to 90%. The joint deflections ranged from about 0.102 to 0.178 mm (4 to 7 mils) for the 40.0 kN (9,000 lbf) load level. Basin testing resulted in maximum deflections of about 0.076 to 0.102 mm (3 to 4 mils) for the 40.0 kN (9,000 lbf) load level. The deflection data indicate a well performing precast concrete pavement project.

**I-66 Ramp, Virginia Jointed Continuous Project**

This jointed precast concrete project, along a ramp exiting off from I-66 highway, was constructed during August to October 2009. Only the outside lane of the two-lane exit ramp was rehabilitated. The precast pavement section, with typical 4.6 m (15 ft) joint spacing, was tied to the inside lane 229 mm (9 in.) thick jointed reinforced concrete pavement (JRCP) with a joint spacing of 18 m (60 ft). It should be noted that it is not considered a good practice to tie different pavement types with different joint spacing. On this project, the precaster's shop drawings indicate the tie-in at the longitudinal joint between the two lanes. Several precast panels were custom fabricated to account for the curvatures in the ramp. Pavement details are as follows:

- Precast pavement system: Fort Miller's Super Slab system
- Panel thickness: 222 mm (8.75 in.)
- Panel dimensions: length – 4.6 m (15 ft) (a few shorter panels to accommodate ramp curvature); Width – 3.7 m (12 ft)
- Number of panels installed: 224
- Base: existing granular base with upto 6 mm (0.25 in.) leveling stone dust
- Joints: Doweled transverse joints; inside longitudinal joint (ramp centerline joint) tied to existing JRCP
- Total project length: 1083 m (3,552 ft)
- Traffic level: Moderate commuter traffic (Washington, DC area) with low level of truck traffic (55,000 vpd)

The project was tested during the nighttime during December 2010 before opening to traffic. Some very tight transverse cracking was noted at 52 panels before opening to traffic, but this is not considered to be of concern because of the steel reinforcement in each panel. About 50% of the cracking in the precast panels was associate with a joint or patch in the adjacent JRCP lane and the tying of the two lanes with different joint spacing and different joint/slab behavior. Such a crack is shown in Figure 6.

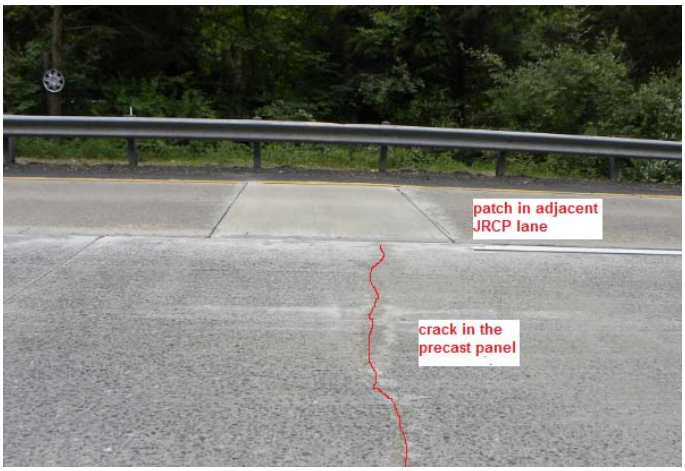



Figure 6: Crack in a precast panel associated with a patch in the adjacent lane

The deflection data for the outer wheel path test indicated LTE at transverse joints of about 85 to 95% for the 45 panels tested. The joint deflections ranged from about 0.203 to 0.381 mm (8 to 15 mils) for the 4082 kg (9,000) lb load level. Basin testing resulted in maximum deflections of about 0.076 to 0.127 mm (3 to 5 mils) for the 40.0 kN (9,000 lbf) load level. The deflection data do not indicate any potential for structural performance issues for the precast concrete pavement project.

**I-295, New Jersey Jointed Intermittent Repair Project**

The precast panel were installed during late 2007 to mid-2008. Panel details are as follows:

- Precast pavement system: Fort Miller's Super Slab system
- Panel thickness: 222 mm (8.75 in.) (existing JRCP thickness – 229 mm [9 in.])
- Panel dimensions: length – variable (2.4, 3.0, 3.7 m [8, 10, 12 ft]); Width – 3.7 m (12 ft)
- Number of panels installed: 277
- Base: existing sandy granular base
- Joints: Doweled transverse joints; longitudinal joints not tied
- Traffic level: Heavy freeway traffic with heavy truck volume (140,000 vpd)
- Existing JRCP joint spacing: 24 m (78 ft)

The project was tested during the daytime during April 2010. A total of 20 precast panels and 20 existing pavement slab panels (adjacent to each precast panel) were tested. Overall condition of the project is good. However, the precast panels exhibited settlement (as measured along the outside longitudinal joint) of upto 25.4 mm (1 in.). This slab settlement also corresponded to a transverse crack in the leave side of the existing slab panel, about 2.4 to 3.7 m (8 to 12 ft) from the end of the repair. The pavement condition is shown in Figure 7. The joint sealing in the repair area was not consistent as some of the joints appeared not to have been sawed or sealed. Also, the dowel slot grout in the joint had separated from the joint face and many of the repair joints.

Figure 7: Views of the I-295 precast panels

The deflection data for the outer wheel path test indicated LTE at transverse joints of the precast panels of about 65 to 85%, compared to LTE of about 70 to 85% in the 50-year old existing JRCP slab panels. The joint deflections at the precast panel joints ranged from about 0.102 to 0.178 mm (4 to 7 mils) for the 40.0 kN (9,000 lbf) load level, compared to about 0.102 to 0.127 mm (4 to 5 mils) for the existing JRCP joints. Basin testing resulted in maximum deflections at the precast panels of about 0.102 to 0.178 mm (4 to 7 mils) for the 40.0 kN (9,000 lbf) load level, compared to 0.076 to 0.127 mm (3 to 5 mils) at the interior of the existing JRCP panels. The deflection data indicate a well performing transverse joint system for the precast panels. However, the panel settlement indicates that a better technique may be warranted to upgrade the bedding when granular bases are used as these bases cannot be effectively re-compacted as part of the precast panel installation process.

## DISCUSSION OF FIELD TEST DATA

Only a limited summary of the data related to performance of older and recently constructed precast concrete pavements are presented in this paper. Evaluation todate of the field data, as summarized here, indicate that precast pavements that are designed and installed well have the potential to provide long term service for both repair and continuous applications. Both the PPCP and the jointed precast pavement systems can be considered as equally viable candidates for rapid rehabilitation of concrete as well as asphalt pavements and for overlay applications. The service life expectations in the US for continuous applications would be about 40 years, similar to what is currently being designed for new conventional concrete pavements. For intermittent repair applications, both the proprietary systems, such as the Fort Miller's Super Slab system, and generic systems, such as the Michigan system, appear to have the potential to provide service life ranging from 10 to 20 years.

The following specific observations can be made on the basis of the evaluation of the performance of the field tested precast concrete pavements:

- PPCP systems
    - The wider joint opening of the expansion joints is resulting in variable LTE. Limiting the length of the post-tensioned segments may allow reduction in the expansion joint width and improve the LTE at these joints. Based on the experience of the authors, good LTE (greater that 85%) is necessary in new construction for highways that carry higher truck traffic.
    - Even though some PPCP panels exhibited cracking, this cracking is not considered of concern as the reinforcement used in each panel and the "bonded" prestressing will keep the cracks tight. However, it is necessary that there be adequate prestress available in the panels in the mid-segment area. The low LTE measured at an interior joint at the Missouri project, between adjacent 3 m (10 ft) long panels, will be analyzed further to determine the cause for the low LTE.

  - The spalling over the reinforcement indicates a need to re-design the reinforcement details. Reinforcement should be epoxy-coated and at least 51 to 64 mm (2 to 2.5 in.) below the panel surface to minimize the potential for surface spalling.
  - The Virginia I-66 project indicated higher deflections and loss of support at the expansion joints. This may be a result of night-time slab curling over the full length of the post-tensioned segments and/or the use of the stone dust bedding/leveling material.
  - The Delaware project shows potential for use of PPCP systems for specific applications, such as, high volume intersections.

- Jointed systems for continuous applications
  - The jointed Super Slab system for continuous applications has performed well. The oldest project, the Tappan Zee project, and the Minnesota project indicate good structural performance – low deflections at joints and at panel interior and good LTE values.
  - Similar to the cracking in PPCP panels, the tight cracking observed in the jointed panels at the Tappan Zee and the Virginia ramp projects are not considered of concern because of the reinforcement used in these panels. For long-term service, it is important that the top layer of the reinforcement in the panels be at least 64 mm (2.5 in.) below the panel surface.
  - For the Super Slab system, the joint sawing detail needs improvement. Joint sawing that leaves a sliver of the dowel slot grout material (possibly not bonded to the joint face) at the surface can result in non-effective joint sealing, begging the question whether these joints should be sawed and sealed or not.
- Jointed systems for repair applications
  - For full-depth or full-slab repair applications, there is a critical need to ensure that there is adequate LTE at joints and that there is good support under the repair, irrespective of whether the repairs involve precast panels or cast-in-place concrete. The Super Slab system, with dowel slots at the bottom, appears to provide good load transfer at joints. The systems incorporating dowel slots at the surface, similar to conventional dowel bar retrofit method, are also capable of providing good LTE at joints. However, these systems require care with the installation of the dowel slot patch material.
  - The use of dowel bar caps is strongly recommended to minimize failure of the dowel bar slot patches, especially when the repairs are used for repairs of joints in JRCP.
  - The panel settlement at the New Jersey I-295 project indicated that more attention needs to be paid to improve the bedding support under precast panels placed over “disturbed” granular bases. This is weakness in the full-depth repair technique that is a major cause of the failure of cast-in-place full depth patches. The use of precast panels is not a warrant to pay any less attention to making sure the support

under the precast panels is not compromised during existing concrete removal and panel installation.

Overall, there does not appear to be any concern about the quality or the durability of the concrete used for the precast panels. For most of the precast applications, repair or continuous, there is a requirement to grind the surface of the completed project and therefore any joint elevation differences during placement of the precast panels can be addressed and a desirable surface texture can be obtained.

To summarize, the precast concrete pavement technology for rapid repair and rehabilitation of high volume highways is an evolving technology, with many highways agencies having implemented it for production use, some investigating it, and some waiting on the sidelines before moving forward with it. Some of the findings presented in this paper and additional work underway in the SHRP 2 Project R05 will provide improved guidance for the application of precast concrete pavements for rapid repair and rehabilitation of distressed pavements.

## ACKNOWLEDGEMENTS

Work reported in this paper was conducted as part of the US SHRP 2 Project R05: Modular Pavement technology. The SHRP 2 Project Manager for Project R05 is Mr. James Bryant. His support during the conduct of this greatly appreciated. The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented. The contents do not necessarily reflect the official views and policies of SHRP 2. This paper does not constitute a standard, a specification, or a regulation. The support of all agencies whose projects were tested and reported in this paper is sincerely acknowledged and greatly appreciated. This support included gathering of project information, site access, traffic control, as well as direct testing support by several agencies.

## REFERENCES

Hall, K. and Tayabji, S., 2008. *Precast Concrete Panels for Repair and Rehabilitation of Jointed Concrete Pavements*, TechBrief No. FHWA-IF-09-003, Federal Highway Administration, Washington, DC.

Merritt, D. and Tayabji, S., 2009. *Precast Prestressed Concrete Pavement for Reconstruction and Rehabilitation of Existing Pavements*, TechBrief No. FHWA-IF-09-008, Federal Highway Administration, Washington, DC.

Tayabji, S. Buch, N. and Kohler, E., 2009. *Modular Pavement Technology: Interim Report*, report prepared by Fugro Consultants, Inc. for SHRP 2, SHRP 2 Project R05, Transportation Research Board, Washington, DC.

# SHRP 2 Project R23 - Development of Guidelines for the Design and Construction of Long Life Pavements Using Existing Pavements

N. C. Jackson[1], Joe P. Mahoney[2], Jason Puccinelli[3]

[1] Nichols Consulting Engineers Chtd, 2714 Hibiscus Ct SE Olympia WA 98513 Ph/FAX 360 923-9359 email: Newtoncj@aol.com
[2] William M. and Marilyn M. Conner Professor, University of Washington, 133D More Hall, Box 352700 Seattle Washington 98195-2700, ph 206 685-1760 fx 206 543-1543, jmahoney@u.washington.edu
[3] Nichols Consulting Engineers Chtd, 1885 S Arlington Ave. Suite 111, Reno Nevada 89509, ph775 329-4955 fx 775 329-5098, jason@nce.reno.nv.us

## ABSTRACT

Guidelines are being developed to help State Highway Agencies (SHAs) in the selection and design of long life pavements using existing in-place pavements. There is currently very clear design guidance and procedures for designing new long life pavements in terms of new or reconstruction. Guidance that does consider the use of existing pavement is typically directed toward pavement rehabilitation with some determinate life expectancy (ARA, 2004). There is no clear and specific guidance for designing and building long life pavements using existing pavements in place. Using the existing pavement in the design and construction of long life pavements offers numerous advantages both in facilitating rapid renewal of the roadway and in the sustainability aspect of utilizing existing materials. Using the existing pavement significantly reduces the mass movement of materials in and out of the project, and in many cases can facilitate the staging of traffic during construction as part of a cost effective rapid renewal strategy.

The approaches found for designing long life pavements using existing pavements presented in the R-23 guidelines are similar to what many agencies currently use, with a few additions from international practice. It was discovered that most agencies use one or more of the approaches contained in the guidelines but have abandoned or no longer consider several of the approaches because of past construction problems. It was found that the assessment of the existing pavement and use of the appropriate design and attention to construction details are critical features of the process. An overview of the form of the guidelines and the detailed information contained in those guidelines are presented, including project assessment, design guidance, critical construction aspects, and guide specifications.

## INTRODUCTION

This project falls within the SHRP 2 Renewal area, which focuses on improving the ability of the industry to design and construct long-lasting highway projects quickly

with minimal disruption to the traveling public. Key components to achieving these objectives include the application of innovative methods and materials for preserving, rehabilitating, and reconstructing pavements. Specific to the R-23 project, construction costs and time can be greatly reduced if the existing pavement can be used in-place as part of the rehabilitation solution. As such, the goal of this project is to develop reliable procedures that identify when existing pavements can be used in-place and the methods necessary to incorporate the original material into the new pavement structure while achieving long-life. The following components were considered in this project:

- Identification of alternatives for using existing pavements in-place for rapid renewal.
- Analysis of advantages and disadvantages for each approach under different site conditions.
- Development of detailed criteria on when an existing pavement can be used in-place, with or without significant modification.
- Identification of practices and techniques available to construct pavements with the above characteristics.
- Determination of the optimal methods to integrate the renewed pavement with adjacent pavements and structures.

The ultimate deliverables required for this project are:

- Design Guidelines
- Construction Procedures
- Guide Specifications

## APPROACHES FOR USING EXISTING PAVEMENT ON RENEWAL PROJECTS

A literature review was conducted as well as person-to-person contact both nationally and internationally. This was followed with a detailed questionnaire which identified the range of renewal approaches and those specifically used by agencies for long life pavements.

The approaches identified and considered were:

- Asphalt over rubblized jointed concrete pavement,
- Asphalt over cracked and seated jointed plain concrete pavement,
- Asphalt over broken and seated jointed reinforced concrete pavement
- Asphalt over sawed, cracked and seated jointed reinforced concrete pavement,
- Asphalt over crushed and shaped asphalt pavement,
- Asphalt over reclaimed or milled asphalt pavement (*major reconstruction*),
- Asphalt over asphalt pavement (*major reconstruction*),
- Asphalt over continuously reinforced concrete pavement,
- Unbonded concrete overlay (whitetopping) over asphalt pavement,

- Bonded concrete overlay of jointed or continuously reinforced concrete pavement,
- Unbonded concrete overlay of jointed or continuously reinforced concrete pavement,
- Asphalt or concrete replacement/inlay of existing lane,
- Precast concrete slab replacement.

There is no national or international standard in defining what a long life pavement is, nor what the actual service life should be. The pavement design life used by the various SHAs was found to range from 20 to 50 years. The latest European guidance indicates that a service life of 30 to 40 years is considered long life (FEHRL Report 2004/01). The long life requirement for this project was stated as providing a minimum of 50 years service.

The potential use of these approaches for application in long life pavements was considered based on the typical performance reported from the various national and international agencies. Since most agencies contacted either did not design for, nor had pavements with 50 years of performance, additional analysis was required to support selection of potential long life approaches. The performance of the appropriate experiments in the Long Term Pavement Program (LTPP) was analyzed to identify or confirm which approach was providing reasonably long term performance. In addition, performance estimates were made based on current and emerging design programs like the Mechanistic-Empirical Pavement Design Guide (MEPDG), (ARA, 2004) and the PerRoad software. The most recent views on long life pavement design and performance criteria was also considered for both flexible and rigid pavement approaches (Newcomb et al) (Jordan et al) (D. Harrington et al) (FEHRL Report 2004/01) (FEREL Report 2009/01).

Based on this analysis, the advantages and disadvantages for each approach under different conditions were enumerated. Criteria were developed on when an existing pavement can be used in-place, with or without significant modification.

The final list of approaches that were selected which could reasonably be expected to provide long life (50 years service) using existing pavements was narrowed down to the following.

- Asphalt over cracked and seated, or rubblized jointed concrete pavement,
- Asphalt over sawed, cracked and seated jointed reinforced concrete pavement,
- Asphalt over reclaimed or milled asphalt pavement (*major reconstruction*),
- Unbonded concrete overlay (whitetopping) over asphalt pavement,
- Unbonded concrete inlay of asphalt pavement,
- Unbonded concrete overlay of jointed or continuously reinforced concrete pavement.

## DESIGN GUIDELINES

The information was assembled into a set of design tables based on the existing pavement type, surface condition, soil support and other pavement assessment information. A sample portion of a typical decision table is shown as Table 1.

Table 1 Sample Portion of Existing Flexible Pavement Actions

| Specific Distress or Description | Decision (Is this distress present?) | Renewal Pavement Type Option | Action | Design Resources |
|---|---|---|---|---|
| Stripping | Yes | Flexible | If stripping is found through all layers, pulverize pavement structure full-depth followed by a thick AC overlay. | Pulverize and use residual material as untreated base. Apply AC thickness from Table DR-1a-2. |
| | | | | Pulverize and treat residual material with emulsion or foamed asphalt resulting in a treated base. Apply AC thickness from Table DR-1. |
| | | | If stripping is found in specific layers, remove AC to maximum depth of stripping followed by a thick AC overlay. | Use Table DR-1 and the subgrade $M_R$ to determine total depth of AC thickness then subtract the thickness of AC removed due to stripping. |
| | | Rigid | Place unbonded PCC overlay. If grade limits require, mill existing pavement. AC overlay over stripped pavement may be required, to stabilize HMA. | Apply unbonded PCC overlay. Use Table DR-2b for JPCP thickness determination |
| | No | -- | Continue to Environment Cracking. | -- |

The study results revealed that specific minimum thicknesses were required to assure long life performance. In the design guidelines, it is assumed that an agency will independently compute the design thicknesses based on their specific design procedures. To ensure that there are reasonable expectations regarding the design thicknesses required to provide long life performance, a set of design thickness tables were developed that could be applied to the various approaches. The design thicknesses have been determined by use of: (1) performance data obtained in this project, (2) design runs using 1993 AASHTO Design Guide for Pavement Structures along with updated modifications (Peters-Davis and Timm, 2009), (3) the MEDPG, and (4) Per Road. Additional thickness verification will be done before the study is complete. The resulting tables were also compared to design tables being developed by various national and international agencies.

It should be noted that in most cases the design thickness developed from long life design approaches are thinner than that indicated by standard application of the 1993 AASHTO design procedure for both flexible and rigid pavements. This is largely

caused by the consideration of low working stress ratios in M-E design for rigid pavements and a similar effect based on endurance limits for flexible pavements.

Table 2 Sample Thickness Design Table
(HMA over Rubblized PCC Pavement)

| ESALs (millions) | HMA Thickness (inches) |
|---|---|
| <10 | 8.0 |
| 10-25 | 9.0 |
| 25-50 | 9.5 |
| 50-100 | 10.0 |
| 100-200 | 11.0 |

In addition to the design guidelines, the project deliverables include construction procedures as well as guide specifications. It is the authors' view that for a project to actually provide long life performance, the agency must follow both the best design practice as well as construction procedures. For this reason both the general design guidance as well as construction procedures will be presented in the form of a best practices document that goes with the design guidelines, construction procedures, and guide specifications. The best practices discussions will be presented for the different pavement types and approaches considered in the design guidelines.

An agency has to consider traffic and staging implications for the selected approaches for specific project conditions. They also need to consider life cycle cost comparisons between the approaches considered based on the agencies policy. Maintenance considerations for long life pavements should also be considered. These discussions, though very important to the selection of the best approach for a given project, must take place outside of the basic design guidelines as they are dependent upon unique agency policy. Detailed guidance will be provided on traffic staging, life cycle cost analysis, required maintenance for long life pavements and life cycle assessment considerations for informational purposes only.

There is an extensive amount of information associated with the design guidelines that should be considered for each approach. An interactive program was created which will simplify applying the guidelines as well as providing an interactive platform on which to house and access the detailed information associated with the guidance on the various approaches.

## R23 GUIDELINES PROGRAM

The interactive program was developed in Adobe Flash which provides an interactive program that can be easily web enabled. The primary reasons for selecting the Flash platform include:

- The Adobe Flash player is nearly ubiquitous on PC, Mac, and many mobile devices, which promotes wider use of the application.

- The software package can be deployed in a variety of environments with ease including standalone CD-ROM, installable executable package, client-side web-page, and hosted web-page.
- It is developed using Adobe's latest freely available technologies (ActionScript 3.0, Flex, AIR) providing assurance that it will be compatible for the foreseeable future.

Other technologies that were considered in this application development included .NET, Java, HTML 5 with Javascript, and a web/server architecture using PHP. The Adobe Flash platform met the program needs best.

The program provides a platform for placing a lot of information in the form of PDF files which are linked in the program by tabs and within the PDF documents.

The program requires the user to move through five steps in applying the guidelines:

- Specify existing and proposed section information,
- Specify existing pavement condition,
- Confirm section design parameters,
- Select renewal strategy,
- Recommended section design, best practices.

**Steps 1 and 2** require basic project information and the assessment of the existing pavement and supporting soil. A project assessment guide has been developed and is imbedded in the program as a searchable PDF file. An example of the introduction screen is shown in Figure 1.

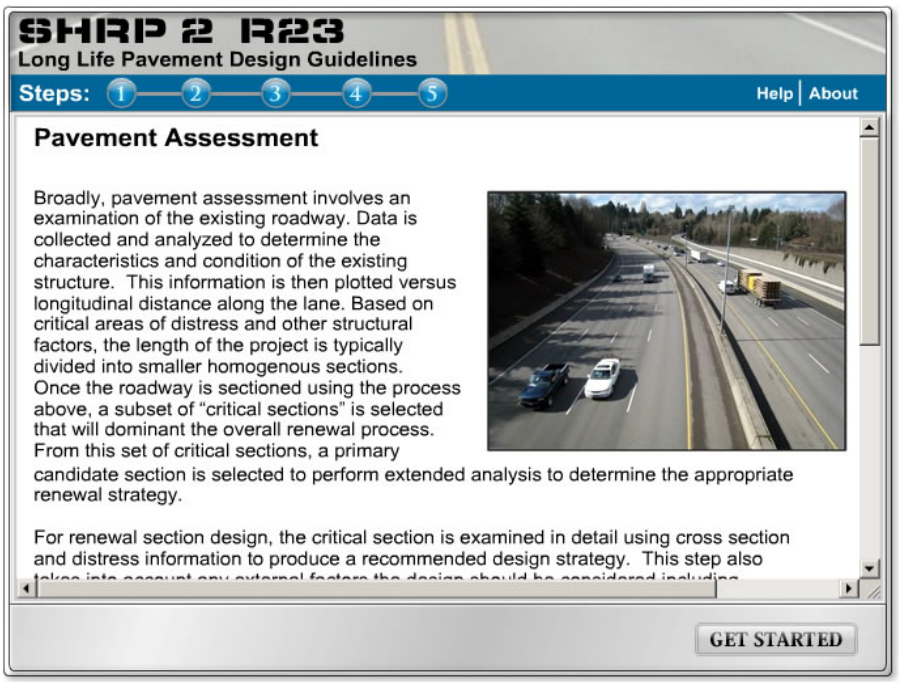



Figure 1 Introduction Screenshot

The program leads the user through the basic information needed to help select the appropriate approach for the site and pavement conditions and design the new pavement sections.

The introduction screen is followed by the project assessment screen where the user is asked to provide information on the general geometrics of the existing pavement and design information for the new roadway. There are three tabs that must be completed. The first tab contains the project identification information. The second tab requests general layer information on the existing pavement and the third for design information for the reconstructed roadway including lane addition and potential height restrictions. An example of that screen is shown in Figure 2. Only the tab for the existing pavement layers is shown for this step.

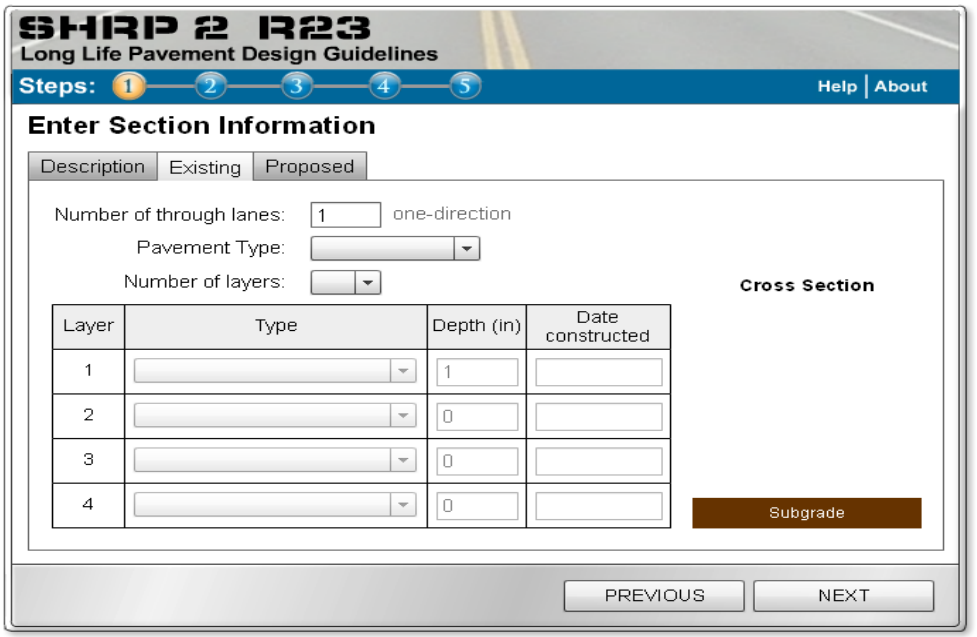



Figure 2 Project and Pavement Layer Information Screenshot

**Step 2** requests the user to provide the pavement condition information from a pavement condition survey that would have been performed for the project. The required pavement information is illustrated in Figure 3. For flexible pavements the program also considers if stripping is present and what layers are involved as well as fatigue cracking, top down cracking, and environmental cracking. For rigid pavement it considers if the panels are cracked and if faulting is present as well as differential joint movement for jointed pavement.

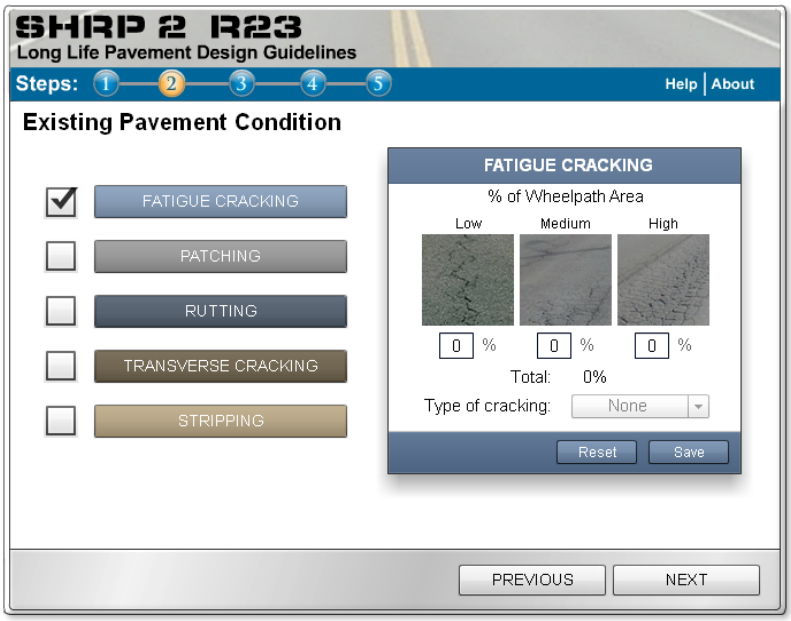



Figure 3 Sample Pavement Condition Screenshot

**Step 3** requests the user to confirm the prior information entered. In confirming that the information entered is correct, it also implies that there is no other distress present which eliminates having to go through each distress category in a normal decision tree. For instance, where the user indicates fatigue cracking and patching is present (but no other distress), they are also confirming that they don't have stripping or other distresses. This confirmation simplifies the steps that one would need to consider in a standard decision matrix or flow chart. An example of this screen is shown in Figure 4.

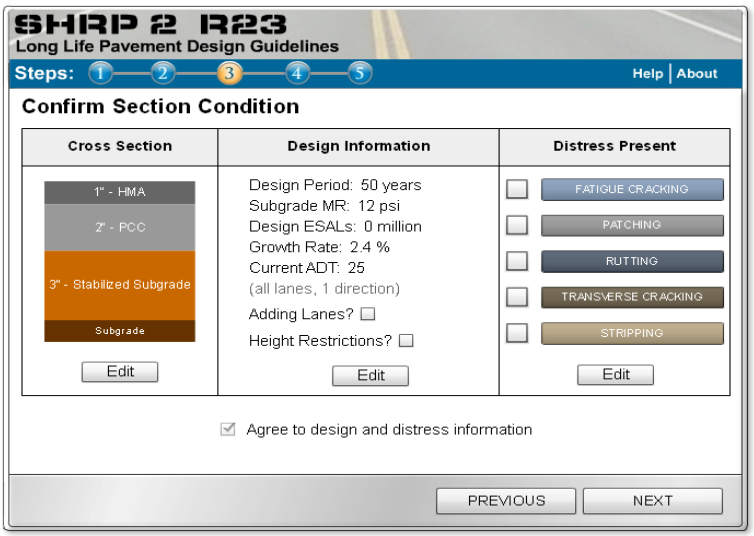



Figure 4 Sample Confirmation Screenshot

In **Step 4** the user is asked to select a renewal option that they want to consider which is followed by an additional decision step that guides the user to the most appropriate approach for the type of pavement the user wants to consider. Additionally, a design thickness is determined based on a predetermined subgrade and base condition; however, the user will have the option to modify those conditions if they wish. A typical screen for this step is shown below in Figure 5.

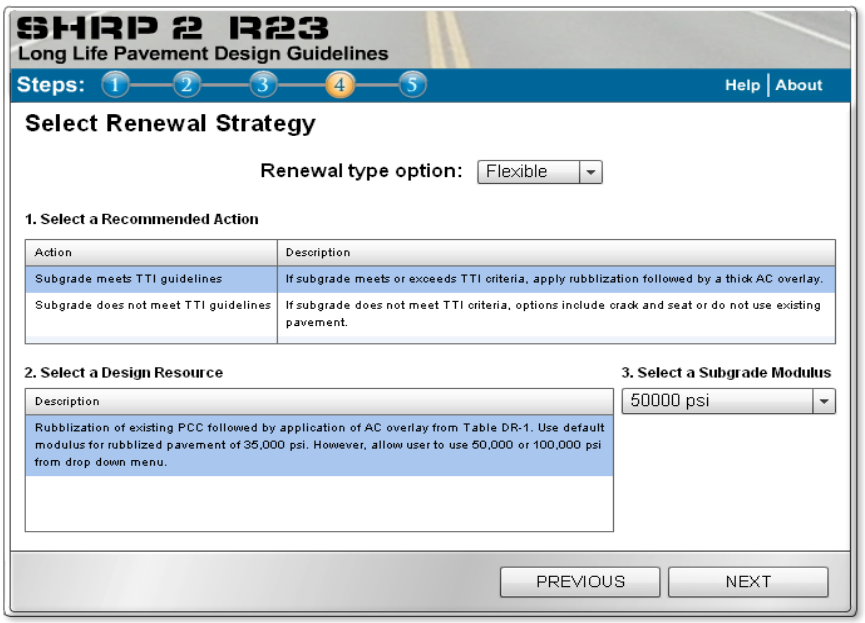



Figure 5 Sample Approach Selection Screenshot

**Step 5** shows the recommended approach and potential thickness required to provide long life performance (as can be seen in Figure 6).

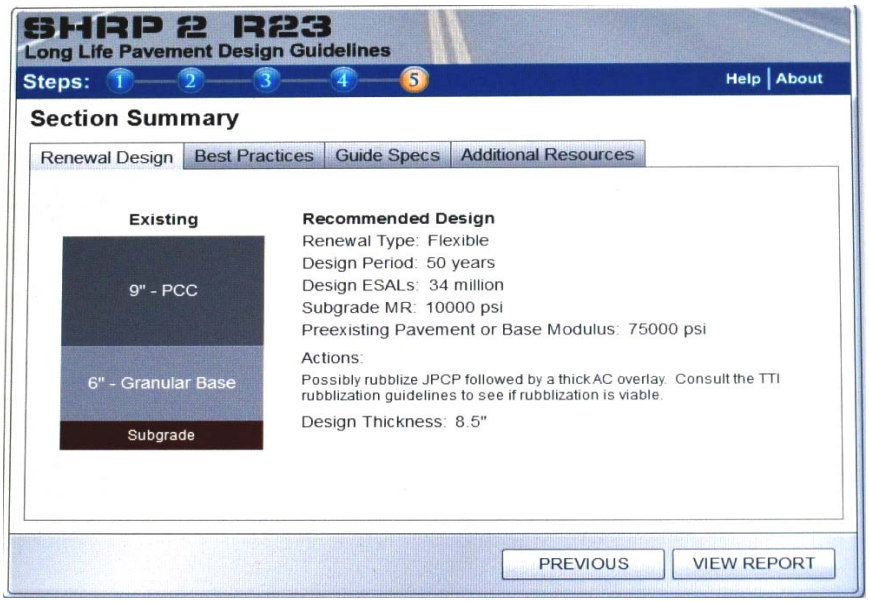



Figure 6 Sample Summary Screenshot

Most of the screens provide access to the large amount of information assembled in this project. This final screen includes tabs that direct the user to appropriate best practices, construction specifications, and additional information. The best practices guidance includes detailed descriptions of the different approaches, construction techniques, and design considerations. The Specifications tab contains guide specifications as well as quality control and assurance considerations. The additional information tab provides information on traffic and project staging considerations, life cycle costs analysis, maintenance requirements, emerging technology, and life cycle assessment considerations. The traffic and staging section will provide a general discussion of traffic considerations, staging, and construction analysis tools that are available. The life cycle costs information will provide guidance on the life cycle cost analysis that should be performed to determine the best approach for a specific project considering the standard material costs over time but also the staging, traffic control and user costs associated with these types of projects. The maintenance section will provide a discussion of the maintenance considerations that go hand in hand with long life pavements. The emerging technology information will discuss some of the emerging processes such as modular pavement systems, thin CRC pavements, etc that are under development in the rapid renewal area that may have application in the future with more usage and field performance experience. As noted earlier in the paper there is also a detailed project assessment manual that is included in the guidelines but that document is accessed in the first screen to guide the user at beginning of the process.

The guidelines and their development will be fully described in the final report that is being prepared for SHRP 2 R-23 project. The Guidelines will be presented in the report in a set of decision tables with all of the supporting information. The interactive program described above will be provided on a CD for agency use. The interactive program and the final report will contain much of the same information, however the interactive program will provide that information in a linked format so that the user will be directed quickly to the information they need to consider for their particular project and site conditions.

## REFERENCES

Applied Research Associates, Inc., *Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures*, Final Report, NCHRP Project 1-37A, National Cooperative Highway Research Program, Transportation Research Board, National Research Council, Washington, DC, 2004.

D. E, Newcomb, R Willis, D.H. Timm, *Perpetual Asphalt Pavements A Synthesis*, Asphalt Pavement Alliance IM-40 2010, AsphaltRoads.org.

D Harrington et al, *Guide to Concrete Overlays Sustainable Solutions for Resurfacing and Rehabilitating Existing Pavements*, Second Edition, September 2008, National Concrete Pavement Technology Center, Iowa State University, Ames, Iowa.

FEHRL Report 2004/01, *ELLPAG Phase 1Report*, *A guide to the Use of Long-Life Fully-Flexible Pavements*, Forum of European National Highway Research Laboratories, Brussels, Belgium, 2004

FEHRL Report 2009/01, Making Best Use of Long-Life Pavements in Europe ELLPAG Phase 2, *A guide to the Use of Long-Life Rigid Pavements*, Forum of European National Highway Research Laboratories, Brussels, Belgium, 2009

Peters-Davis, K. and Timm, D. (2009), "Recalibration of the Asphalt Layer Coefficient," Report 09-03, National Center for Asphalt Technology, Auburn, AL, August 2009.

RW Jordan, C Coley, HM Harding, I Carswell, KE Hassan, *Best Practice Guide for Overlaying Concrete*, Road Note 41, Transportation Research Laboratories 2008, Wokingham Berkshire, United Kingdom.

# Behavioral Housing Search Choice Set Formation: A Hazard-Based Screening Model of Property Value and Work Distance

Taha H. Rashidi[1] and Abolfazl (Kouros) Mohammadian, Ph.D.[2]

[1]Corresponding Author, Ph.D. Candidate, Department of Civil and Materials Engineering, University of Illinois at Chicago, 842 W. Taylor St. Chicago, IL 60607, Phone: 312-996-0962, Fax: 312-996-2426 Email: thosse2@uic.edu

[2]Associate Professor, Department of Civil and Materials Engineering, University of Illinois at Chicago, 842 W. Taylor St. Chicago, IL 60607, Phone: 312-996-9840, Fax: 312-996-2426, Email: kouros@uic.edu

## Abstract

Residential location search has been an interesting topic to both practitioners and researchers. The housing search process starts with an alternative formation and screening practice. At this level households evaluate all potential alternatives based on their lifestyle, preferences, and utilities to form a manageable choice set with limited number of plausible alternatives. This paper attempts to study this screening and filtering practice in more details and to develop a modeling framework that can replicate this process.

The average desired property value and work distance are considered as two potential attributes, among others, which are examined by the household to evaluate feasible housing alternatives. It is postulated that alternatives will be included in the choice set only if the average property value meets the household budget and the work distance satisfies the household distance threshold. The search process will then continue until the choice set is complete. A hazard-based model is formulated that can model the desired property value and average work distance together. Several household socio-demographic attributes from eight waves of the Puget Sound Transportation Panel (PSTP) were utilized for model estimation along with built environment variables, characteristics of the supply side of the market, and several other economic indicators. The approach presented in the paper provides a remedy for the large choice set problem typically faced in discrete choice modeling."

# 1. Introduction

Residential location search has been an interesting research topic in many fields including transportation, urban planning, geography, economics, and other relater disciplines. Metropolitan planning organizations, real estate companies, insurance companies and financial institutions are also among the non-academic organizations that are interested in having an accurate housing search model. Since the early introduction of the discrete choice paradigm, the individual's alternative selection behavior has been primarily modeled using this approach (McFadden 1974). The prediction potential and the accuracy of a discrete choice model parameter estimation process itself are highly dependent on the choice set composition. There has been great effort to improve the theory and applications of the discrete choice models, however, researchers seldom examine the importance of the choice set composition on parameter estimation (Ben-Akiva and Lerman 1985 and Timmermans and Golledge 1990). Even though recent advances in computational power allows researchers to work with large datasets, in practical applications, the difficulty of handling many alternatives makes it necessary to reduce the number of alternatives in the choice set into some manageable size. Accordingly, in the literature, there has been two extreme approaches for selecting the set of alternatives; first, randomly selecting a finite number of alternatives from the *universal choice set*, as it is defined by Ben-Akiva and Lerman (1985), and second, considering all plausible alternatives (Salomon and Ben-Akiva 1983 and Thill and Horowitz 1991). It can be shown that both approaches can raise some concerns. Although inclusion of all possible alternatives may seem to be a conservative approach, nonetheless, it can be unrealistic as it assumes decision makers have perfect knowledge about all alternatives. This approach can result in assigning non-negative selection probability to some alternatives that otherwise may not be known or be available to the decision maker. On the other hand, random selection of few alternatives for the choice set by stratified sampling or other similar approaches can result in bias and possibly inaccurate parameter estimation.

In addition to the two abovementioned approaches, there are other methods to address the choice set formation issue. Non-econometric methods, either heuristic or non-heuristic have been used in the literature, in which alternatives are evaluated by certain criteria for being included in the choice set. For instance, Arnold *et al.* (1983) utilized a preference ranking method to rank the shopping destination where only stores ranked as *excellent* are included in the individual choice set.

Probability of selecting an alternative to the choice set can also be formulated jointly with the main discrete choice model. In this approach, the error term of the random utility of being in a choice set can be correlated with the discrete choice

model's error term and the two random utility models can be jointly estimated (Thill 1992). Regardless of the type of the method which is used for choice set composition, it is critical to employ an appropriate filtering/screening method. As asserted by Manski (1977), accurate estimation of parameters of discrete choice models and correct prediction of choices by them is conditional on correct information about the choice sets. One can argue that the estimation of the parameters can be questionable if deficient information about potential alternatives and the choice set is utilized in the model. Nonetheless, as long as the individual's true alternatives are considered in the choice set, under specific conditions, a logit model can accurately estimate the parameters, as it has been discussed by McFadden (1978).Similarly, Manski mentioned in his paper that under specific conditions, the approximation of the true choice set by a systematically selected subset of it does not put at risk the consistency of the choice model estimates (Manski 1977). However, this will not be true in the case of mis-specified choice set that can result in erroneous parameters estimates, even if such effects are marginal (Lerman 1985). For instance, if an alternative has a very low possibility of being considered by the decision maker, its inclusion in the choice set will have minimal effects on the choice prediction results. Nonetheless, behavioral insight of the methods of choice-set formation may considerably improve the accuracy of the model and parameter estimates, while the absence of this insight can be problematic (Burnett and Hanson 1979 and Burnett and Hanson 1982).

This study aims to introduce a behavioral method for housing search choice set formation. The residential location choice process starts with an alternative evaluation and screening practice. Intuitively, people first scan their alternatives and then filter them based on their priorities, lifestyle, preferences, and utilities. Finally, among the filtered alternatives the most desired option with the highest utility is selected. The screening stage in which the choice set is selected is modeled in this study in which intra- and inter-household effects are considered. While there are several factors affecting the selection of housing alternatives (e.g., property value, commute distance, school quality, safety, tax rate, etc), in order to show the practicality of the approach, only two most influential factors on household residential relocation decision were considered in the screening process model of this study. These include household average property value and work distance. Many studies have considered either one of these two variables separately and few considered the both. Housing unit price has usually been modeled in a hedonic format while accounting for market supply and demand in the model. Travel distance is also typically obtained from the transportation network simulation module.

The rest of the paper is organized as follows. First, a brief literature review is presented and the study approach is discussed. The datasets used in this study are then explained and their key variables are discussed. Model derivation and the mathematical formulations of the system of equations are presented next. Following

that, experimental results of different steps of the parameter estimation process are presented. Finally, conclusions and future research directions are discussed in the final section.

## 2. Background and Study Approach

Price and commute distance are typically considered as two dominant factors influencing the residential location search (Kim 1992). Similarly, commute distance is commonly used in the job location search models. These two variables are commonly modeled in the literature separately (Kim 1992 and Clark *et al.* 2003). In a interdisciplinary study Waddell (1996) modeled the interactions between workplace, residential mobility, tenure, and location choices in a Nested Logit framework. In transportation field, commute distance has been one of the most commonly used measures in various applications. The application of housing price variable in transportation applications has been limited mainly to land-use models but is commonly used in other disciplines such as marketing and economics.

Similar to other spatial choice models, housing search process deals with numerous plausible alternatives. While number of alternatives in a mode choice problem or other similar choice situations is limited, other applications such as activity location, destination choice, or housing search decisions deal with large-size choice sets which can complicate the modeling process. To ease the choice set generation process, some studies have suggested considering all possible alternatives for each individual and assigning a non-negative selection probability to each alternative (Richards and Ben-Akiva 1975 and Domencich and McFadden 1975). Such an approach is not suitable for a housing search choice problem where numerous correlated alternatives exist. In such cases, an accurate choice set generator model is needed.

The choice set configuration problem can be traced back to early applications of discrete choice models. Ben-Akiva and Lerman (1985) proposed the stratified sampling procedure to generate the alternative set and showed the efficiency of that approach. Srinivisan (1987) introduced three levels in screening the alternatives and finding the final choice set: awareness set, evoked set and choice set. He borrowed the term evoked set from another study by Howard who originally introduced it in 1963. According to his model, the awareness set consists of all alternatives the consumer is aware of. This set is then filtered to the evoked set which is a subset of the awareness set and consists of those alternatives that meet certain criteria for further consideration. Finally, the choice set is a subset of the evoked set in which there are very few alternatives including the final choice which is the immediate group of alternatives before making a decision. Shocker *et al*. (1991) employed the

term consideration set for evoked set which was originally introduced in a study by Wright and Barbour in 1977.

Other than the different definitions for the choice set, various solutions have been introduced to deal with the choice set problem. Willumsen and Ortuzar (2001) listed three ways for tackling the choice set problem available in the literature:

1- Rule-based heuristic or deterministic choice set generation methods,
2- Simply asking the individuals in the survey regarding their preferences about the feasible alternatives,
3- Application of random choice sets.

Lerman (1984) proposed a two step process; initially, the probability distribution function across all possible choice sets is defined, then, conditional on the specific choice set, the choice probability of each alternative in the choice set is defined. This study focuses on the first step of the two-step approach by introducing a hazard-based approach that is shown to result in generating a realistic random choice set. The second step of this two-step approach will remain as a future research task.

An extensive curve fitting exercise suggested that the average property value and commute distance follow certain distributional forms. In this study, property values with log-logistic distribution and work distance with Weibull distribution are used to estimate the probability density functions of the possible choice sets. Knowing that, the choice set probabilities are modeled by using two continuous variables of price and work distance. Coupling these probabilities and a discrete choice model that can predict the probability of each alternative in the choice set can be formulated in a continuous-discrete format (Bhat 2005). Developing such a framework is currently underway. Alternatively, one can use the modeling approach presented in this study to generate the choice set probabilities and couple them with a traditional discrete choice models similar to the ones suggested by Manski (1977) or Fotheringham (1983).

## 3. Data

Puget Sound Transportation Panel (PSTP) was used as the primary source of data that is used in this study. The PSTP is a panel data for Seattle Metropolitan Area (Murakami and Watterson 1992) .Nonetheless, only household observations of the King county area are used for the modeling practice due to need for auxiliary data (e.g., property values, etc) that were not available for other counties. The last eight waves out of the existing ten waves in the PSTP covering the last decade of the 20th century plus the two first years of the twenty-first century are included in this study. The PSTP provides a wide range of variables in the household level including household socio-demographic attributes. Furthermore, person level attributes such as home to work distances are also provided in the PSTP.

As noted earlier, there are two dependent variables that were used for screening the household choice; work distance and price. Average household work distance is directly obtained by running some queries on PSTP data. The property value, on the other hand, is not provided in the PSTP whereas the census tract number in which the household resides is the smallest geography of the household residential location. Land values and house prices are mainly attained by county assessment departments. This information is mainly provided for property tax preparation purposes and is cast away after a decade or so. The only county in the Seattle area for which historical parcel level land values and housing unit prices are available as far as early 90s is the King county. Therefore, the PSTP data was filtered based on the county in which household resides as well.

The data retrieved from the King county assessment department (King County Assessment Department 2009) is at the very detailed parcel level and it should be aggregated into the census tract level to be coordinated with the PSTP data. Parcel level addresses are mapped to the census street 2000 file and then aggregated up to the tract level as a GIS application. Finally, the aggregated land values and housing prices are merged to the PSTP data.

King county built-environment characteristics are borrowed from an adjunct survey of the PSTP in which different job category counts, intersection density, transit availability and many other land-use related variables in a grid of 150 meters by 150 meters are presented.

Finally, historical macroeconomic data are also merged to the abovementioned data sets. Variables like interest rate, inflation rate, gas price and unemployment rate are all tested in the models and their impact on the household decision on residential location attributes are examined.

## 4. Model Formulation and Methodology

Housing search process has been the topic of many researches and has been studied from different angles. Among copious variables which may represent and/or reflect the household residential location choice behavior, this study focuses on property value and work distance. As noted earlier, location selection process can be broken into two consequent and correlated sub-processes; initially household members form their choice sets by screening available alternatives and filtering them based on their priorities, and preferences. Following this step, they pick the most desirable alternative among the filtered alternatives of the choice set. When modeling this process, it is recommended (Leman 1984) that the probability of being in a choice set for each alternative also be considered in the alternative selection process. Accordingly, for each household, all alternatives with higher probability of not being excluded will be selected to the choice set.

An extensive curve fitting exercise was undertaken. It was found that the aggregate property values in census tract level follow a log-logistic distribution while work distance follows a Weibull distribution. Table 1 shows the results of the distribution test on two continuous variables of work distance and property value based on Kolmogorov-Smirnov statistics (Chakravarti *et al.* 1967 and Eadie *et al.* 1971).

Table 1 Best Fitted Distribution to the Dependent Variables

| | Average Work Distance | | Average Property Value | |
|---|---|---|---|---|
| Distribution | Kolmogorov-Smirnov Statistic | Rank | Kolmogorov-Smirnov Statistic | Rank |
| Beta | 0.04001 | 2 | 0.13549 | 7 |
| Chi-Squared | 0.13611 | 15 | 0.62077 | 16 |
| Exponential | 0.15026 | 16 | 0.2451 | 15 |
| Gamma | 0.04053 | 3 | 0.09653 | 4 |
| Gen. Extreme Value | 0.04818 | 5 | 0.08302 | 2 |
| Laplace | 0.15608 | 17 | 0.14729 | 11 |
| Log-Logistic | 0.09731 | 11 | **0.0655*** | **1** |
| Logistic | 0.11387 | 13 | 0.1487 | 12 |
| Lognormal | 0.08385 | 10 | 0.1398 | 9 |
| Normal | 0.10011 | 12 | 0.15968 | 13 |
| Weibull | **0.03533*** | **1** | 0.13572 | 8 |

** The smaller the KS statistic is for a distribution, the closer that dstribution is to the data*

It is assumed that depending on household's attributes and its cash flow and available budget, the household will have a price threshold beyond which housing alternatives are not affordable to the household. Alternatives with the price range higher than the threshold will not be accepted by the decision maker. The rate of accepting a price conditional on not accepting it before reaching to that threshold (price survival) is considered in this study for explaining the price acceptance behavior. Similarly, decision makers have some value in mind for the maximum commute distance beyond which housing alternatives will not be attractive to the household. In such cases, increasing work distance does not survive and the household will reject any alternative with the distance that surpasses the threshold defined for the household. This interpretation of the two continuous dependent variables can suggest using a hazard-based formulation framework. In a mathematical language, this can be formulated as:

$$\lambda(t)dt = \Pr(t + \Delta t \geq T \geq t \mid T \geq t) = \frac{f(t)dt}{S(t)} = \frac{S'(t)dt}{S(t)} \qquad [1]$$

where $\lambda(t)$ is the probability of failure for individual *i* given that it has survived until time *T*, $f(t)$ is failure probability density function and $S(t)$ is the survival function.

The survival function can be calculated using Equation [1] as:

$$S\ (t) = \exp\left[ -\int_0^t \lambda(u)du \right] \quad [2]$$

In addition to the baseline hazard function, other covariates like socio-demographic attributes, built-environment variables and macroeconomic factors can also be incorporated in the hazard function using a proportional hazard formulation which was initially introduced by Cox (1959). The proportional hazard formulation for average property value and average work distance with log-logistic and Weibull distributions are as follows:

Weibull baseline hazard (for work distance): $\lambda_i(wd) = \gamma\ wd^{\ \gamma-1} \exp(-\theta_x X_i)$ [3]

Log-logistic baseline hazard (for price): $\lambda_i(p) = \dfrac{\frac{\beta}{\alpha}(\frac{p}{\alpha})^{\beta-1}}{1+(\frac{p}{\alpha})^{\beta}} \exp(-\theta_x X_i)$ [4]

where $\gamma$ is the shape parameter of the Weibull distribution, $\alpha$ and $\beta$ are scale and shape parameters of log-logistic distribution, X denotes explanatory variables, $\theta_x$ is the vector of parameters, *p* stands for the average property value, and *wd* stands for the average work distance.

Using the same definitions, the survival function with Weibull or log-logistic assumption for the baseline hazard can be shown as:

Weibull baseline hazard (for work distance): $S_i(wd) = e^{-wd^{\gamma}\exp(-\theta_x X_i)}$ [5]

Log-logistic baseline hazard (for price): $S_i(p) = (1+(\frac{p}{\alpha})^{\beta})^{-\exp(-\theta_x X_i)}$ [6]

In a mathematical language, the likelihood of failure in accepting a price or a distance while examining different alternatives is equal to the hazard of failure to accept the alternative times the probability of surviving without accepting it. The joint likelihood function that is formulated for the average work distance and property value based on their hazard and survival functions across all alternatives, prices, and distances can be written as:

$$L = \prod_{i=1}^{N} \prod_{j=wd,p} \lambda_{ij}(t) \times S_{ij}(t) \quad [7]$$

where $N$ is the number of observations. This function can be maximized to estimate its parameters. The probability density functions estimated by using the results of parameter estimation of Equation [7] are then utilized to generate individual choices.

## 5. Explanatory Variables

The PSTP data set provides a long list of household socio-demographic attributes including household size, income, auto ownership, number of adults, number of workers, number of youths and children, among others. Several other dummy variables were generated that represent changes in household status such as lifestyle transitions but were not found to be statistically significant in the model.

Furthermore, two built environment variables and land-use characteristics of the area in which household resides, are included in the models. Frequency of the transit service during the day, especially mid-day, was found significant in both work distance and property values models. Number of certain employment opportunities in each gridcell surrounding the household residence were also examined for their affects on the household decision.

In addition, macroeconomic related factors like inflation rate, mortgage rate and unemployment rate were included in the explanatory variable pool. In order to have all prices and income values to be comparable, the first used wave of the PSTP was assumed to be the base year and the property prices and incomes referring to years after the base year were deflated to the base year using the historical inflation rates. Since the actual (not nominal) property values are available in each year, for each household, the equivalent base-year property value can be estimated. The historical nominal 30-year mortgage rates (HSH Associates Financial Publishers 2009) were also converted to the real mortgage rates by subtracting the nominal values from the corresponding year inflation rates. Macroeconomic effects on the household work distance are captured through the unemployment rates obtained from the US 125 years Bureau of Labor Statistics (U.S. Bureau of Labor Statistics 2009).

The average value and standard deviation of the explanatory variables that were found statistically significant in the models are presented in Table 2.

Table 2 Explanatory Variable Used in the Models

| Explanatory Varibels | Average | St. Dev. |
|---|---|---|
| Income | 51537.12 | 26985.79 |
| Number of employed | 1.20 | 0.85 |
| Number of Vehicles | 1.76 | 0.83 |
| Number of adults | 1.82 | 0.64 |
| Number children 6-17 | 0.26 | 0.64 |
| Number of Children 1-5 | 0.20 | 0.54 |
| Job count in a small gridcell* | 0.76 | 1.54 |
| Mid-day transit availability in large gridcell ** | 5.09 | 9.88 |
| Unemployment rate | 5.82 | 1.09 |
| 30 year Mortgage Rates (lag Var.) | 8.05 | 0.76 |
| 30 year Mortgage Rates (Forward Var.) | 7.67 | 0.66 |

** 150 meters by 150 meters gridcells*

*** 750 meters by 750 meters gridcells*

## 6. Modeling Results and Analysis

The results of parameter estimation of joint choice set formation model in which average work distance and property value are determining are presented in Table 3. Model parameters are estimated by maximizing the likelihood function presented in Equation [7] using the *nlp* procedure provided by SAS 9.1.3 package. Before evaluating the quality of the estimated parameters, it should be noted that the effect of covariates in a hazard model is facilitated by incorporating negative sign for parameters in formulation. In other words, if a covariate gets a negative sign, the chance of failure or the probability of accepting a price/work distance is increased. Alternatively, having a positive sign means that any increase in the covariate decreases the chance of failure for the household which implies that the household tends to increase the price/work distance.

Table 3 Results of Joint Model of Household Average Work Distance and Property Value

Household Average Property Value

| Parameter | Estimate | t Value | Pr > \|t\| |
|---|---|---|---|
| Alpha | 2.006 | 4.341 | 0.000 |
| Beta | 3.130 | 6.234 | 0.000 |
| Constant | 6.764 | 2.884 | 0.004 |
| Previous Prop. Value (X1,000) | 0.008 | 4.180 | 0.000 |
| Income (X100,000) | 0.451 | 1.298 | 0.196 |
| Number of Children Age 1-5 | -0.300 | -1.802 | 0.074 |
| Change in Number of Age 6-17 | -0.688 | -3.146 | 0.002 |
| Change in Number of Employeds | 0.206 | 1.538 | 0.126 |
| Mid-Day Transit Availability | -0.015 | -2.365 | 0.019 |
| Educational Service Job in the Gridcell | -0.086 | -1.438 | 0.153 |
| Lag_ 30 Year Mortgage Rate | -1.249 | -4.584 | 0.000 |
| Forward_ 30 Year Mortgage Rate | -0.501 | -2.984 | 0.003 |

Household Average Work Distance

| Parameter | Estimate | t Value | Pr > \|t\| |
|---|---|---|---|
| Sigma | 1.828 | 12.505 | 0.000 |
| Constant | 2.847 | 6.680 | 0.000 |
| Previous Work Ditsance | 0.074 | 4.274 | 0.000 |
| Change in Income (X100,000) | 0.683 | 1.165 | 0.246 |
| Number of Vehicles | 0.281 | 2.195 | 0.030 |
| Number of Employeds | 0.193 | 1.834 | 0.069 |
| Change in Number of Adults | -4.630 | -1.740 | 0.084 |
| Mid-Day Transit Availability | -2.398 | -3.576 | 0.000 |
| Unemployment Rate Change | -0.238 | -1.486 | 0.139 |

| | |
|---|---|
| *Likelihood value with only constant* | *-592.96* |
| *Likelihood value at convergence* | *-531.73* |
| $-2[L(C)-L(\beta)]$ | *122.47* |

The beta parameter of the log-logistic distribution of the property value model is greater than one which presents a non-monotonic function. Furthermore, the Weibull distribution of the work distance models has a monotonically increasing shape.

The property value of the previous residence can be considered as one of the factors that can approximate the price range of the next residence to which the household will be moving. It is expected that wealthier households and household that have received income raise purchase more expensive houses. Interestingly, the negative sign of the number of children (1-5 and 6-17 years variables) in the model

suggests that if the total number of children and/or total number of youths increases, the residential relocation choice can shift towards the less expensive areas. This can be justified by the fact that larger households with more children in the family tend to need more space, therefore, all else being equal, they tend to obtain more housing for their limited budget resulting in higher probability of selecting less expensive properties. On the other hand, as the total number of workers in the household increases, the household tends to move to more expensive neighborhoods. This is in line with earlier findings on the income variable since more workers in a household should result in higher earned income.

Two built environment related variables are included in the household average property value model and both are shown with negative sign in the model. Household living in transit accessible areas tend to consider less expensive neighborhoods. Households living in a school zone or area with more educational jobs, (e.g., university students living around campus), are also more likely to consider lower cost alternatives. Finally, the effect of mortgage rate which is an important external variable from the supply side of the market on the housing decision is examined. Both lag (previous year) and forward (next year) mortgage rates are considered to account for the effects of market experience and future market expectations. Intuitively, as the mortgage rate increases, households with limited budgets tend to select less pricey areas to meet their budget constraints.

Similar to the case of average property value, household's current average work distance is considerably affecting the household decision about its new residence. Annual income which is positively correlated with the number of vehicles is also important on the household decision about the average work distance. The higher a household income is, the farther they can select their work location from the residence. Wealthier households are also more likely to live in suburban areas and commute farther distances. Similarly, total number of vehicles in the household is positively correlated with the work distance. Households with more workers can commute to farther work destinations whereas households with more changes in the number of adults are likely to work closer to their home. Households living in areas with more available mid-day transit are also more likely to reduce their work distance. Finally, unemployment rate as a representative of the supply side of the market, found to be significant in the household average work distance. Results shown in Table 3 imply that any increase in the unemployment rate shifts the households' tendency to reduce their average work distance.

The likelihood function value at convergence is -531.6. Therefore, the statistic $-2[L(C)-L(\beta)]$ would be $-2[592.96-531.6] = 122.47$. It is noteworthy that this statistics is asymptotically Chi-square distributed with degrees of freedom of 21 which is highly significant.

## 7. Simulation and Sensitivity Analysis

The housing search choice set formation model that is developed in this study is validated both internally and externally. The internal validation involved evaluating the overall goodness of fit of the model, likelihood value at convergence, and estimated parameters that were all statistically significant. The explanatory variables were also selected such that choice specific, household taste variation and market characteristics are included in model.

The external validation was performed through a simulation practice. The parameter estimates of the model that are presented in Table 3 were used to estimate the probability of accepting a price/work distance for each household. As noted earlier, the probability density function for accepting a price/work distance can be obtained by estimating the product of hazard and survival functions. The probability density function can be easily written using Equations [3] to [6] as:

The probability density function for work distance is:

$$f_i(wd) = \left[\gamma\ wd^{\ \gamma-1} \exp(-\theta_x X_i)\right] \times \left[e^{-wd^{\gamma} \exp(-\theta_x X_i)}\right] \quad [8]$$

and the probability density function for price can be shown as:

$$f_i(p) = \left[\frac{\frac{\beta}{\alpha}\left(\frac{p}{\alpha}\right)^{\beta-1}}{1+\left(\frac{p}{\alpha}\right)^{\beta}} \exp(-\theta_x X_i)\right] \times \left[(1+\left(\frac{p}{\alpha}\right)^{\beta})^{-\exp(-\theta_x X_i)}\right] \quad [9]$$

As shown in equations 8 and 9 above, the probability density functions of price and work distance are a function of household characteristics. The probability of accepting a price and a work distance were estimated for each household using Equations [8] and [9]. These equations can generate a probability density function for each household similar to the one shown in Figure 1 for average household work distance.

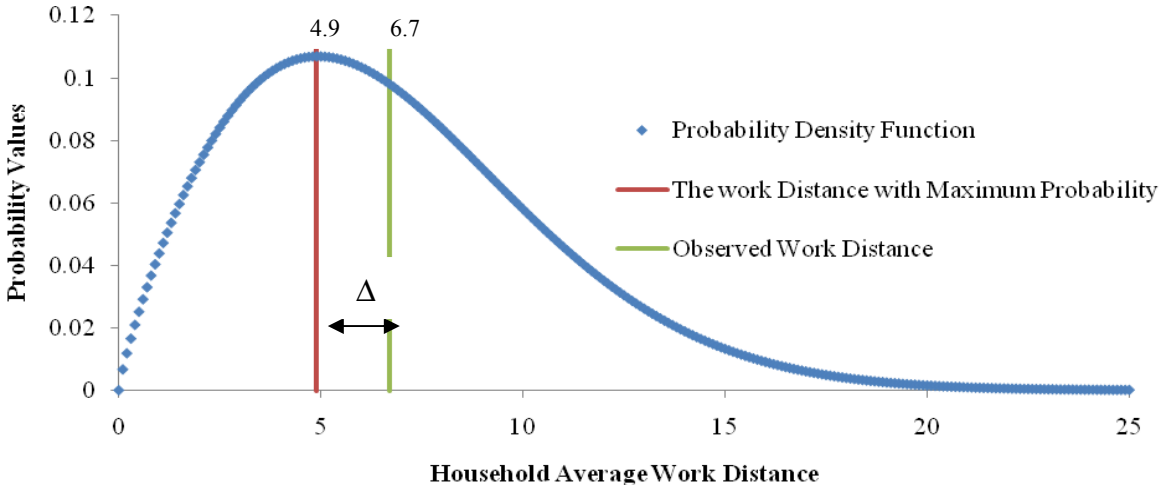



Figure 1 An example of estimating the percentage distance between the maximum probability density and the observed value of household work distance ($\lambda = (6.7 - 4.9)/4.9 = 36.73\%$)

Given the observed household work distance, one can compare and estimate the difference ($\Delta$) between the observed and estimated maximum probability values for each household. Then the percentage difference ($\lambda$) between the observed and simulated prices and work distance values can be calculated for each observation.

Various distributional forms are fitted to these $\lambda$ values and the best fitted density functions are selected based on the Kolmogorov-Smirnov test (Chakravarti *et al.* 1967 and Eadie *et al.* 1971) that was shown to be statistically highly significant in both cases. The cumulative density functions of the best fitted distributions for property value and work distance variables are presented in Figures 2 and 3.

Figure 2 presents the cumulative density functions for the estimated percentage differences between the observed and simulated property values along with the best fitted density function parameter values. Similarly, Figure 3 presents the difference between observed and simulated average work distance values.

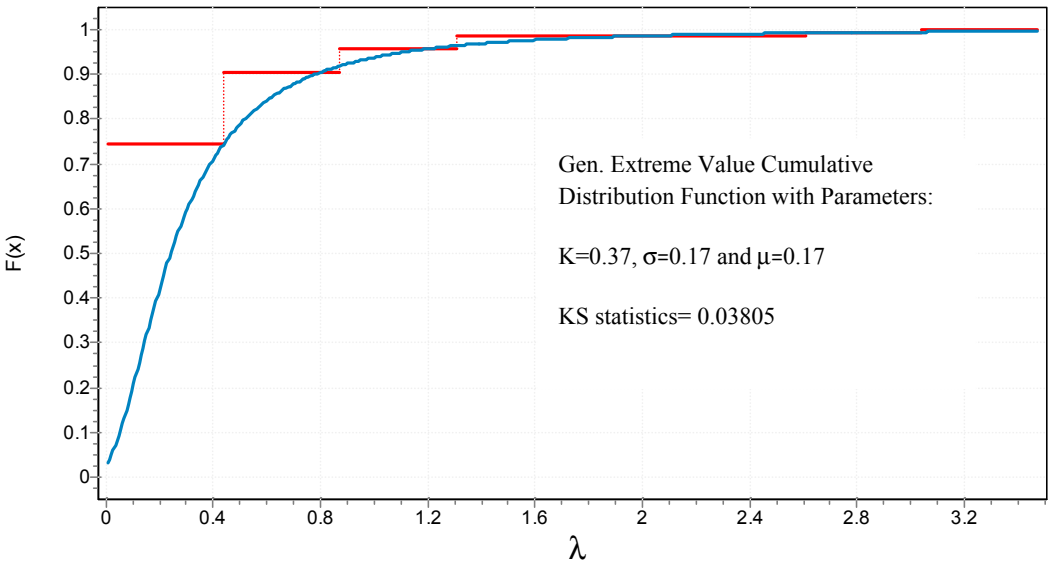



Figure 2 Sample frequencies and cumulative density function for percentage differences between simulated and observed property values.

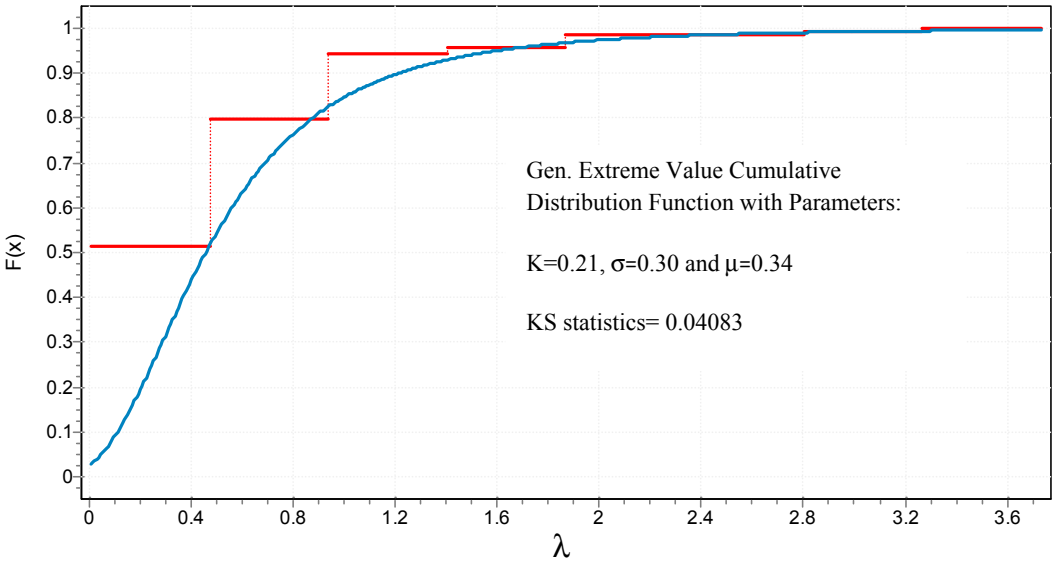



Figure 3 Sample frequencies and cumulative density function for percentage differences between simulated and observed average work distance.

As shown in Figure 2, over 90 percent of the predictions are at most 80 percent different from the observed property values. In other words, if the housing search choice set formation method that is presented in this study is utilized to simulate property price threshold values (y) and used while considering 0.8y variation to filter potential alternatives, the resulting choice set will be 90% accurate.

Similarly, in the case of average work distance simulation shown in Figure 3, it can be observed that if the choice set includes areas with work distance standard deviation of 0.8y, then the resulting choice set will be 80% accurate.

Other than model validation purpose, Figures 2 and 3 can indicate that the model developed in this study are capable of being used in a simulation process in which household housing search choice sets can be accurately generated. Therefore, there is no need to consider a *universal choice set* for all the households which will cause computational difficulties and potential risk of erroneous estimates. Rather, the choice set can be cut down into a manageable but highly accurate set of alternatives that can result in more efficient and unbiased model estimation.

## 8. Conclusions and Future Directions

This study presents a behavioral model of alternative set formation for residential location choice problem. A two-step approach is considered in which alternatives are evaluated and screened based on household priorities, lifestyle, and preferences and for each alternative, the probability of being selected in the choice set is estimated. Following that, the alternative with highest utility can be selected using traditional choice models. This study focused on the first step of the household housing search process but the choice set formation methodology that is introduced in this paper can be then coupled with a discrete choice model to present a comprehensive location choice model..

Two continuous variables of average property value and average work distance were jointly modeled using a hazard-based formulation. It should be noted that the application of property value and travel distance variables in this paper does not limit the use of other variables in the proposed modeling framework. The modeling framework can account for the joint effects of multiple factors, therefore, other potential variables such as school quality, safety and crime rate, property tax, etc can be also considered. The likelihood function was formulated based on the assumption that Weibull distribution is the preferred distribution for the average work distance while log-logistic distribution provides better fit to the average property value. Incorporating the failure and survival concept in defining the choice acceptance level threshold is one of the contributions of this study.

The Puget Sound Transportation Panel of Seattle Metropolitan Area was used in this study for the modeling practice along with other sources of data such as built environment, land-use, and economic factors. Many household socio-demographic attributes and several land use indicators were tested in the modeling process among which total number of children, youths, adults, employed members, vehicles and income were found to be statistically significant in the model. Among built environment variables the mid-day transit availability and educational job counts in

the region were significant in the final model. From the supply side of the market, inflation, mortgage rate and unemployment rate were proved to influence the household residential location decision. Household's perception about last year and next year mortgage rates were also found to be critical factors that can influence the taste of the household and its housing price threshold.

The models developed in this study were validated internally and externally and overall, it was shown that they are capable of generating highly accurate choice sets that can result in more efficient and unbiased housing search models.

Further improvements to the model include: incorporating heterogeneity in the choice set formation, investigating the importance of other variables on housing search choice set formation besides property value and work distance, and integrating the choice set probabilities with a discrete choice model. These improvements remain as future research tasks. It should be also noted that the application of the proposed modeling framework is not limited to the housing search problem. Such a framework can be used in other contexts where large number of alternatives should be evaluated. For instance, in the case of activity location choice (e.g., shopping) a similar approach can be used, however, instead of price and distance, other appropriate factors such as size (e.g., number of stores, or retail jobs) can be used along with distance.

## 9. References

Arnold, S. J., Oum T. H. and Tigert D. J., (1983), Determinant attributes in retail patronage: seasonal, temporal, regional, and international comparisons, *Journal of Marketing Research*, 20, pp. 149-57

Ben-Akiva M. and Lerman S.R., 1985, Discrete choice analysis. Theory and application to travel demand, Cambridge: MIT Press

Bhat C.R.,(2005), A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions, *Transportation Research Part B*, 39 (8), pp. 679–707

Burnett K.P. and Hanson S., (1979), Rationale for an alternative mathematical approach to movement as complex human behavior, *Transportation Research Record,* pp.723, 11-24

Burnett K.P. and Hanson S., (1982), The analysis of travel as an example of complex human behavior in spatially-constrained situations: definitions and measurement issues, *Transportation Research A,* 16, pp. 87-102

Chakravarti I.M., R.G. Laha, J. Roy, 1967, Handbook of Methods of Applied Statistics, vol. I, John Wiley and Sons

Clark W. A. V., Huang Y. and Withers S. D., (2003), Does commuting distance matter? Commuting tolerance and residential change, *Regional Science and Urban Economics* 33, pp. 199-221

Cox, D. R., (1959), The analysis of exponentially distributed life-time with two types of failures, *Journal of Royal Statistical Society,* Vol. 21B, pp. 411-421

Domencich T. A. and McFadden D., (1975), *Urban travel demand: a behavioral analysis*, Amsterdam: North-Holland

Eadie, W.T., Drijard D., James F.E., Roos M. and Sadoulet B., 1971, Statistical Methods in Experimental Physics. Amsterdam: North-Holland, 269-271.

Fotheringham A. S., (1983), A new set of spatial interaction models: the theory of competing destination, *Environment and Planning A*, 15, pp. 15-36

Howard J.A., (1963), Marketing Management, Homewood, IL, Richard Irwin

HSH Associates Financial Publishers, (2009), Historical Mortgage Rates, http://www.hsh.com/mtghst.html last accessed on July 2009

Kim S., (1992), Search, Hedonic Prices and Housing Demand, *The Review of Economics and Statistics*, 74, pp. 503-508

King County Assessment Department, (2009), Parcel Level Property Values, King County, Washington, http://info.kingcounty.gov/assessor/DataDownload/default.aspx last accessed on July 2009

Lerman S. R., (1984), Recent advances in disaggregate demand modeling, In M. Florian (Ed.), *Transportation Planning Models*, (Amsterdam: North-Holland)

Lerman S.R., (1985), Random utility models of spatial choice. In Hutchinson, B.G., Nijkamp, P., Batty, M., editors, *Optimization and discrete choice in urban systems*, Berlin: Springer-Verlag, pp. 200-217

Manski, C.F. 1977: The Structure of Random Utility Models, *Theory and Decision*, 8, pp. 229-254

McFadden D., (1974), Conditional Logit Analysis on the Temporal Stability of Disaggregate Travel Demand Models, Transportation Research Part B, 16, pp. 263-278

McFadden D., (1978), Modeling the choice of residential location. In Karlqvist, A., undqvist, L., Snickars, F. and Weibull, J.W., editors, *Spatial interaction theory and planning models*, Amsterdam: North-Holland, pp. 75-96

Murakami E. and Watterson W. T., (1992), The Puget Sound transportation panel after two waves, *Transportation*, Vol. 19, No. 2, pp. 141-158

Richards M. G. and Ben-Akiva M. E., 1975), *A disaggregate travel demand model,* Lexington: Lexington Books

Salomon I. and Ben-Akiva M., (1983), The use of the life-cycle concept in travel demand models, *Environment and Planning A*, 15, 623-38

Shocker A. D., Ben-Akiva M. E., Boccara B. and Nedugadi P., (1991),Consideration set influences on consumer decision-making and choice: issues, models and suggestions, *Marketing Letters*, 2, pp. 181-197

Srinivasan T., (1987), An integrative approach to consumer choice, *Advances in consumer research*, 14, pp. 96-101

Thill J. C. and Horowitz J. L., (1991), Estimating a destination-choice model from a choice-based sample with limited information, *Geographical Analysis*, 23, pp. 298-315

Thill J. C., (1992), Choice set formation for destination choice modeling, *Progress in Human Geography,* **16(3)**, pp. 361-382

Timmermans, H.J.P. and Golledge, R.G., 1990, Applications of behavioral research on spatial problems II: preference and choice, Progress in Human, Geography 14, pp. 311-54

U.S. Bureau of Labor Statistics (2009), Local Area Unemployment Statistics, http://www.bls.gov/lau last accessed on July 2009

Waddell P., (1996), *Accessibility and Residential Location: The Interaction of Workplace, Residential Mobility, Tenure, and Location Choice*, Presented at the Lincoln Land Institute TRED Conference

Willumsen L. G. and Ortuzar J. de D., (2001), Modelling Transport, John Wiley & Sons, New York

Wright P. and Barbour F., (1977), Phased decision strategies. In: M. Starr and M. Zeleny (eds.) management Science, Amsterdam, North Holland, pp. 91-109

# An Analysis of Person-to-Person Car Sharing

*R.C. Hampshire[1] and C. Gaites[1]*

**Abstract:** Person-to-Person (P2P) car sharing allows car owners to convert their personal vehicles into share cars which can be rented to other drivers on a short-term basis. Using data from the Longitudinal Employment Dynamics (LED) program and the US Census, we develop a methodology to assess the market feasibility of, and economic incentives, P2P car sharing. Further, we apply the methodology to develop a case study of P2P car sharing in Pittsburgh, PA.

## 1. Introduction

As of 2007 there were over 237 million private vehicles owned and operated in the United States (USBTS, 2007). Estimates of the average cost to own, maintain, insure, and park a private vehicle ranges from $460 to $913 per month (AAA, 2007). All of these private vehicles also represent an enormous environmental burden. Car sharing is a demand side approach to reducing the environmental impact of the transportation sector. It also simultaneously reduces private transportation costs for some drivers with only intermittent need for vehicle transportation. Because it changes the economics of driving by converting vehicle transportation from a fixed cost into a variable cost, car sharing has been shown to reduce mode adjusted vehicle miles traveled (MVMT) among members by 67% (Cervero, 2007). Previous studies have demonstrated that because members of car sharing services are much less likely to purchase their own cars, and may even sell a car after joining a car sharing service, each new share car added to existing car sharing fleets removes 4.6 to 20 private vehicles from the road (Shaheen S. C., 2005). The economic benefits to members are significant as well. The average member of the City CarShare car sharing service in San Francisco spends only about $540 per year on automotive transportation (Sullivan & Magid, 2007). This represents a tenfold cost savings when compared to owning a small sedan (AAA, 2007).

A range of market demand studies conducted in the US and Europe have estimated cost savings alone would drive between 3% and 25% of the driving population to forego car ownership, or to replace their privately owned cars, and instead take up membership in a car sharing service (Millard-Ball, 2005). Other research estimated that if a sufficient number of conveniently located vehicles were available, then 10% of the individuals over the age of 21 in metropolitan areas of North America would adopt car sharing (Shaheen S. C., 2005). Today, despite the enormous potential environmental benefits, and despite considerable consumer demand, adoption rates for car sharing are currently 12 to 30 times lower than projected by market research (Millard-Ball, 2005).

[1] H. John Heinz III College, School of Public Policy and Management, Carnegie Mellon University, 2102B Hamburg Hall, 4800 Forbes Ave, Pittsburgh, PA, 15217; hamp@cmu.edu; cgaites@andrew.cmu.edu

Despite significant expected consumer demand, the traditional car business model is difficult to scale geographically. In the traditional car sharing model the business that manages the share car fleet must bear the high upfront, fixed cost of leasing or purchasing all of the vehicles in the fleet. Each new share car must be located in a high density area with many potential members in order to drive utilization of that vehicle and offset these costs. In order for the fleet operator to realize a profit on its initial investment, about 25 active members must live within a ¼ mile of each vehicle to ensure sufficient utilization of the share car (Sullivan & Magid, 2007). The five largest US car sharing services maintained member to vehicle ratios of 66:1 in 2005 (Shaheen S. C., 2005). Even when acceptable locations are found, and vehicles placed into service, there is still substantial risk that membership will never reach sufficient levels to drive enough utilization to offset the capital costs associated with those vehicles (Zipcar, 2010).

This paper explores the feasibility of a more scalable form of car sharing, called Person-to-Person (P2P) car sharing. P2P car sharing allows car owners to convert their personal vehicles into share cars which can be rented to other drivers on a short-term basis. With a P2P car sharing service upfront fixed costs for the business that manages the fleet, and therefore the utilization requirements, are greatly reduced. There are currently at least 5 operating P2P car sharing services: Getaround, Go-Op, RelayRides, SprideShare and WhipCar. Using data from the Longitudinal Employment Dynamics (LED) program and the US Census, we develop a methodology to assess the market feasibility of, and economic incentives and externalities associated with, P2P car sharing. Further, we apply the methodology to develop a case study of P2P car sharing feasibility in Pittsburgh, PA.

In Section 2, we review the literature on market segmentation and demand estimation for traditional car sharing and summarize the key success factors. In Section 3, we present a methodology to assess the feasibility of P2P car sharing from both the demand and supply side. A queueing theory analysis is used to match supply and demand to determine the viability of P2P car sharing in a given area. Section 4, presents the economic case for car owners and renters to participate in P2P car sharing. Finally, Section 5 applies the P2P feasibility methodology to Pittsburgh, PA.

## 2. Success Factors for Car Sharing

For the sake of clarity, in this paper a traditional car sharing service will be defined as: "A membership program intended to offer an alternative to car ownership under which persons or entities that become members are permitted to use vehicles from a fleet on an hourly basis (Celsor & Millard-Ball, 2007)." As P2P car sharing is largely indistinguishable from traditional car sharing when viewed from the vantage point of a prospective renter member, the methodology employed in this paper to assess the potential renter demand for P2P car sharing is based on the current state of practice for assessing demand for traditional car sharing services.

The commercial success of any car sharing operation, be it a community based service or a for-profit service, is directly tied to the level of vehicle utilization achieved. Utilization rate is the primary indicator of a car sharing organizations financial health. Point of Departure (POD) location selection is the primary driver of

achieving a high level of vehicle utilization (Sullivan & Magid, 2007). Utilization rate is generally measured in terms of revenue-hours-per-vehicle day. Successful car sharing services have achieved utilization rate as high as 40%, meaning that the average vehicle in the operator's fleet is being driven by a paying customer for more than 9 hours of every 24 hour day (Sullivan & Magid, 2007).

In order to select appropriate locations for PODs operators of car sharing services must perform careful market segmentation to identify locations where high utilization can be achieved,. Market segments for car-sharing can be usefully divided into two categories: 1) *Geographic markets* – the neighborhoods where car-sharing vehicles can be placed to best effect; and 2) *Demographic markets* – the demographic groups that are most likely to join a car-sharing program (Celsor & Millard-Ball, 2007). The phrase geographic market refers to the macro-scale characteristics of the neighborhood in which a shared car is located. The phrase demographic market on the other hand refers to the individual characteristics of car sharing users themselves.

### 2.1 Geographic Markets

Previous literature describes the geographic markets where car sharing has taken root and several common market characteristics have been identified. Residential density is generally cited as an important success factor for car sharing. High residential density is often closely connected with quality and availability of transit, and also neighborhood walk-ability (Sullivan & Magid, 2007). If the geographic reach of existing car sharing services is any indication, then it is clear that density is a strong predictor of car sharing success. In 2004, 94% of car sharing membership was concentrated in eight densely populated metropolitan regions: San Francisco, Los Angeles, San Diego, Portland, Seattle, Boston, New York, and Washington, DC (Shaheen, Schwartz, & Wipyewski, 2004).

Density is also important because it is an indication of the size of the potential customer base of a POD. About 80% of all members of City CarShare's car sharing service live within ½ mile (or 10 minute walk @ 3mph) of the nearest POD (Cervero, 2007). More than half live within ¼ mile of the nearest POD (Cervero, 2007). Doubling density around a POD doubles the number of potential users for the vehicles located at that POD (Sullivan & Magid, 2007). In the City CarShare publication *Bringing Car Sharing to Your Community*, Sullivan & Magid indicate that 50-100 members living within ½ mile radius is threshold for a successful 2-3 car POD (25-33 members/car). They also provided a table indicating the levels of market penetration required to achieve 25 members per vehicle living within a five minute walk (about ¼ mile) of a POD at various population densities as a guide for assessing location feasibility.

In addition to density, Celsor and Millard-Ball found that vehicle ownership had the strongest, most consistent correlation to the amount of car sharing service in a neighborhood (Celsor & Millard-Ball, 2007). Car ownership rates of less than 1 car per household were strongly predictive of the presence of car sharing in a neighborhood. Several other factors also predict the success of car sharing. 1) *Parking Difficulty/Cost*; 2) *Mix of Land Uses* ; 3) *Transit Access and Commute Mode Split*; and 4) *Household/Family Composition.*

### 2.2 Demographic Markets

Demographic markets are closely interrelated with geographic markets as demographic characteristics in a neighborhood are generally influenced by the geographic characteristics of the neighborhood and vice-versa. Regardless of the direction of causality, several demographic characteristics correlate with participation in car sharing. 1) *Age* – A survey of North American car sharing participants found 39% were in the 25 to 34 year old age group; 27.4% were in the 35 to 44 year old age group (Millard-Ball, 2005). 2) *Education* - Residents living in POD neighborhoods are also far more likely than the average person to hold a Bachelor's degree or higher (Celsor & Millard-Ball, 2007).

## 3. Feasibility

In this section, we present a methodology to assess the feasibility of P2P car sharing in a given geography. The analysis first considers the total supply of parked cars. Next, the total demand is estimated. Finally, the minimum total number of P2P cars needed to support the demand while maintaining a quality renter experience is calculated. This minimum required number of P2P cars is compared to the estimated number of parked cars to determine the feasibility of P2P car sharing in a given study area.

### 3.1 Supply

The fundamental difference between determining the feasibility of P2P car sharing and traditional car sharing is the need to estimate the number of parked cars. In this section, we describe a methodology to estimate the total supply, or number of parked cars per hour, in car hours for an area. We consider both the cars parked due to work trips and cars parked at residences. The number of cars parked due to employment is estimated using data from the 2008 Longitudinal Employment Dynamics (LED) program of the US Census. The number of cars parked at residences is estimated using data from the 2000 US Census. The LED combines data from federal and state sources on employers and employees to provide job counts on the census block level. The total number of jobs on a block multiplied by the mode share of driving to work, (see the Census Transportation Planning Package CTTP), is a proxy for the total number of cars parked due to employment. To estimate the temporal arrival and departure of parked cars, we use "time of departure to work" data on the county level from the CTTP. Here, we will assume that the arrival to work temporal behavior on the individual block level is consistent with the arrival to work behavior at the county level.

The US census provides data on the total number of vehicles available to households at the block level. Also, "means to work" columns in the census give the total number of vehicles driven to work per block. To estimate the temporal arrival and departure of parked cars at residences, we again use the CTTP data on the "time of departure to work." For simplicity, we assume all cars that are not driven to work are assumed to be parked.

### 3.2 Demand

In order to determine if a sufficient level of demand exists within a given area to support a P2P car sharing service, our analysis first projected the *number of potential car sharing service members* in that area and then compared that number to the area's total residential population. The ratio of potential renter members to the total residential population is considered the *maximum achievable market penetration* within that area. Population density in each area, measured in persons per square acre, was then calculated by dividing the total residential population by the geographic area in square acres. Next, the maximum achievable market penetration is compared to the *required level of market penetration* necessary to support a single car POD at the calculated population density. If the maximum achievable market penetration for an area was found to exceed the required level of market penetration for achieving 10 members within ¼ mile of a single car POD, then the area was determined to have sufficient renter demand for P2P car sharing.

First, we determine the *maximum achievable market penetration* in each area using US Census data. For this study the following data were obtained: 1) The percentage of the population within each area aged 25-40; 2) The percentage of households with two or fewer people; and 3) The percentage of workers over 16 that do not drive a vehicle to work. The total population of each area is multiplied by the percentage in an urban area, the percentage aged 25-40, the percentage in households with 2 or fewer people, and the percentage of workers that do not drive to work. The resulting population was classified as the number of potential car sharing service members in the area. In order to calculate the *maximum achievable market penetration*, the calculated number of potential car sharing service members in each area is divided by the total population in that area.

Next, we compute the *required levels of market penetration* to which the calculated maximum achievable market penetration should be compared. The table of Sullivan & Magid's showing required penetration rates for achieving 25 members within ¼ mile of a single car POD is expanded to include required penetration levels for achieving 20, 15, 10, & 5 members in ¼ mile radius of a single car POD. This is done by scaling Sullivan & Magid's required penetration rates linearly with the required number of members. The expanded table is presented below (TABLE 1).

**TABLE 1: Penetration rates required to achieve 25, 20, 25, 10, or 5 members per vehicle.**

| Density (Persons/ Acre) | Penetration Rate Needed (25 members) | Penetration Rate Needed (20 members) | Penetration Rate Needed (15 members) | Penetration Rate Needed (10 members) | Penetration Rate Needed (5 members) |
|---|---|---|---|---|---|
| | 25 | 20 | 15 | 10 | 5 |
| 10 | 1.99% | 1.59% | 1.19% | 0.80% | 0.40% |
| 15 | 1.33% | 1.06% | 0.80% | 0.53% | 0.27% |
| 20 | 0.99% | 0.79% | 0.59% | 0.40% | 0.20% |
| 25 | 0.80% | 0.64% | 0.48% | 0.32% | 0.16% |
| 30 | 0.66% | 0.53% | 0.40% | 0.26% | 0.13% |

A penetration threshold of 10 renter members living with a ¼ mile radius of each single car POD selected based on the authors projection that a P2P car sharing model would lower a car sharing providers cost basis by approximately 60%. The author's projection was informed by interviews with car sharing operators, and is supported by an examination of the financials included in Zipcar's recently filed Securities and Exchange Commission Form S-1 Registration Statement. The Statement of Operations presented in Zipcar's filing indicates that "Fleet Operations" accounted for, on average, 69% of total operating expenses between 2008 and 2010 (Zipcar, 2010). The item "Fleet Operations" consist principally of costs associated with operating vehicles, such as lease expense, depreciation, parking, fuel, insurance, gain or loss on disposal of vehicles, accidents, repairs and maintenance as well as some employee-related costs (direct labor on vehicles). The P2P car sharing model eliminates all of those expenses except insurance and fuel.

The determination of market viability (in terms of demand) for P2P car sharing is based on the possibility of achieving an effective member to vehicle to ratio of ten renter members per car at a given area's population density. Therefore, in each viable area the maximum number of cars for which there is sufficient demand is equal to the number of potential car sharing service members divided by ten.

If we assume that the maximum numbers of cars are supplied in each area (a consistent renter member to vehicle ratio of 10:1), then the expected utilization of each vehicle can be calculated. Cervero's five year study of the impacts of City CarShare's car sharing service in San Francisco found that the average renter completed 1.31 trips each month. The average duration of each trip was found to be 3.93 hours, during which time approximately 21.2 miles were covered (Cervero, 2007). Millard-Ball's survey of car sharing members in North America found that members completed an average of 3.34 trips per month, and that the median number of trips taken per month was 2 (Millard-Ball, 2005). Assuming an average trip frequency of 2.33 trips per month (the average of 1.31 and 3.34 trips per month), an average trip duration of 3.93 hours, and an average trip distance of 21.2 miles, then the average owner supplied vehicle would be rented 23 times per month (280 times per year), driven 494 miles per month (5928 miles per year), and occupied by a renter for 92 hours per month (1099 hours per year).

### 3.3 Matching Supply and Demand

In this section we compute the minimum number of cars needed in each target area to accommodate the demand generated by the target renters in that area. Given that the demand occurs randomly throughout the day, we must consider contention for booking a vehicle. Thus, we employ a simplified queueing theory analysis ***(Kleinrock, 1976)***, which allows us to estimate the number of cars needed to support the generated demand while providing an acceptable car availability level for renters.

First, we characterize the random temporal arrival of requests from target renters. The average car sharing member takes 2.33 trips per month. For simplicity of analysis, we assume that these 2.33 trips per month are requested uniformly throughout the days of the month. We further assume that the trips are taken uniformly within a given day. The hourly demand rate per target renter is then 2.33/(30*24) where each of the trips has an average length of 4hrs.

We now estimate the car utilization of traditional car sharing organization. This utilization serves a target for the P2P service. Traditional car sharing companies places at most one share car for every 25 members within ¼ mile of the car. If a POD has one car, then single server queueing analysis results in a vehicle utilization of 12.5%. In this uniform demand setting, this result is equivalent to a renter having a 12.5% chance of being blocked from reserving the car. This utilization value varies based on the operators renter to vehicle ratio.

In Section 4, we estimated that P2P car sharing removes up to 60% of the operational costs of car sharing. This reduction leads to economic viable share cars with only 10 members that live within ¼ mile of the car. We can now compute the number of cars needed in each target area, such that the rental availability is identical to that of traditional car sharing. In this way, P2P sharing provides identical quality of service as traditional car sharing. For each target area, we assume that all of the cars are indistinguishable and desirable to every target renter in that target area. In this simplified setting, the probability that a car is not available when a renter requests a reservation is given by the Erlang C formula, $\gamma(C,\rho)$, (Kleinrock, 1976)) with a traffic intensity parameter, $\rho = 4n \cdot 2.33/(30 \cdot 24)$ and $C$ is the number of P2P shared cars, $n$ is the number of target renters in a study area. The minimum number of cars needed per target area is then the smallest number of cars given a total demand rate that keeps the probability of a blocked request less than or equal to 12.5%, or $\gamma(C,\rho) \leq .125$. This type of calculation has a unique solution and is routine in the management of call centers (Noah, Koole, & Mandelbaum., 2003). The resulting minimum number of needed P2P shared cars is compared to total supply of cars available to determine the viability of the study area.

## 4. Economic Incentives

For P2P car sharing to function there must be an economic incentive for both parties to the P2P exchange to participate. If per hour and per mile pricing for a vehicle rental is established, then existence of an economic benefit for each party at that price level can be determined using the vehicle usage projections presented in Section 3.2. For the purposes of the analysis presented here rental pricing of $5.50 per hour, plus $0.35 per mile is assumed. This pricing is consistent with that offered by San Francisco's City CarShare as part of their ShareLocal & SharePlus pricing plans, and was selected because data taken from Cervero's study of that service were used to project frequency of usage. We focus on the owner economic incentives given that the renter economic incentives are well studied in the literature.

The economic benefit to the owner equals the sum of all payments made by renters less any transaction fees paid to the P2P provider who coordinates the network, the cost of the additional depreciation incurred, and the opportunity cost of time invested by the owner when managing the rental process. This calculation ignores income taxes and assumes that the costs of insuring and fueling the vehicle are borne by the P2P service provider who coordinates the network in exchange for a 30% transaction fee.

If the owner's vehicle is assumed to be a small sedan, then AAA estimates the cost of deprecation for a vehicle driven 10000 to 15000 miles per year at $0.17 to $0.22 per mile (AAA, 2007). Determining opportunity cost requires determining both the owner's personal income and the setup time invested into each rental. The economic benefit is $ 5,084.39 for an owner with a small sedan and a personal income of $20,000 per year who invests ½ of an hour of setup time into each rental assuming usage of 5928 miles and 1099 hours per year. Rental payments are based on usage charges of $5.50 per hour and $0.35 per mile.

The net economic benefit varies with the owner's income and also varies with the setup time invested. The table below (TABLE 3) shows the net economic benefit calculated for various levels of income and various per rental setup times.

**TABLE 3: Net economic benefit (loss) at various levels of income for various setup times**

| | | Owner Per Capita Annual Income | | | | |
|---|---|---|---|---|---|---|
| Setup time (Hours) | | **$ 10,000** | **$ 30,000** | **$ 50,000** | **$ 70,000** | **$ 90,000** |
| | **0.25** | $5,420 | $4,748 | $4,076 | $3,404 | $2,731 |
| | **0.5** | $5,084 | $3,740 | $2,395 | $1,051 | ($292) |
| | **0.75** | $4,748 | $2,731 | $715 | ($1,300) | ($3,317) |
| | **1** | $4,412 | $1,723 | ($964) | ($3,653) | ($6,341) |

## 5. Case Study: Pittsburgh

Applying the methodology outlined here to analyze the potential demand for P2P car sharing service in Pittsburgh finds that there are up to 14,460 potential members living in viable markets in that city. The spatial distribution of these markets is shown in FIGURE 1. FIGURE 1 also shows the minimum number of full time cars that would need to be supplied in each market as determined by the queueing analysis described in this paper. The spatial distribution of parked cars available in terms of full-time car equivalents (FTCE) was computed for each census block group. Each FTCE is represents 24 car hours. The required level of market penetration amongst car owners who live and/or work in each census block group was calculated and found to range from a minimum 0.06% to a maximum of 25%. The spatial distribution of this car owner market penetration required is shown in FIGURE 1.

## 6. Conclusion

Given the increasing costs of private car ownership and the increasing worldwide focus on reducing greenhouse gas emissions attributable to the transportation sector, development of both supply side and demand side innovations are required. One such demand side innovation is the mainstream adoption of car sharing. An established stream of academic literature has documented the many benefits of car sharing. However, the current business models of traditional car sharing providers hinder their scalability to areas with lower population density.

P2P car sharing aims to provide a scalable form of car sharing that has the potential to address both the high costs of private vehicle ownership and decrease

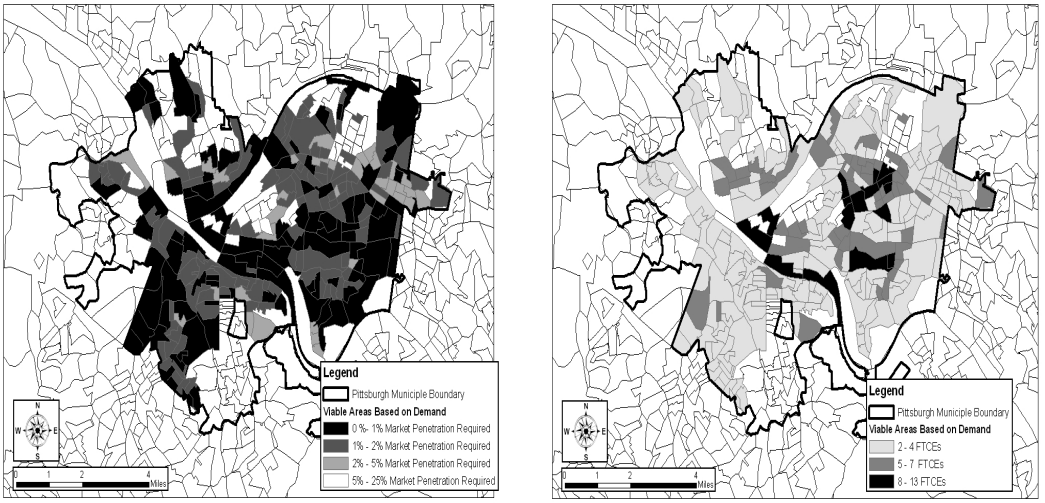



**FIGURE 1: Number of required full time cars and and required market penetration for demand viable areas**

greenhouse gas emissions. This paper develops a methodology to assess the feasibility of P2P car sharing in a given area, the economic incentives of the market participants and the resulting societal externalities.

The framework developed in this paper is a first step in establishing a rigorous framework to analyze the feasibility of P2P car sharing. For the sake of simplicity, assumptions have been made regarding the temporal distribution of demand, and the availability of parked cars. The authors plan to reexamine these assumptions in further works. Also, this paper does not consider the legal or insurance regulatory regimes currently governing P2P car sharing. These topics are important, but beyond the scope of the present paper, and are topics of future research.

## Bibliography

AAA. (2007). *Your Driving Costs 2007, Behind the Numbers*. Retrieved November 2009, from AAA Exchange: http://www.aaaexchange.com/Assets/Files/20073261133460.YourDrivingCosts2007.pdf

Celsor, C., & Millard-Ball, A. (2007). Where Does Carsharing Work?: Using Geographic Information Systems to Assess Market Potential. *Transportation Research Record* , 61-69.

Cervero, R. G. (2007). City CarShare: Longer-Term Travel Demand and Car Ownership Impacts. *Transportation Research Record* .

Grossberg, R., & Newenhouse, S. (2002). *Community Car: A New Transportation Option for Madison, Wisconsin, Carsharing Feasibility Study.* Madison Environmental Group, Inc.

Holtzclaw, J. e. (2001). Location Efficiency: Neighborhood and Socio-Economic Characteristics Determine Auto Ownership and Use – Studies in Chicago, Los Angeles and San Francisco. *Transportation Planning and Technology , 25* (1), 1-27.

Kleinrock, L. (1976). *Queueing Systems.* Wiley.

Millard-Ball, A. M. (2005). *Car Sharing: Where and How it Succeeds, TCRP Report 108.* Transportation Cooperative Research Program.

Noah, G., Koole, G., & Mandelbaum., A. (2003). Telephone call centers: Tutorial, review, and research prospects. *Manufacturing & Service Operations Management* .

Shaheen, S. C. (2005). Car Sharing in North America: Market Growth, Current Developments, and Future Potential. *Submitted for the Transportation Research Board* , 19.

Shaheen, S., Schwartz, A., & Wipyewski, K. (2004). Policy Considerations for Carsharing and Station Cars: Monitoring Growth, Trends, and Overall Impacts . *Transportation Research Record* , 128-136.

Sullivan, E., & Magid, L. (2007). *Bringing Car Sharing to Your Community – Long Guide.* San Francisco: City CarShare.

USBTS. (2007). *Number of U.S. Airecraft, Vehicles, Vessels, and Other Conveyances*. Retrieved November 2009, from U.S. Bureau of Transportation Statistics, Research and Innovative Technology Administration: http://www.bts.gov/publications/national_transportation_statistics/html/table_01_11.html

USEPA. (2005). *Emissions Facts: Greenhouse Gas Emissions from a Typical Passenger Vehicle*. Retrieved November 2009, from US Environmental Protection Agency: http://www.epa.gov/OMS/climate/420f05004.htm

Zipcar, I. (2010). *Form S-1 Registration Statement under the Securities Act of 1933;.* http://www.sec.gov/; Accessed June 2010.

Paper 459

# Urban Development and Traffic Congestion

H. S. Levinson[1] and J.C. Falcocchio[2]

[1]Transportation Consultant, 5305 Ashlar Village, Wallingford, CT, 06492; PH (203) 949-9700; email: hslevingon@aol.com
[2]Department of Civil Engineering, Polytechnic Institute of N.Y.U., 6 Metro Tech Center, Brooklyn, N.Y. 11201; PH (718) 260-3612; email: jfalcocch@poly.edu

## ABSTRACT

The objectives of this paper are to view contemporary traffic congestion as a phenomenon that has always existed within cities throughout history, and to recommend pragmatic modern day solutions to address/manage the congestion problem. This has been achieved by a review of the factors that create congestion, and of the roles that electric streetcars, rapid transit lines, suburban rail lines, and automobiles have played in the decongestion of cities. Examples of outcomes from past transportation improvements are illustrated to show that regional changes in transportation access create new development nodes of activities outside cities that reduce development pressures in the central city (hence reducing congestion growth there) and increase the attraction of peripheral areas to major land developments that lead to increased congestion in these areas. The paper proposes a mix of transportation and land development strategies to minimize the negative effects of urban traffic congestion in the future.

## INTRODUCTION

Traffic congestion permeates metropolitan America. Once limited to city centers and their immediate environs, congestion now permeates the urban region. Attempts to alleviate congestion – such as building new freeways – often generate new land developments that are soon followed by congestion. This paper presents a broad overview of traffic congestion, how it relates to land development, and what could be done to alleviate it, rather than merely transfer it. The focus is on land development policies that would complement transportation improvement options. In many respects, it is a follow-up to a paper entitled Traffic Congestion – Past-Present-and Future published by the ASCE in 1997 (Levinson, 1997).

Understanding and addressing congestion is important for many reasons. Traffic congestion increases delays to users. Equally, if not more important, it increases energy consumption, emissions, and greenhouse gases. And it also influences where activities locate or relocate, and people's choice of travel modes.

## CONCEPTS AND CAUSES

Congestion to the traveler is slow travel with frequent delays. It includes overcrowded conditions on sidewalks, stairs, buses and rail cars. It is backups of slow-moving traffic on freeways and city streets. It can be "recurrent" (take place every day), or non-recurrent (occur when there are crashes or inclement weather). It affects all means of travel. And it responds to changes in various transportation technologies.

*Traffic Congestion* is defined as the travel time or delay that exceeds the time under light or free-flow conditions; it becomes unacceptable when it exceeds a given norm (Lomax, 1997). It can represent the difference or the ratio of travel times or travel time rates (such as the travel time index

*Mobility* represents the ability of people and goods to move easily, quickly and reliably between their origins and destinations. Mobility usually increases when congestion decreases.

*Accessibility* is the ability to achieve travel objectives within an acceptable amount of time. (For example, a trip to work within 30-minutes).

*Basic characteristics* include duration, extent, and intensity. A fourth characteristic, reliability, reflects the variations in the other three.

Congestion occurs when (and where) the demand for travel exceeds the available supply (capacity). It usually results from four basic causes: (1) population and economic growth that outpace the investment in transportation facilities; (2) the spatial separation of home from workplace, school, or shopping center; (3) physical barriers such as hills and water bodies that constrain movement channels and concentrate movements, and (4) operational deficiencies in the street or public transport system.

## GROWTH AND SPREAD

Congestion is not new! It has been found in cities long before the automobile was invented. Throughout history it has responded to changes in technology, and to the concentrations of people in urban areas.

Each means of travel has increased mobility, and influenced land development and the form of cities. When walking was the main means of travel, communities were compact, and congestion was concentrated in the city center. Eighteenth century London is an example.

- The electric street car extended the radius of the city, focused development along street car lines, and spread congestion outward from the city center. Large cities such as Boston and Philadelphia placed their street car lines underground to avoid congestion.

- A handful of cities built rapid transit lines that complemented suburban rail lines in improving mobility. These facilities had the dual effects of further concentrating development in the city center and extending urban development outward along the rapid transit lines. In a few cases, parallel rapid transit lines were built to accommodate the increased demand.

- Automobiles and the roadways that were built to serve them further decentralized development and traffic congestion.

- The changes in transport technology progressively flattened the population density gradient – the decline in population density with increasing distances from the city center. These changes are illustrated in Table 1 that gives illustrative population, and employment densities for pedestrian, electric transit and automobile cities.

The changes that have taken place in the Chicago area over the last century illustrate the dispersion of people, jobs and congestion. Population within the city of Chicago increased from 2.2 million in 1910 to 3.6 million in 1950 and then declined to almost 3 million by 2000. During the 90-year period, the metropolitan area increased from 2.8 million to more than 8 million. The proportion of people living in the suburbs increased from 22% in 1910 to 65 percent in 2000. Congestion reduced on many city streets and increased dramatically in the suburbs.

- In 1910, congestion concentrated in the downtown area and its environs. The relocation of the South Water Street market and the construction of Wacker Drive. Both reduced congestion.

- A 1940 map of traffic congestion in Cook County showed no congestion in suburban areas – including the environs of today's O'Hare International Airport. Today this area has recurrent congestion.

Suburban traffic congestion has progressively emerged in urban America over the last half century. The dispersion of population and commercial activities into areas without transit supportive densities and frequent transit service have increased both motor vehicle travel and congestion. Unimodal site designs and building arrangements further exacerbate the problems – walking is difficult and transit access is limited. Disjointed suburban street systems and frequent driveways contribute to both crashes and congestion. Heavy left-turn movements at major junctions complicate traffic signal phasing and result in excessively long traffic signal cycles.

## REDUCING CONGESTION

Congestion can be reduced by making better operating existing facilities or by building new facilities. In the first case, the gains are usually modest, and there are usually minimum spin-off effects to other areas. In the second case, major increases in accessibility can significantly affect development as congestion shifts from one place to the next.

***Operational Improvements.*** Operational improvements have been used for many years to reduce delays and crashes along both freeways and arterial streets. Reported speed and travel time benefits are shown in Table 1. The gains are usually modest and there are few, if any, spin-off effects to other areas.

- Freeway improvements include providing additional lanes by using shoulders or narrowing lane widths, ramp metering, and providing bus or high-occupancy vehicle lanes. Speed increases of 20 to 30% are common. HOV lanes – especially on approaches to water crossings can save up to 9 minutes per mile.

- Arterial street improvements include restricting curb parking (all-day, or just peak hours), prohibiting left or right turns, providing left-turn lanes, installing one-way streets and effectively coordinating traffic signals. Speeds can increase up to 30%, although in a few cases higher gains are reported.

- Traffic signal coordination improvements can reduce route travel times by up to 25% depending upon the type of initial coordination.

TABLE 1
REPORTED TRAVEL TIME BENEFITS OF
VARIOUS TRAFFIC OPERATIONS IMPROVEMENTS

| | % INCREASE IN SPEEDS | % DECREASE IN TRAVEL TIMES | TIME SAVINGS MIN/MILE |
|---|---|---|---|
| Freeways | | | |
| Additional Lanes by using | | | |
| Shoulder or narrowing lanes | 20 | | |
| Ramp Metering | 20-30 | 30 | 0.3 – 0.4 |
| HOV/Bus Lanes at | | | |
| Water Crossings | | | 3 to 9 |
| Elsewhere | | | 1 to 2 |
| Arterial Streets | | | |
| Curb Parking and Restrictions | 20-30 | 25 | |
| Left-turn Prohibitions | 25-35 | 10 | |
| One-way Streets | 20-50 | | |

***Physical Improvements***. Major transportation facility improvements have three possible effects. They can (1) reduce the extent and duration of congestion; (2) transfer congestion from one location to another, and/or (3) stimulate new development and congestion.

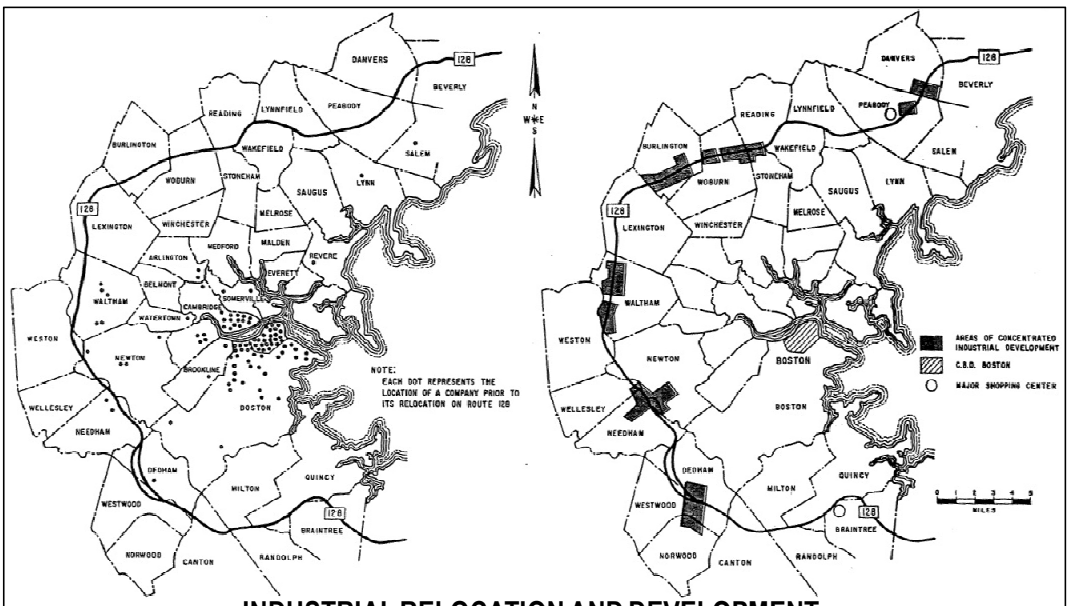
A simple example of transferring congestion is the construction of a "flyover" or "fly-under" at a complex multi-phase intersection. Congestion is transferred to the downstream touchdown point, resulting in a smaller net gain in travel times than initially anticipated.

The building of new rail and road transport facilities has more complex effects. There is a gain in accessibility that fosters new development, and in many cases new congestion elsewhere. Extensions of rapid transit lines from Manhattan into the Bronx transformed farm land into apartments between 1890 and 1930. The growth in capacity was quickly utilized, and parallel lines were built to accommodate the increased demand.

Land development effects have been more pronounced where circumferential express highways were built around cities. These expressways had three major effects: 1. Reduce travel times *around* cities, and reduce congestion *within* the cities, 2. Expand accessibility by making interchange areas attractive for new development, sometimes shifting economic activity from downtown to the beltway environs; and 3. Increase congestion at key interchanges and along beltways.

Examples of suburban traffic congestion are found along many beltways, and their interchanges. (1) The Capital Beltway around Washington D.C. (along with the Tyson's Corners mega complex) is a clear example. (2) The economic and transportation effects of building (and continuously upgrading) Route 128 around Boston (now mainly I-95) are well documented. (See Figure 1). By September 1957, there were 99 new industrial and commercial plants located along the highway, more than 70 plants were previously located within a four-mile radius of Boston (Bone, 1959).

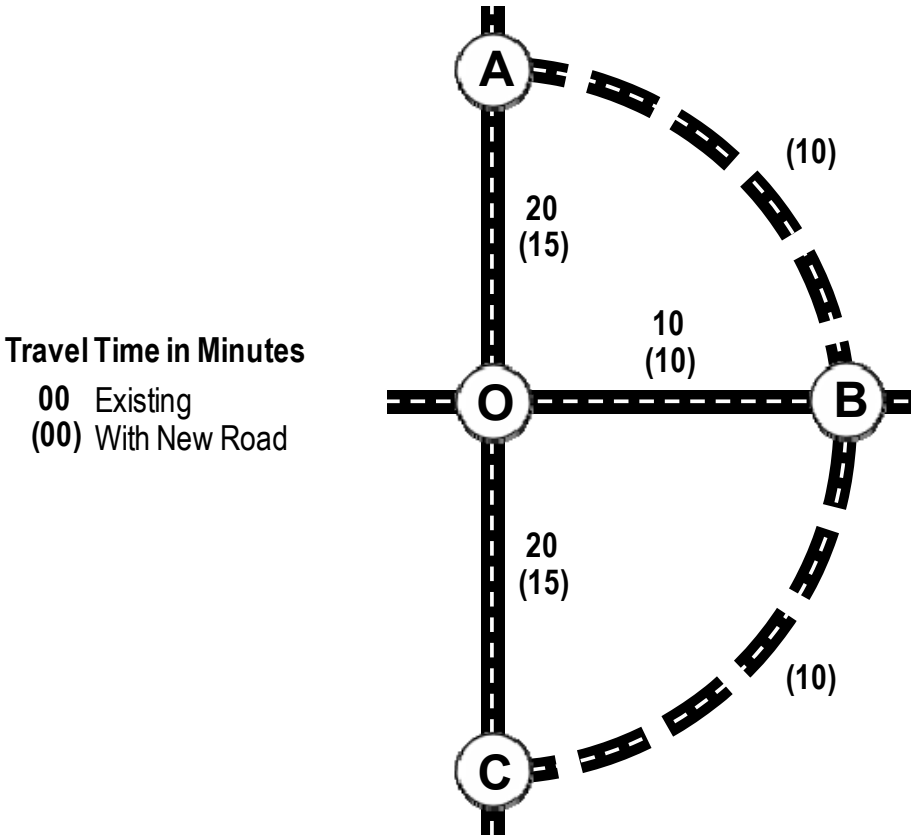
Congestion is also common on radial highways that enter the ring freeways from suburban settings and sometimes within the commercial developments that emerge.




**INDUSTRIAL RELOCATION AND DEVELOPMENT ALONG MASSACHUSETTS ROUTE 128**

FIGURE 1

Figure 2 explains why belt express highways attract new development. A new beltway more than doubles the accessibility of an interchange, as compared with only a modest increase in the city center.




| | Travel Times (Min.) | |
|---|---|---|
| | Before | After |
| O-A | 20 | 15 |
| O-B | 10 | 10 |
| O-C | 20 | 15 |
| A-B | 50 | 10 |
| A-C | 40 | 30 |
| B-C | 30 | 10 |

| | 100 Divided By Sum of Travel Times From All Points | |
|---|---|---|
| | Before | After |
| A | 1.1 | 1.8 |
| B | 1.4 | 3.3 |
| C | 1.4 | 1.8 |
| O | 2.0 | 2.5 |

Source: Estimated

**CONCEPTUAL EFFECTS OF PERIMETER ROAD ACCESSIBILITY**

FIGURE 2

## PLANNING POLICY AND DIRECTIONS

Alleviating congestion, especially in future years calls for new bolder approaches to both development policy and transport improvements.

- From a transportation supply perspective the goals should be capacity, continuity and connectivity. In addition to better traffic engineering and design, there is a need for more effective corridor/access management, better suburban street systems, and more street system connectivity. Public transport should be encouraged where development densities and street patterns permit.

- From a development perspective, the goals should be livability, and sustainability. Public agencies should work together in coordinating transportation and land development policies and actions. These policies, together with selective improvement of roads, transit, and pedestrian circulation can help reduce existing congestion and avoid future congestion.

  It is essential to anticipate development effects of new transportation facilities, and to design the facilities accordingly.

  Suggested actions include:

- Provide public transport and pedestrian opportunities for new developments.

- Foster diverse activity centers that allow people to make multi-purpose trips.

- Anticipate major development at key interchanges along new or expanded freeways, and design roadways and interchanges accordingly.

- Limit access to 1,000 feet or more away from freeway interchanges along arterial roads.

- Require large office developments to be located where there are good public transport and residential catchment areas.

- Concentrate commercial development at key junctions and limit it elsewhere.

- Establish urban design and density controls, including form-based zoning.

- Improve the continuity, connectivity and capacity of the roadway system.

- Extend rail or bus rapid transit into developing areas. Zone for supportive residential and commercial densities and selected suburban stations in advance of development.

- Where feasible develop combined rapid transit lines and freeways in the same corridor (or envelope) to create corridors of high accessibility that can serve as a development spine and catalysts. Consider "multi-modal intercity corridor overlay zones" to reinforce the accessibility framework.

- Utilize value-capture strategies and economic disincentives to help locate and organize development where it can be supported by public transportation.

- Consider road pricing and parking policies, including CBD parking space limitations where there is off-street transit.

## LOOKING AHEAD

Traffic congestion will likely grow and disperse as urban regions expand. Coping with this congestion calls for concerted physical and operational improvements and public policy actions. As new transport facilities are built, or existing ones are expanded, development pressures will result in congestion where new markets are created and new sites are made accessible. Anticipatory development planning, multi-modal transport policies, and designs for livability, walkability and public transport will be needed to better manage congestion in the years ahead. Concerns about greenhouse gases, and climate change will make such actions essential. The goal is integrated transportation and land development for a better tomorrow.

## REFERENCES

(1) Bone, A.S., and M. Wohl "Massachusetts Route 128 Impact Study", Bulletin 227, Highway Research Board, National Research Council, Washington D.C., 1959.

(2) Levinson, H.S., T.J. Lomax, and S. Turner. "Traffic Congestion – Past – Present – Future" in Traffic Congestion and Traffic Safety in the 21st Century, challenges, Innovations and Opportunities, Proceedings of the Conference sponsored by the Urban Transportation Division and Highway Division, ASCE, R.H. Benekohal, ed., Chicago, Illinois, June 8-11, 1997.

(3) Levinson, Herbert S., "Operational Improvements – An Overview" in UTRC Research News – Region 3, University Transportation Research Center, City University of New York, Winter 2005.

(4) Lomax, T., S. Turner, G. Shunk, H.S. Levinson, R.H. Pratt, P. Bay, B. Douglas, NCHRP Report 398, Quantifying Congestion, Volume 1, Final Report, Transportation Research Board, National Research Council, Washington D.C., 1997.

# Relating Land Use and Transport Modeling with Transportation Asset Management

**Luis Esteban Amador-Jiménez**[1]

[1]PhD, Assistant Professor, Dept. of Building, Civil and Environmental Engineering, Concordia University, Montreal, PQ, Canada, H3G 1M8. Email: amador@.encs.concordia.ca

## ABSTRACT

Increasing computing capacity has made possible the use of advanced simulation and optimization methods to solve complex problems in transportation and urban planning. Methods such as linear programming and heuristic optimization have made it possible to find the most cost-effective solution for allocating resources among competing alternatives. Agent based approaches and regional econometric models have succeeded to simulate urban development, economic growth and their impacts across modes of transportation. This paper examines the advantages of incorporating Land Use and Transport Modeling into Asset Management and presents a framework for accomplishing such an objective; especially for timing maintenance, rehabilitation and capital projects to improve safety, mobility and condition. It suggests that decisions for scheduling rehabilitation of existing links and construction of new projects should be based on user impacts and expected gains in terms of regional economic benefits.

**CE Database subject headings**: transportation asset management; integrated land use and transportation; decision making analytical tools, capital projects; mobility; safety; condition.

## Introduction

Even though literature on pavement management can be traced back to the 1960s (Haas and Hudson 1994), it was not until the 1990s that pavement management systems started to be adopted in developed countries (TAC 1997). Transportation asset management (TAM) seems to have evolved from pavement management systems. Earliest developments of transportation asset management date back to the 1970s with progressive legislation found in the Local Governments Act of New Zealand. In 1995 the National Asset Management Steering Group from New Zealand became one of the first organizations to formally adopt such methods for the management of civil infrastructure (NAMS 2006). A few years later, the Federal Highway Administration (1999) followed a similar path identifying transportation asset management as the decision making process to achieve long-term sustainability of transportation assets. The FHWA (1999) along with the NAMS (2006), recognized that modern civil infrastructure's asset management is an approach that

incorporates the economic assessment of trade-offs between competing alternatives, and that information is the critical factor to make cost-effective decisions (Ouertany *et al.* 2008).

However, to date practical implementations of Transportation Asset Management (TAM) have faced several limitations (NCHRP 2005): (1) Static performance modeling that fails to recognize the dynamic nature of a changing network of assets (*i.e.,* condition, safety and capacity) that in reality respond to changes on the demand, which in turn represents the evolution of urban patterns and establishment of new industries and economic development. (2) Analytical tools with limited capability for incorporating multiple objectives across modes of transportation. (3) Other practicalities related to the need of a modeling mechanism capable of capturing treatment effectiveness from historical observations.

According to Batty (1979), initial developments on integrated Land Use and Transportation (LUT) date back from the 1950s. However, it wasn't until the 1960s with Lowry (1964) that the first operational model, based on spatial interaction and gravity theory, appears (Iacono *et al.* 2008). During the 1980's a second generation of models based on random utility (*i.e.,* multinomial and nested logit) exhibited better representation of urban dynamics incorporating more detailed modeling of spatial economics and land market changes. More disaggregated models (*i.e.,* agent or cell based) appeared during the 1990s modeling urban dynamics from a bottom up approach. Since then, advances on LUT modeling have attempted to refine the simulation in order to reproduce urban development and its impacts on the transportation network (Hunt *et al.* 2005).

## *Paper Objective*

To propose an improved decision making framework capable of supporting a more comprehensive Transportation Asset Management by incorporating simulation capabilities from Land Use and Transport modeling into Asset Management.

## *Review of TAM and LUT models*

### Transportation Asset Management

Recent advances on TAM have focused on improving analytical tools to achieve the most cost-effective solution for maintenance and rehabilitation of civil infrastructure. Today's state of the practice on infrastructure asset management relies on trade-off global optimization for selecting the optimal set of actions among competing alternatives to maintain, rehabilitate and upgrade networks of infrastructure assets (NCHRP 2005). It's well known than applying an adequate treatment at the right time will positively impact the service life of any asset. While some treatments translate into a rejuvenation of the asset, others only slow down the deterioration rate (Figure 2). However, treatment availability depends on asset type; while for pavements there is a wide range of options across different stages of the lifespan (Figure 3); for water

systems and pipe networks there are very few. In addition, more research is required to formally document the effectiveness and post-treatment performance of any asset.

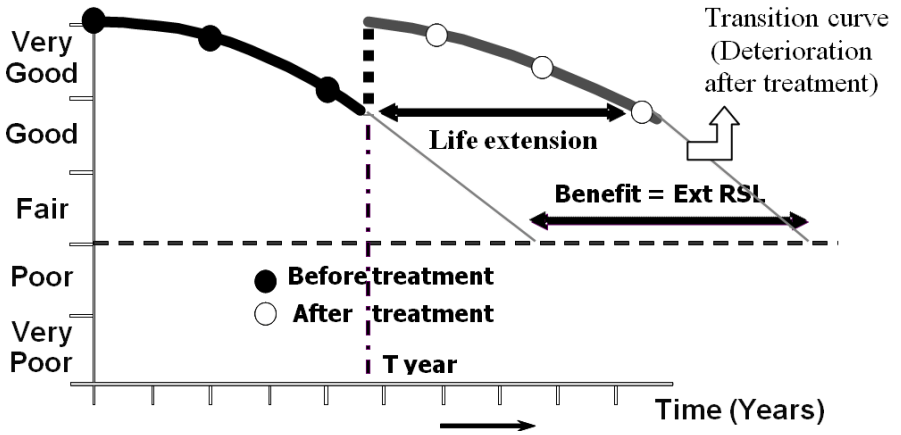



*Figure 2. Performance and Treatment effectiveness*

Once treatments have been identified, and the objectives selected, a decision making analytical tool chooses among competing alternatives. Historical, advances in decision making tools can be traced back to the 1980s with some state level implementations of pavement management systems. Some tools such as the PAVER and the HDM3 were based on cost-benefit analysis for single periods of time with no consideration on how today's decisions impact the long term achievement of agency's goals. In addition, such models were incapable of trade-off analyses across asset types. Another drawback was the large amount of data required to calibrate these models.

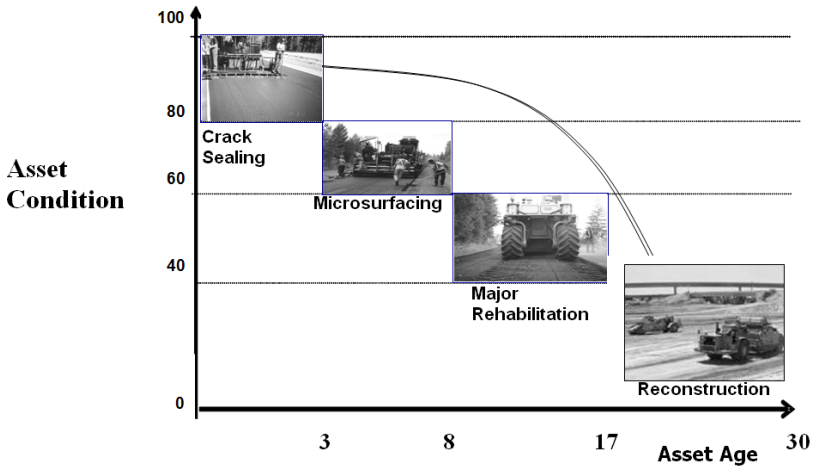



*Figure 3. Sample of common treatments for pavements*

The adaptation of linear programming and other optimization techniques for asset management addressed most of those issues. Linear optimization was capable of finding the optimal path (Figure 4) to take full advantage of cost-effectiveness of individual treatments (what treatment?), associated with individual asset elements

(where?), and benefits of advancing or deferring a certain treatment (at what time?). It will seek an allocation that will minimize costs (or maximize the benefits, or any other measures of return on investment) over the whole network of assets in the long run.

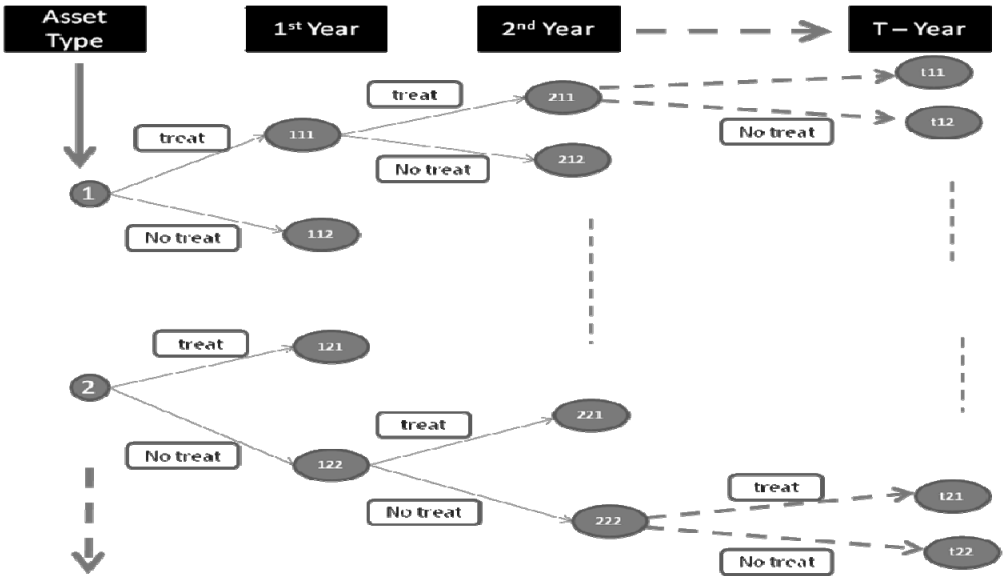



**Figure 4. Graphical representation of possible paths**

Today's state of the art in civil infrastructure management systems is advancing on two fronts: more reliable performance modeling and improved techniques for decision making capable of allocating actions across different networks of assets, across time and for several objectives. Such decision-making tools can even be expanded to make decisions across different modes of the transportation. Table 1 summarizes historical advances on civil infrastructure management systems.

**Table 1. Historical comparison of Infrastructure Management Systems**

| Time | Description | Analytical Tools | Sample Analytical tool | Performance Modeling |
|---|---|---|---|---|
| 1980s | Pavements only. Some with GIS interface | Cost-benefit analysis | HDM3, HERST, PAVER | Deterministic |
| 1990s | Originally designed for pavements, extended to Civil Infrastructure | Linear Programming Optimization | VEMAX | Stochastic (Semi-Markov) |
| 2000s | Adapted from Forestry, Goal /weighted objectives, custom defined modules. GIS interface | | REMSOFT | Deterministic |
| 2010s | Across assets, multiple objectives, across assets, spatial-temporal coordination of activities | Heuristic - Evolutionary Algorithms | Under Development | Mechanistic-with some empirical, Stochastic (MCMC-Bayesian), Nested |

## Land Use and Transportation modeling

Land use and transportation models strive to capture processes of change on land development due to economic interactions & development in order to forecast levels of demand across modes of the transportation network. Two main approaches have been historically employed to simulate spatial dynamics and transport modeling: (1) Regional economic models and, (2) Micro simulation models. Although both pursue the same objective, their approach departs from opposite ends. On one hand, regional economic models base their simulation on a top-down approach in which macro aggregated trade zones are used to predict flows of commodities, which are then converted into demand for commercial and passenger traffic. Congestion from the transport network feed back into the original trade system by increasing a disutility function which penalizes modal links.

On the other hand, micro simulation models use a bottom up approach, which departs from disaggregated agents and small economic units to reach a macro-level of interactions in which trade flows between economic activities and levels of demand for the transportation network are determined. Other forms based on cellular automata and complexity theory return to the original aims of spatial interaction models attempting to capture the spatial dynamics of a region disregarding close form econometric relationships. Table 2 summarizes advances on Land Use and Transport modeling.

**Table 2. Advances on Land Use &Transport models**

| Generation | Description | Theory | Model | Complexity |
|---|---|---|---|---|
| 1st – 1960s | Spatial Interactions | Gravity-based | ITLUP, IRUPD | Low; employ coefficients for attractiveness |
| 2nd –1980s | Regional Economics | Random Utility Multinomial Logit / Nested Logit | TRANUS, MEPLAN | Medium; required aggregated census data at zone level per industry |
| | Land Market | | METROSIM, MUSSSA | |
| 3rd – 1990s | Agent-cell Based | Micro-simulation, Cellular Automata | PECAS, ILUTE Urban Sim, MALUT, LUCI2 | Required detailed disaggregated information from census data |

While more accurate for metropolitan (city) wide level, micro simulation models suffer from drawbacks related to calibration for which they are data hungry. Rather, regional economic models are less complex systems that perform well on a regional basis. This paper recommends the use of regional econometric models for state/province wide implementations and micro-simulation based approaches for city level (metropolitan region) urban systems that are by nature dynamic and complex (Batty 2005, Iacono *et al.* 2008).

The key steps of Land Use and Transport modeling are presented on Figure 5; following a regional economics approach, an initial calibration is performed in order

to fit observed data by estimating a set of factors, checking model agreement with production and land prices by sectors. Then, future location of activities is predicted along with floorspace and land values estimates (de la Barra 2005). Secondly, trade flows are estimated and converted into demand for commercial and passenger traffic. Such demands are allocated across transportation modes to a set of paths or travel options. Demands load different modes of the network, which in turn translate into increased travel time and other disutilities that affect the location of activities. A model based on micro-simulation undertakes a bottom up approach (Iacono *et al.* 2008) in which the urban dynamics are disaggregated into agents and units, therefore the model is initially calibrated to fit observations at that level and other macro activities are generated by aggregating activities to regions (zones) and industries. The rest of the process goes in similar fashion as the abovementioned.

## A framework for improved decision making on TAM

As aforementioned, one of the key features of any Land Use and Transport models is its capability to update the model parameters based on travel disutilities associated to particular links and network modes. Travel time is currently the most employed function to capture disutility (Iacono *et al.* 2008, Hunt *el al.* 2005, De la Barra 2005). The first variant proposed by this paper is to modify the disutilities function to incorporate measures of mobility, condition and safety. As shown on Figure 5, mobility translates not only to travel time but also to vehicle operating cost because it accounts for the level of service on any link (*i.e.,* congestion), fuel consumption and vehicle deterioration.

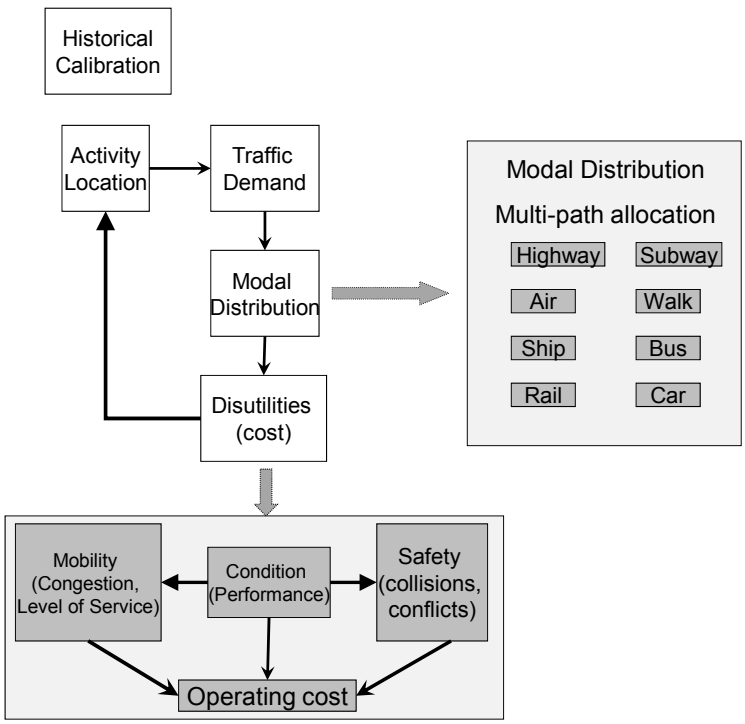



**Figure 5. Conceptual LUT Regional Economic Model**

However, existing LUT models do not incorporate safety and its impacts on travel behavior. There is sufficient research to prove that drivers perception of road safety is a factor in the choice of routes and links for passenger cars (Hildebrand 2003). Another argument can go towards increased travel time as a function of the number of conflicting points (TAC 1999). Finally, drivers in general would tend to choose those roads with better surface condition (*i.e.,* low roughness). Lower levels of condition negatively impact safety and mobility; drivers would be discouraged to drive at maximum operational speed if the ride is rough, which translates on increased travel time. Drivers would tend to avoid damage on the road surface (*i.e.,* potholes, settlements, etc), which may turn into unexpected behavior compromising road safety. In summary all three factors would likely impact travel time, vehicle operating cost and insurance premiums, which in turn translate into additional production cost. This in general goes in detriment of the overall productivity of any industry.

This paper proposes to measure and incorporate the performance on the three factors as part of an improved TAM system (Figure 6). For the same reason projects that have been identified as candidates to improve any of these aspects (*i.e.,* mobility, condition and safety) will be incorporated in the pool of assets, carrying a special index to recognize the fact that they are still not part of the network (Figure 6). The decision making tool will treat them as belonging to a new dimension when generating the full combinatorial space of treatments across time in order to select the optimal path as explained in Figure 4.

The other key element for incorporating LUT into TAM is their capability to measure the economic impact of any new project and the overall impact in the network (Figure 5).

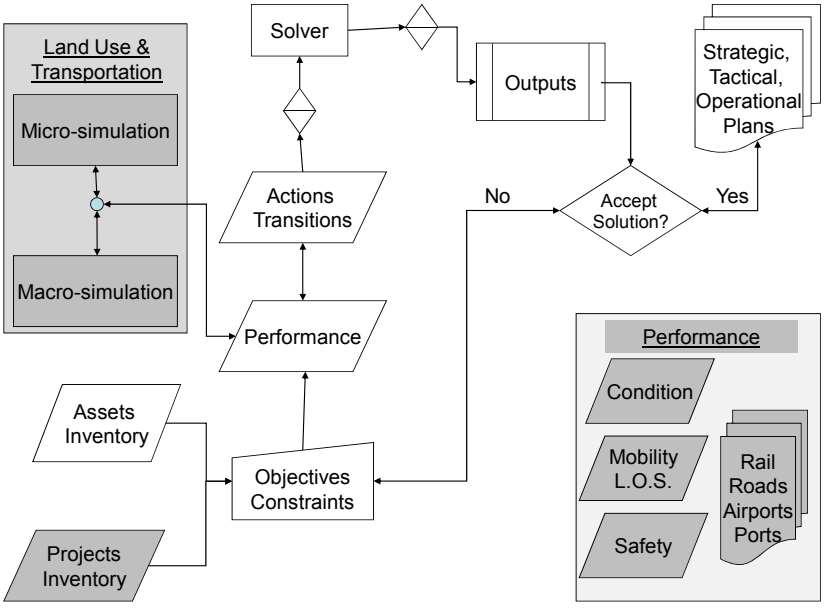



**Figure 6. Improved TAM system which incorporates LUT modeling capabilities**

This key feature will allow planners to improve their decision making by realizing of the cost effectiveness of incorporating new links in the network and how far the benefits of such capital projects will go. Another important consideration is that of redistribution of the demand across links of the network after the construction of a new project and how that affects condition, safety and mobility. Therefore, the individual contribution of any project will be assessed within the decision making analytical tool in terms of economic contribution and improvement in safety, condition and mobility, which are typical objectives on any Asset Management system (Figure 6). Such elements would be incorporated and updated on the Performance model subcomponent of the TAM.

# Conclusions

The integration of LUT modeling capabilities within a TAM system will support a more comprehensive multimodal decision making framework in which decisions are allocated among competing alternatives for improving safety, condition and mobility in order to maximize the cost-effectiveness of the transportation network and the economic benefits (*i.e.,* productivity) of local industries.

# References

Batty, M. (1979). "Progress, success and failure in urban modeling". *Environment and Planning A*, volume 11(8), pp 863-878.

Batty, M. (2005). "Cities and Complexity: Understanding Cities with Cellular Automata, Agent-based models and Fractals". *MIT Press Cambridge*, Massachusetts.

Batty, M. (2008). "Cities as Complex Systems: Scaling, Interactions, Networks, Dynamics and Urban Morphologies*"*. Centre for Advanced Spatial Analysis University College London, Paper 131, London, U.K.

De la Barra, T. (2005). *"*Integrated land use and transport modeling: decision chains and hierarchies*"*. *Cambridge Urban and Architectural Studies*. 1st paperback version. Cambridge University Press, New York, NY.

FHWA. (1999). "Asset Management Primer". Office of Asset Management, Federal Highway Administration, Washington, DC., USA.

Haas, R., Hudson, W.R., Zaniewski, J. (1994). "Modern Pavement Management". Krieger Publishing Company. Malabar, Florida.

Hildebrand, E. (2003). "Dimensions in elderly travel behaviour: A simplified activity-based model using lifestyle clusters". *Journal of Transportation*, Springer Netherlands, pp 285 – 306.

Hunt, J.,D., Kriger, D.,S., and Miller, E.,J. (2005). "Current Operational Land-use- Transport modeling Frameworks: A Review*"*. *Transport Reviews*. Volume 25. No-3. pp 329-376.

Iacono, M., Levinson, D., and El-Geneidy, A. (2008). "Models of Transportation and Land Use Change: A Guide to the Territory". *Journal of Planning Literature*, number 22, pp 323-340.

Lowry, I. S. (1964). A model of Metropolis. Memorandum RM-4035-RC. Rand Corporation. Santa Monica, CA.

NAMS. (2006). "International Infrastructure Management Manual*"*. *National Asset Management Steering Group*. New Zealand.

NCHRP. (2005). "Analytical tools for Asset Management". Transportation Research Board. National Cooperative Research Program. Report 545. Washington D.C.

Ouertani, M.Z., Parlikad, A.K., and Mcfarlane, D. (2008). "Towards an approach to select an asset information management strategy". International Journal of Computer Science and Applications. Volume 5, number 3b. pp25 – 44.

TAC. (1997). "Pavement Design and Management Guide", Transportation Association of Canada, National Project Team, Ottawa.

TAC. (1999). Geometric Design Guide for Canadian Roads. Transportation Association of Canada. Ottawa.

# Case Study:
# Assessment of the Vulnerability of Port Authority of NY & NJ Facilities To the Impacts of Climate Change

Brian J. McLaughlin, P.E., LEEP AP[1], Scott D. Murrell, P.E.[2]
and Susanne DesRoches LEEP AP[3]

[1]Senior Civil Engineer, Port Authority of NY & NJ, 2 Gateway Center – 16th Floor, Newark, NJ 07102; PH (973) 792-4361; FAX (973) 792-4303;
email: bmclaugh@panynj.gov
[2]Chief Civil Engineer, Port Authority of NY & NJ, 2 Gateway Center – 16th Floor, Newark, NJ 07102; PH (973) 792-4327; FAX (973) 792-4303;
email: smurrell@panynj.gov
[3]Sustainable Design Manager, Port Authority of NY & NJ, 2 Gateway Center – 16th Floor, Newark, NJ 07102; PH (973) 792-3559; FAX (973) 792-4303;
email: sderoch@panynj.gov

## Abstract

The Port Authority of NY & NJ (PANYNJ) is a bistate agency, which conceives, builds, operates and maintains infrastructure critical to the New York/New Jersey region's trade and transportation network. These facilities include America's busiest airport system, marine terminals and ports, the PATH rail transit system, six tunnels and bridges between New York and New Jersey, the Port Authority Bus Terminal in Manhattan, and the World Trade Center. For more than eight decades, the Port Authority has worked to improve the quality of life for the more than 17 million people who live and work in New York and New Jersey - a region that supports 8.6 million jobs with an estimated gross regional product of more than $929 billion.

The release of projections for climate change variables published by the Intergovernmental Panel on Climate Change (IPCC) has resulted in greater public awareness and has lead to the launch of state and local government initiatives to evaluate the vulnerability of critical infrastructure. Since PANYNJ facilities are vital to the New York/New Jersey regional economy, their safe and efficient operation is of the utmost importance and an assessment of climate changes impacts was warranted. Many of these facilities are located in close proximity to coastal waters, which further magnifies the potential impacts of climate change variables, such as sustained sea level rise with associated storm surge, increased precipitation and increased temperature.

The following case study details the involvement of the PANYNJ in a climate change assessment effort, lead by the New York City (NYC) Mayor's Office of Long Term Planning and Sustainability, between August 2008 and March of 2010. This case study will provide an overview of the evaluation process with an emphasis on the coastal flooding potential and possible adaptation measures.

## Introduction

The climate change assessment team, in which the PANYNJ participated, was lead by the city of New York, and titled "Climate Change Adaptation Task Force" (CCATF). This effort is part of a comprehensive sustainability plan for NYC, called PlanNYC. As part of this effort, the PANYNJ evaluated the vulnerability of its infrastructure to climate change variables to identify at-risk infrastructure, considering projections for three decades, the 2020s, 2050s and 2080s. This case study will examine: the process by which PANYNJ facilities were evaluated for vulnerability, how they were categorized by degree of risk, and the types of climate adaptation strategies that were developed. A key aspect of this study is the lessons learned during this evaluation, which will benefit both public and private sector engineers, administrators, owners and operators involved with transportation infrastructure, as similar assessments and studies are likely to become more commonplace in the future. Additionally, this case study will provide insight into the challenges of undertaking such an effort and how climate change projections can inform design guidelines, maintenance programs and long-term planning.

## Overview of NYC CCATF

The two main goals of the NYC CCATF were as follows: identify critical infrastructure that could be at-risk from the effects of climate change and develop coordinated adaptation strategies to secure these assets and incorporate climate change adaptation into long-term planning processes.

The NYC CCATF was comprised of five working groups: Energy, Policy, Transportation, Communication, and Water and Waste. The PANYNJ participated in the Transportation Working Group along with other public and private entities that own or operate transportation infrastructure within the NYC region. The Transportation Working Group members also included AMTRAK, CSX, Metropolitan Transportation Authority (MTA), New Jersey (NJ) Transit, New York State Department of Transportation (NYSDOT) and New York City Department of Transportation (NYCDOT). Individual working group members were tasked with creating an inventory of at-risk infrastructure and developing adaptation plans, while the overall goals of the Transportation Working Group were aimed at identifying/creating synergies among stakeholders by developing design guidelines for new infrastructure and identifying citywide strategies for further study.

## Assessment Process

The assessment process can be categorized into the following six major tasks: definition of climate change variables and projections, development of asset inventories, vulnerability assessment, risk analysis, prioritization of assets and development of adaptation strategies.

The listed order of these tasks represents the general sequence in which they were performed, although there was some degree of iteration involved between the various steps.

## Climate Change Variables

The New York City Panel on Climate Change (NPCC) provided guidance, analysis and information on climate change variables and associated risks. The NPCC is composed of academic and research institutions including Columbia University and the NASA Goddard Institute for Space Studies. NPCC provided the scientific expertise and technical tools for this effort in the form of a publication titled "Climate Risk Information". This publication describes the projections and probabilities associated with climate change variables including temperature, precipitation, sea level rise and coastal storms. Figure 1 shows a summary of the projections provided for these climate variables for the 2020s, 2050s and 2080s.

Midway through the effort, the NPCC provided subsequent sea level rise projections accounting for possible accelerated rates of ice melt in the Greenland and West Antarctic Ice sheets, described as the "Rapid Ice-Melt Scenario". Due to the uncertainty associated with these projections and the timeliness in which they were provided, these projections were not incorporated into this process. These projections are shown in Figure 1.

| **NPCC Climate Projections** *(Rosenzweig, C. & Solecki, W. 2010)* | | | | | |
|---|---|---|---|---|---|
| | | **Baseline (1971-2000)** | **2020s** | **2050s** | **2080s** |
| **Air Temperature** | | 12.8 ºC (annual mean) | + 0.8 – 1.7 ºC | + 1.7 – 2.8 ºC | + 2.7 – 4.2 ºC |
| **Precipitation** | | 118.1 cm (annual mean) | + 0 – 5% | + 0 – 10% | + 5 – 10% |
| **Sea Level Rise** | | NA | + 5.1 – 12.7 cm | + 17.8 – 30.5 cm | + 30.5 – 58.4 cm |
| **Coastal Storms** | **1-in-100 yr Flood to recur** | ≈ once every 100 years | ≈ once every 65 to 80 years | ≈ once every 35 to 55 years | ≈ once every 15 to 35 years |
| | **1-in-500 yr Flood to recur** | ≈ once every 500 years | ≈ once every 380 to 450 years | ≈ once every 250 to 330 years | ≈ once every 120 to 250 years |
| **Rapid Ice-Melt Sea Level Rise Projections** *(Not considered in PANYNJ Infrastructure Assessment)* | | | | | |
| **Sea Level Rise** | | NA | + 12.7 – 25.4 cm | + 48.3 – 73.7 cm | + 104.1 – 139.7 cm |

**Figure 1**

## Asset Inventory

The first step in the NYC CCATF effort involved generating a list of assets that may be vulnerable to the climate change variables provided by NPCC. This was quite an undertaking for the PANYNJ, since it operates airports, marine terminals, vehicular tunnels, major bridges, a rail transit system, bus stations and many other smaller facilities. The PANYNJ Engineering Department developed an initial list of the major facilities that may be at-risk to impacts from the climate change variable projections provided by NPCC. The initial list consisted mainly of facilities and any critical elements within those facilities that would affect safety or disrupt operations. The preliminary list was circulated to staff throughout the various agency departments to ensure thoroughness. Some examples of the major assets of the PANYNJ include: John F. Kennedy International Airport, Newark Liberty International Airport, LaGuardia Airport, Stewart International Airport, PATH rail transit system, Lincoln Tunnel, Holland Tunnel, George Washington Bridge, OuterBridge Crossing, Goethals Bridge, Bayonne Bridge, World Trade Center site, Port Newark & Elizabeth Marine Terminals and Howland Hook Marine Terminal.

## Vulnerability Assessment

When selecting potential candidates for classification as at-risk infrastructure, specifically for flooding potential, the criteria used included: site topography (average elevation), proximity to coastal waters, performance during past weather events (i.e. heavy rainfall) and criticality of the facility. During this phase of the analysis, we discovered that the availability of information and ease of access for specific infrastructure is crucial. As is the case with most state agencies that manage and operate transportation infrastructure, the PANYNJ information archive systems are extensive, making retrieval time consuming. Additionally, there are multiple sources of information stored in multiple locations. Intangible sources of information such as institutional knowledge can be very valuable and these sources were also tapped through discussion with facility staff during this effort.

### Information Gathering

Since the majority of PA facilities are situated near the coast, the major threat to our infrastructure was found to be the storm surges cased by extreme events, such as hurricanes or extra tropical cyclones (nor'easters) coupled with sea level rise; therefore, it was crucial to obtain information on site topography. Elevation information was collected from record drawings, although it may not truly represent the as-built condition. Where available, data acquired from topographic site surveys, was used for this evaluation. Based on the challenges involved with gathering record information, we recommend the use of topographic site surveys for any entity considering this type of effort, when record drawings are not readily available. This may result in a cost savings and more accurate data.

### Sea Level Rise and Coastal Storms

Topographic maps and Federal Emergency Management Act (FEMA) Flood Insurance Rate Maps (FIRM) were useful tools for determining the flood potential for facilities. It is important to note that traditionally, FEMA flood maps have been used

to identify the areas that are vulnerable to the 100-yr. flood level, but during our participation in this effort, we discovered that topographic maps acquired through low altitude flyover yielded flood limits that differed from those shown on FEMA flood maps. This difference can likely be attributed to the higher precision of localized topographic surveys. These various sources of flood plain data were compared to identify facilities situated in the vicinity of the current or projected 100-yr flood plain for classification as at-risk infrastructure.

**Temperature and Precipitation**

PA facilities handle millions of patrons each year; therefore, the potential hazards of increased temperature and precipitation were considered. The impacts of the effects of extreme events such as more frequent heat waves and more intense rainfall were considered in assessing the vulnerability of building systems and drainage systems for roadways and airfields. Building components such as HVAC and electrical systems typically have a useful life span of less than 25 years and therefore were not considered at-risk due to the opportunity to improve these systems during replacement cycles.

See Figure 2 for a partial list of the at-risk PANYNJ facilities identified during this evaluation.

| SAMPLE INVENTORY OF AT-RISK INFRASTRUCTURE | | | |
|---|---|---|---|
| **Classes of Infrastructure** | **Specific Infrastructure Locations or Sub-classes of Infrastructure** | **Climate Variable** | **Description of Climate Change Impact** |
| **Berth & Roadway** | Piers, Slips and Roadways at Port Newark, Elizabeth Port Authority Marine Terminal, Howland Hook, Brooklyn, Hoboken Waterfront | Extreme events and storm surge with sea level rise | Increased flooding risk due to nor'easters and hurricanes |
| **Rail** | PATH Stations and open/subsurface Track - PATH Tunnels | Extreme events and storm surge with sea level rise | Increased flooding risk due to nor'easters and hurricanes |
| **Airfield** | John F. Kennedy International Airport, LaGuardia Airport - Runways and Taxiways | Extreme events and storm surge with sea level rise | Increased flooding risk due to nor'easters and hurricanes |
| **Tunnel** | Holland Tunnel | Extreme events and storm surge with sea level rise | Increased flooding risk due to nor'easters and hurricanes |
| **Bridge** | Bayonne | Sea level rise | Sea level rise will lower existing bridge clearances |
| **Building** | Terminal Buildings at John F. Kennedy International Airport & LaGuardia Airport | Rising temperatures/ heat waves | Increased risk for power failures shutting down baggage handling systems |

**Figure 2**

## Risk Analysis

The risk analysis done for each asset was a function of "Likelihood of Occurrence" of impact on infrastructure occurring during an assets useful life and the "Magnitude of Consequence" of impact on that infrastructure. The "Magnitude of Consequence" was evaluated over the following six sectors: internal operations, capital and operating costs, society, health, economics and environment. The "Magnitude of Consequence" was classified at levels of low, medium or high; however, this evaluation was not quantitative. "Likelihood of Occurrence" was classified at degrees of low, moderate, high or virtually certain/already occurring. The key concept when classifying the "Likelihood of Occurrence" is that the impact must occur within the lifetime of the infrastructure. This type of analysis can be quite subjective without the use of established quantitative criteria to differentiate the various degrees. Since the main threat to PANYNJ facilities was found to be extreme events in the form of coastal storms, PANYNJ Engineering established a quantitative scale to assign the degrees of "Likelihood of Occurrence". Figure 3 shows the PA likelihood definitions scale for extreme weather events. Historically the probability of occurrence of flood levels caused by coastal storms has been expressed in terms of a return period of years, i.e. a flood level that has a 1% probability of occurrence during a given year is often described as the 1 in 100-year flood level. As a starting point, we defined any likelihood less than or equal than 1% as a low likelihood. Refer to Figure 3 for a summary of the PANYNJ likelihood definitions and associated probabilities and return periods.

| PA "LIKELIHOOD OF OCCURRENCE" DEFINITIONS FOR EXTREME WEATHER EVENTS | | | |
|---|---|---|---|
| **Likelihood** | **Definition** | **Probability of Impact Range** | **Return Period of Impact** |
| **Virtually certain / Already occurring** | Nearly certain likelihood of the impact occurring over the useful life of the infrastructure and/or variable may already be impacting infrastructure | > 20% | 1 in 5 yrs. or less |
| **High** | High likelihood of the impact occurring over the useful life of the infrastructure. | > 10% and ≤ 20% | 1 in 5 yrs. to 1 in 10 yrs. |
| **Moderate** | Moderate likelihood, with some uncertainty remaining, that the impact will occur over the useful life of the infrastructure. | > 1% to ≤ 10% | 1 in 10 yrs. to 1 in 100 yrs. |
| **Low** | Low likelihood of the impact occurring over the useful life of the infrastructure. | ≤ 1% | 1 in 100 yrs. or greater |

**Figure 3**

## Prioritization of Assets

The prioritization of at-risk assets was determined by using the risk matrix shown in Figure 4. An asset having a high "Likelihood of Occurrence" and high "Magnitude of Consequence" would call for the immediate development of strategies to mitigate impacts. Based on the probability ranges defined by the PANYNJ for extreme events, assets that were at risk of impact by this event were classified as a moderate "Likelihood of Occurrence" level. The overall "Magnitude of Consequence" for PANYNJ facilities across the six sectors was estimated at a moderate level based on discussion with PANYNJ staff from various departments with knowledge of these sectors. When evaluating the "Magnitude of Consequence" impact across the six sectors for PANYNJ facilities an important consideration was that during a major storm event populated areas will be evacuated and transportation systems would be operating on a limited, emergency basis. Airport, berth, and rail transit facilities were found to have the highest risk and therefore were assigned the highest priority for the development of adaptation strategies. These facilities were found to have a more significant impact on regional commerce. Less critical transportation infrastructure such as localized roadways, vehicular tunnels and buildings were assigned a lesser risk level and fell under the category of "Watch", or monitor.

| **RISK MATRIX** | | | | | |
|---|---|---|---|---|---|
| **Magnitude of Consequence** | **High** | Evaluate Further / Develop Strategies | Develop Strategies | Develop Strategies | Develop Strategies |
| | **Medium** | Watch | Evaluate Further / Develop Strategies | Develop Strategies | Develop Strategies |
| | **Low** | Watch | Watch | Evaluate Further / Develop Strategies | Evaluate Further / Develop Strategies |
| | | **Low** | **Moderate** | **High** | **Virtually Certain** |
| | **Likelihood of Occurrence** | | | | |

**Figure 4**

## Adaptation Strategies

Adaptation strategies were divided into three categories: maintenance & operations, capital investments and regulatory. Maintenance and operation type strategies include: sandbagging, portable pumps, temporary flood gates, cleaning of drainage systems, repositioning of rolling stock and performance of more detailed studies. Capital investments are permanent improvements and include: installation of new flood barriers; elevating elements of critical infrastructure to levels above projected flood elevations; relocation of critical facilities to higher ground, and designing new

assets for quick restoration after an extreme event. Regulatory strategies include modification of city building codes and design standards.

**Engineering Redundancy**

An interesting aspect to this study was the discovery that past capital improvement investments that involved engineering design redundancy have the unintended consequence of mitigating against future impacts that may be induced by climate change. Security projects that involve the construction of barriers and walls have also provided further protection against the future impacts of storm surge.
The following two examples show how engineering design redundancy provides the added benefit of mitigating against future climate change impacts.

The first example involves the asphalt concrete pavement mixes used by the PANYNJ. Over the years, PANYNJ Civil Engineering and Materials Division have developed asphalt mixes that contain modified polymers for better performance under sustained heavy truck traffic and aircraft loads. These modified asphalt mixes also have the added benefit of performing well through a much higher temperature range.

The second example involves a project at LaGuardia Airport that constructed an interconnection between electrical substations. LaGuardia Airport requires significant electrical power to ensure efficient and reliable operation with a peak load power demand of twenty million volt-amperes. The airport is supplied power via two independent local utility company (Con Edison) networks that feed two airport substations to provide redundancy to the electrical network. During the summer of 2006, a heat wave disabled the network feeding the airport's Central Electrical Substation (CES). The interconnection between the two substations was proven invaluable as it allowed the PANYNJ to feed most of the critical loads of the CES from the West End Electrical Substation (WECS) and ensured uninterrupted airport operation. . The projections for an increase in temperature, along with more frequent and longer duration heat waves, make electric grid reliability and redundancy a major issue for infrastructure in the NYC region.

**Design Criteria for Adaptive Strategies**

The PANYNJ has also implemented "Interim Design Criteria" that will consider the impact of climate change variables during the design for new construction and major rehabilitation projects. These criteria are shown in Figure 5. The interim design criteria will be in effect until further information is available on climate change projections and will be reviewed and/or updated on a two-year basis. PANYNJ is evaluating the impacts of these design criteria on future and current capital investment projects.

| PANYNJ INTERIM DESIGN CRITERIA | | |
|---|---|---|
| **Climate Variable** | **Baseline 1971-2000** | **2080's** |
| **Air Temperature** | 12.8 ºC (annual mean) | 3.33 ˚C increase over baseline |
| **Precipitation** | 118.1 cm (annual mean) | 10% increase over baseline |
| **Sea level rise** | Current Mean High Water (MHW) | 45.7 cm increase over current MHW |
| **Flood Elevation** | 45.7 cm increase over the current FEMA 100-year flood level plus 30.5 cm criteria (Current FEMA 100-year flood level plus 76.2 cm) | |
| Where prohibiting factors preclude the application of these design criteria to all project elements, focus should be centered on critical project elements for which disruption of service would result in significant impacts to facility operations. | | |

**Figure 5**

One of the first projects to consider these criteria involves the replacement of a substation. The engineering design effort for Substation No. 7 of the PATH rail transit system was recently completed and is ready for construction. This substation is part of a larger network of substations that provides power for approximately seven miles of at-grade rail between Penn Station (Newark, NJ) and the Journal Square Station (Jersey City, NJ). The existing substation No. 7 has reached the end of its useful life and is situated at elevation 10 (NAVD88), which is one foot above the 100-yr. flood level. The design for the replacement substation was based on the newly established interim design criteria and therefore specifies a base elevation of 11.5, which will help to ensure protection against future coastal flooding events for this critical infrastructure.

## Conclusion and Lessons Learned

Through involvement in this effort, the PANYNJ has begun evaluating its facilities' emergency plans and is in the process of initiating more detailed studies to identify operational and maintenance strategies that will improve these plans. The PANYNJ has performed flood studies of LaGuardia Airport and portions of the PATH rail transit system and is in the process of initiating a flood study of the Holland Tunnel. Some major capital improvement projects that are currently in the conceptual design phase and will consider the interim design criteria include the LaGuardia Airport Terminal Modernization and PATH projects such as Washington Street Power House, Newport Station Access, Harrison Station Replacement and Grove Street Modernization.

One of the major benefits provided by involvement in this assessment was the lessons learned, which are as follows:

- Consistent and quantitative regional climate projections from a single reliable source are critical to the success of a climate change assessment for a given region. General information on climate change can be obtained from publications by the IPCC, however regional downscaling of the IPCC models should be performed to determine regional effects, as was done for this effort by the NPCC.
- Institutional knowledge of facility performance during major storm events can aid in determining level of vulnerability of a facility.
- The level of accuracy of information for a given facilities' construction type, site topography and layout is important to determine the level of risk. Actual site surveys may save time and money if record drawings are not readily available. Example - FEMA FIRM maps are suitable for identifying a rough level of vulnerability, but may not represent precise topographic information.
- Infrastructure having a useful life of less than 25 years (for example, HVAC systems) may not warrant evaluation for the impacts of climate change variable projections due to the ability to upgrade during replacement.
- Past capital improvement investments that involve engineering design redundancy or security improvements may have the added benefit of providing mitigation against future impacts of climate change.
- Risk analyses should be as quantitative as possible to ensure facilities are evaluated consistently. Regional stakeholders should develop and use these quantitative criteria in a coordinated manner, when possible.
- Relatively simple and cost effective protective measures can be taken to significantly reduce the potential impacts of climate change in the near-term. Examples – sandbags and Jersey Barriers.
- The cost of an adaptation measure should be weighed against the costs incurred by the loss of that facility, repairs, and lost revenue. The safety of patrons using a facility is the paramount consideration in this evaluation.
- Inclusion of mitigation measures against climate change variables in capital project planning over long term time periods can help to distribute and lower mitigation costs.
- The connectivity of transportation infrastructure under operation by various owners requires a coordinated approach to planning and adaptation.
- A comprehensive assessment of an existing transportation network is necessary for identifying the degree of vulnerability and for prioritizing the protection of its' assets.
- While uncertainty exists in the climate data provided by the NPCC, this uncertainty should not prevent evaluating facilities and capital projects.

**Acknowledgements**

The authors would like to thank the New York City Mayor's Office of Long Term Planning and Sustainability for providing the guidance and oversight throughout this assessment effort.

**References**

Rosenzweig, C. & Solecki, W., New York City Panel on Climate Change, "Climate Change Adaptation in New York City: Building a Risk Management Response", Annals of the New York Academy of Sciences, Vol. 1196, Wiley-Blackwell, May 2010

# Vegetated Roof Installations at O'Hare and Midway International Airports

Gene Peters[1]

[1]Ricondo & Associates, on behalf of the Chicago Department of Aviation (CDA), November 2010. 10510 West Zemke Road, Chicago, Illinois 60666; PH (773) 462-7333; FAX (773) 462-8554; email: epeters@ricondo.com.

## ABSTRACT

The *Sustainable Airport Manual* (SAM) developed by the Chicago Department of Aviation (CDA) established contractual requirements to consider sustainability in the design and construction of every project at O'Hare and Midway International Airports. SAM requires that vegetated roofs be installed on airport facilities wherever possible to reduce the heat island effect, conserve energy, and reduce storm water runoff.

The CDA has installed 232,534 $ft^2$ of vegetated roofs and is installing 126,456 $ft^2$ more (as of November 2010). Installations include the first on-airport Federal Aviation Administration facility in the nation to feature a vegetated roof and the construction of nearly 175,000 $ft^2$ of vegetated roof on a major cargo building, one of the 15 largest vegetated roof areas in the U.S.

The CDA's installations demonstrate that installing vegetated roofs at airports is practical and cost effective. Operational and maintenance cost savings are expected from increased roof life span, energy use reduction, and storm water quantity/quality management. Additional benefits achieved include a reduction in noise and heat islands, air quality protection, and enhanced aesthetics.

## I. VEGETATED ROOF INSTALLATIONS – AN INTEGRAL PART OF THE CITY OF CHICAGO'S SUSTAINABLE INITIATIVES AND SUSTAINABLE AIRPORT MANUAL

The Chicago Department of Aviation (CDA), which manages O'Hare International Airport (ORD) and Midway International Airport (MDW), introduced the Sustainable Design Manual (SDM) in 2003 under the direction of Mayor Richard M. Daley. The SDM was created to ensure that sustainable initiatives were implemented during the build out and modernization of ORD. Initiated in 2001, the $6.6 billion O'Hare Modernization Program (OMP) includes extensions of existing runways, development of a new runway, and associated airfield components, terminal facilities, and landside improvements.

Through the SDM, the CDA developed contractual requirements for the consideration of sustainability in the design and construction of airport projects, including a detailed tracking and rating system. In 2009, with input from over 200 contributors including airport operators, consultants, research institutions, governmental agencies, and aviation organizations, the SDM was updated to form the Sustainable Airport Manual (SAM).

A unique component of the SAM is the requirement that vegetated roofs be installed on airport facilities wherever possible to reduce the heat island effect, conserve energy, and reduce storm water runoff. Vegetated roofs are a potential strategy to obtain the SAM credits and corresponding LEED® 2009 credits for New Construction and Major Renovations listed below in **Table 1**.

**Table 1. Applicable Sustainable Airport Manual and LEED® 2009 Credits**

| | Applicable Credit |
|---|---|
| SAM Credit: | Sustainable Sites, Section 2.5.1 Stormwater Design, Quantity Control |
| LEED® Credit: | Sustainable Sites Credit 6.1: Stormwater Design–Quantity Control |
| LEED® Credit: | Sustainable Sites Credit 5.2: Site Development–Maximize Open Space |
| SAM Credit: | Sustainable Sites, Section 2.5.2 Stormwater Design, Quality Control |
| LEED® Credit: | Sustainable Sites Credit 6.2: Stormwater Design–Quality Control |
| SAM Credit: | Sustainable Sites, Section 2.6.2 Landscape & Exterior Design to Reduce Heat Islands, Roof |
| LEED® Credit: | Sustainable Sites Credit 7.2: Heat Island Effect–Roof |
| SAM Credit: | Energy & Atmosphere, Section 4.4 Optimize Energy Performance |
| LEED® Credit: | Energy & Atmosphere Credit 1: Optimize Energy Performance |

## II. VEGETATED ROOF STRUCTURE AND COMPONENTS

The CDA recognized a need for low maintenance, drought-resistant, and non-wildlife attracting plant species that are acceptable for an airport environment. The selection process included the creation of Specification 02905, *Sustainable Airport Landscaping*, to incorporate vegetated roof considerations. The varieties of sedum species selected are listed below in **Table 2**.

**Table 2. Typical Plant Species Installed on Vegetated Roofs by the Chicago Department of Aviation**

| Genus | Species | Common Name | Flower | Height (cm) |
|---|---|---|---|---|
| Sedum | album (brevifolium) | White Stonecrop | White | 5-10 |
| Sedum | sexangulare | Tasteless Stonecrop | Yellow | 5-10 |
| Sedum | spurium | Dragon's Blood Stonecrop | Pink | 10-15 |
| Sedum | relfexum (altissimum) | Jenny's Stonecrop or Blue Spruce | Yellow | 10-15 |
| Sedum | floriferum | Weihenstephaner Gold or Bailey's Gold | Yellow | 10-15 |
| Sedum | kamtschaticum | Orange Stonecrop | Yellow | 10-15 |
| Sedum | acre (Aureum) | Goldmoss Stonecrop | Yellow | 5-10 |

These sedum species were chosen for their tolerance to drought, lack of food production (berries or seeds), and deficiency of habitat, thereby reducing their attractiveness to wildlife. All species are perennials naturalized to North America and can be found as ground cover in many states.

Extensive vegetated roof systems were installed by the CDA to reduce irrigation needs and the load added to the building. Extensive systems have a profile less than 6 inches and a fully saturated weight ranging from 10-50 pounds per square foot. In contrast, an "intensive" vegetated roof system requires at least one foot of soil depth and contains elaborate irrigation and drainage systems that add considerable load to a structure. Installation of an extensive system typically consists of a single ply roofing membrane. The systems installed at ORD are fully adhered, meaning the delivered components were assembled directly on the roof. For example, **Figure 1** illustrates some of the structural components of the extensive vegetated roof installed on the ORD FedEx Cargo Sort building.

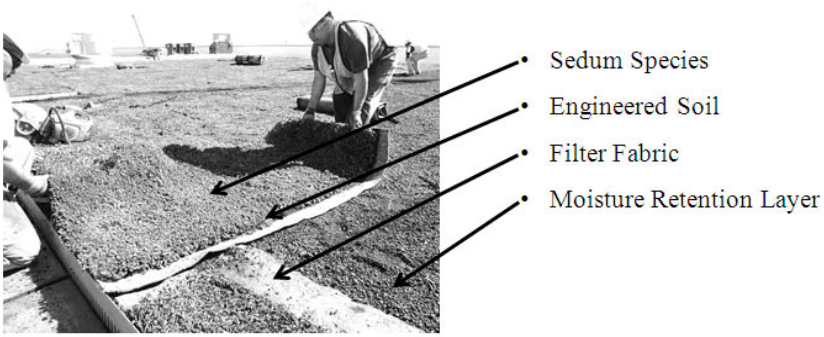



**Figure 1: Installation of the ORD FedEx Cargo Sort Building Vegetated Roof**

Vegetated roof installation is typically bid by the CDA to require the installer to provide care and maintenance, including periodic irrigation, for up to two years; this ensures that mature vegetation is established and covers approximately 90 percent of the available surface. Although species are drought-tolerant, irrigation is necessary for plant establishment due to the heavy wind updrafts and direct exposure to sunlight which increases the loss of moisture from the growing media. The typical two year maintenance period helps prevent soil erosion and reduces the likelihood that plants are uprooted during storm events or high winds.

To reduce the potential of Foreign Object Debris (FOD) from vegetated roofs, the CDA and its tenants designed vegetated roofs in compliance with the wind speed requirements identified in Factory Mutual (FM) Global Property Loss Prevention Data Sheet 1-35: "Green Roof Systems" (Factory Mutual 2007).

The initial cost of a vegetated roof system is typically \$9 per square foot ($ft^2$), or about twice the initial cost of a conventional roof. The increase in cost accounts for structural support and the purchase and installation of the vegetated components. Typical vegetated roof plantings cost approximately \$0.40 per $ft^2$ (Lazarus 2007). A conservative estimate of the average annual maintenance cost is \$1 per $ft^2$, which includes periodic monitoring and plant upkeep. Installers typically provide a two year warranty for the installation and components. To ensure proper installation, the roof may be intentionally flooded to detect leaks or a leak detection system may be installed.

Where there is only partial roof coverage by a vegetated system, the CDA uses high reflectance roofing materials for the remaining roof area where possible. For example, 47 percent of the ORD United Airlines cargo building will be vegetated and the remaining area will contain a high Solar Reflectance Index (SRI) roof surface.

## III. VEGETATED ROOF INSTALLATIONS

With over 500,000 $ft^2$ of vegetated roofs (as of June 1, 2010), Chicago is the green roof capital of the U.S. (The Associated Press 2010). The CDA has installed 232,395 $ft^2$ of vegetated roofs at ORD and MDW and plans to install 126,456 $ft^2$ more (as of August 2010). The CDA has taken a "lead-by-example" approach, installing vegetated roofs atop airport-owned facilities and encouraging tenants to follow suit. Vegetated roofs have been installed on a guard post canopy, a lighting control system facility, and the first vegetated roof ever has been installed on a FAA airport traffic control tower administration facility. Additionally, the FedEx cargo building constructed at ORD in 2010 features nearly 175,000 $ft^2$ of vegetated roof coverage, one of the largest green roof areas in the U.S. **Table 3** lists the size of existing and proposed vegetated roof installations at ORD and MDW and their respective world ranking.

**Table 3. Vegetated Roofs Installations at O'Hare and Midway International Airports**

| Location of Existing Vegetated Roofs | Vegetated Area ($ft^2$) | Year Completed |
|---|---|---|
| ORD Aircraft Rescue and Firefighting Facility #3 | 3,440 | 2006 |
| ORD Guard Post #1 Canopy | 6,500 | 2007 |
| ORD South Airfield Lighting Control Vault | 14,200 | 2008 |
| ORD North Airport Traffic Control Tower Base Building | 8,910 | 2008 |
| ORD North Airport Traffic Control Tower Electrical Building | 917 | 2008 |
| ORD Booster Pump Station (Building 815) | 1,278 | 2008 |
| ORD FedEx Vehicle Maintenance Building | 3,170 | 2009 |
| ORD FedEx World Services Center | 10,024 | 2010 |
| ORD FedEx Sort Building | 174,442 | 2010 |
| ORD Enterprise Rental Car Maintenance Building | 3,627 | 2010 |
| ORD Enterprise Rental Car Customer Service Center | 2,847 | 2010 |
| MDW Economy Elevated Parking Structure Buildings and Electrical Vault (Combined) | 3,179 | 2008 |
| *SUBTOTAL (Existing)* | *232,534* | |
| Vegetated Roofs under Construction/Proposed | | |
| ORD United Airlines Cargo Building | 108,816 | |
| MDW Consolidated Rental Car Facility | 17,640 | |
| **TOTAL** (Existing + Proposed) | **358,990** | |

The sections below describe the completed vegetated roof installations at ORD and MDW. See **Figure 2** for photos of 3 of the 12 existing installations at ORD and MDW.

**Figure 2: ORD North Airport Traffic Control Tower, ORD Aircraft Rescue and Firefighting Facility #3, and the MDW Electrical Vault Vegetated Roofs (pictured left to right)**

**ORD Aircraft Rescue and Firefighting (ARFF) Facility #3**

A 3,440 $ft^2$ extensive vegetated roof was installed on ARFF Facility #3 in 2006 to help improve air quality, conserve energy, reduce storm water runoff, reduce the urban heat island effect, and serve as an insulated roof and a radiant barrier. The installation of the tray-based system marked the first vegetated roof installed by the CDA. The planting, which includes sedum and other groundcover, is visible by Airport Transit System (ATS) riders who can see the roof as they travel to and from the terminal and remote parking lots. The vegetated roof on the firehouse also features 10 solar thermal collectors that capture the sun's energy; this energy is used to heat water that is circulated in metal tubes to provide hot water for the firefighters who work in the building.

**ORD Guard Post #1 Canopy**

The Guard Post #1 Canopy, completed in 2007, is covered by a 6,500 $ft^2$ vegetated roof. The roof design reduces the urban heat island effect, reduces the peak flow rate of storm water, and allows for the retention of storm water runoff that would otherwise flow to the storm water system. The vegetated roof on the guard post is a visible display of the City of Chicago's commitment to the environment at the northwest airport entrance.

**ORD South Airfield Lighting Control Vault (SALCV) and Duct Bank Building**

Completed in 2008, the 14,200 $ft^2$ vegetated roof installed on the South Airfield Lighting Control Vault (SALCV) and Duct Bank Building (ORD Building 607) consists of an extensive system with 6 inches of growing medium. The SALCV provides and regulates electrical power for runway and taxiway lighting. Workers monitoring the facility have reported cooler temperatures inside the building compared to conventional facilities of the same size. The protected membrane roof and waterproofing assembly consists of a rubberized asphalt membrane, polyester reinforcement, drainage/retention mat, root barrier, and extruded insulation (GreenRoofs.com 2009).

**ORD North Airport Traffic Control Tower**

The FAA North Airport Traffic Control Tower (NATCT), which opened in November 2008 to accommodate Runway 9L-27R, features nearly 10,000 $ft^2$ of vegetated roof space on top of the administration (8,910 $ft^2$) and electrical (917 $ft^2$) buildings, making it the first on-airport FAA facility in the nation to feature a vegetated roof. Critical airport traffic control equipment is housed underneath the vegetated roof in the administration and electrical buildings. The installations demonstrate the CDA's confidence in vegetated roof technology and willingness to consider vegetated roofs on all airport facilities.

The vegetated roofs and energy efficient electrical devices are estimated to help decrease energy use by approximately 36 percent compared to baseline consumption levels (Rink 2009). As a result of the sustainable initiatives incorporated into the project, the tower earned a four Green Airplane Certification based on the SDM rating system and is the first tower in the nation to earn United States Green Building Council (USGBC) LEED® Silver certification.

### ORD Booster Pump Station

In 2010, a pump station (ORD Building 815) was constructed southeast of the Aviation Administration Building (Building 4) along Bessie Coleman Drive. As part of the construction, a 1,278 $ft^2$ vegetated roof system was installed that covers about 75 percent of the building's total area (approximately 1,716 $ft^2$).

### ORD FedEx Sort Building, World Services Center, and Vehicle Maintenance Building

The construction of future Runway 10C-28C as part of the OMP required the relocation of the FedEx Metroplex and other existing facilities in 2010. As part of the relocation, a 174,442 $ft^2$ vegetated roof was installed on the relocated FedEx Cargo Sort Building, the largest vegetated roof at any airport in the U.S. and the largest vegetated roof on a free standing building in the greater Chicago area. The Sort Building is located on ORD property adjacent to runways and airplane traffic and is used to sort packages, documents, and non-conveyable items. The roof design and material selection required close consideration to prevent Foreign Object Debris in an active airspace. The installation included 3.9 acres of pre-vegetated mat, more than 2,200 cubic yards of media and 3 miles of edging. It is estimated to have a saturated weight of 4.3 million pounds and retain approximately 2 million gallons of storm water annually (Airport Business.com 2010). In addition, a 3,170 $ft^2$ vegetated roof was installed on the FedEx vehicle maintenance building and a 10,024 $ft^2$ vegetated roof was installed on the World Services Center.

### ORD Enterprise Rental Car Maintenance Building & Customer Service Center

The Enterprise Rental Car facility at ORD opened in the summer of 2010 and features vegetated roofs on the maintenance building (3,627 $ft^2$) and the Customer Service Center (2,847 $ft^2$). The Enterprise facility incorporates several sustainable initiatives such as high efficiency fixtures to reduce water consumption and the use of recycled construction materials. The facility is anticipated to obtain USGBC LEED® Silver certification.

### MDW Economy Elevated Parking Structure Buildings and Electrical Vault

The installation of 3,179 $ft^2$ of vegetated roof at MDW was completed in June 2008. Modular systems were installed on two entrance vestibules to the economy parking garage and the 1,400 $ft^2$ electrical vault building. In these modular systems, the vegetation is fully mature at the time of installation, reducing maintenance costs and the potential for erosion from wind and rain. The 4 inch system installed on the seven-level economy parking garage features an architectural design that matches the MDW terminal. The vegetated components were assembled and planted off-site, hoisted to the roof, and anchored in place.

## IV. VEGETATED ROOF BENEFITS

Vegetated roofs provide a number of benefits to both the individual building owner and the Chicago community. Vegetative roofs reduce storm water runoff quantity and improve storm water quality, reduce the heat island effect, lower operating costs, increase roof lifespan, decrease interior and exterior noise, improve

aesthetics, and reduce air pollution and greenhouse gases. The benefits of installing vegetated roofs at Chicago's airports are described below:

**Reduce the Heat Island Effect**

Conventional roofs absorb solar radiation and re-radiate it as heat, contributing to the heat island effect. A heat island is an urban area where temperatures are significantly hotter than the surrounding countryside because of large impervious surface areas such as asphalt and concrete typically found at an airport.

The urban heat island effect causes overheating of cities during the summer, leading to excessive energy consumption on cooling and air conditioning, and pollution from the generation of electricity. Heat islands also accelerate chemical reactions in the atmosphere, increasing ozone concentrations that are a risk to human health. Installing vegetated roofs atop airport facilities has proved to be an effective method for counteracting the heat island effect at the CDA's airports.

**Lower Operating Costs**

Vegetated roofs improve thermal insulation, helping the airport reduce heating and cooling costs. They may also reduce the amount of roofing insulation used and the size of HVAC equipment. The Federal Energy Management Program reported that the maximum temperature fluctuation in a year on a vegetated roof was 95°F (35°C) while a conventional roof experienced a maximum temperature fluctuation of 212°F (100°C) (Federal Energy Management Program 2004). In the areas immediately under the roof, vegetated roofs reduce operating expenses by 95 percent for cooling and 26 percent for heating (Lazarus 2007).

**Life Cycle**

A vegetated roof can prolong the life of the average roof from 15-20 years to 40-50 years, reducing maintenance and replacement costs. Unlike a vegetated roof, a conventional roofing membrane suffers from exposure to ultraviolet radiation, temperature extremes, and precipitation. Although conventional roofs typically have lower installation costs, the re-roofing costs increase the life cycle cost of the roof.

**Storm Water Retention**

Vegetated roofs increase storm water retention, filtration, and evaporation, and decrease the need to expand or rebuild drainage infrastructure. Water is stored by the substrate where it is taken up by plants and returned to the atmosphere through transpiration and evaporation. This process helps slow the rate of storm water flow, reduces stress on sewer systems and the risk of contamination and flooding, and may reduce storm water management charges. Depending on the plants and depth of growing medium, vegetated roofs typically retain 70 to 90 percent of the precipitation that falls on them during the summer and between 25 to 40 percent in the winter (The Encyclopedia of Earth 2008).

**Reduce Air Emission Impacts**

In addition to absorbing heat, vegetated roof foliage absorbs dust and greenhouse gases and filters airborne particles. Plants improve air quality by using carbon

dioxide to produce oxygen through photosynthesis. One square-meter of vegetated roof removes approximately 4.4 pounds of airborne particulates per year (Blok 2009).

**Noise Reduction**

Sound waves produced by machinery, traffic, and/or airplanes are absorbed or deflected by the soil, plants, and the trapped layer of air in airport vegetated roofs. A vegetated roof with a 4.7 inch substrate layer can reduce sound by 40 decibels; a 7.9 inch substrate layer can reduce sound by 46-50 decibels (City of Phoenix 2006).

**Aesthetics and Visibility**

Millions of visitors observe ORD's vegetated roofs while traveling on the ATS, driving to the airport along Mount Prospect Road, or when flying overhead, making the CDA's commitment to the environment very visible. The sedum plant species used by the CDA are low-growing plants that have white or yellow flowers and green, light green, or maroon foliage and are aesthetically pleasing yet are tolerant to drought and do not attract wildlife.

## V. CONCLUSIONS AND RECOMMENDATIONS

The CDA's commitment to sustainability has encouraged other airports, industry associations, and the federal government to incorporate sustainable initiatives into their projects. The successful vegetated roof installations by the CDA demonstrate that installing vegetated roofs at airports is practical and cost effective.

Installing vegetated roofs on existing and future roof areas provides potential economic and operational benefits to the airport and environmental benefits to the surrounding community. Operational and maintenance cost savings are expected from increased roof life span, energy use reduction, and storm water quantity/quality management. The payback period of a vegetated roof installation varies on the size, available government subsidies, and the cost of electricity and municipal water. Typically, a vegetated roof pays for itself within 25-30 years. Additional vegetated roof benefits such as reducing the heat island effect, noise reduction, air quality protection, and aesthetics and marketability cannot be quantitatively valued.

Airports interested in vegetated roof technologies should pursue the following steps:

- Develop a sustainable design/construction policy for vegetated roof installation.
- Develop guidelines for sustainable design and construction (checklist).
- Understand mission-critical considerations unique to an airport environment, including concerns over FOD generation and wildlife attractants.
- Develop a detailed cost-benefit evaluation of recommended installations.
- Coordinate with the owner/end user to address design, construction, operation and maintenance issues.
- Develop a material and installation specification for vegetated roofs. Include drought tolerant, low maintenance, and non-wildlife attracting plant species.
- Install meters/monitoring devices to analyze actual gained economic, operational, and environmental benefits from vegetated roof installations.
- Develop roof vegetation maintenance specifications; include a 2-year maintenance warranty that establishes maintenance procedures such as watering, weeding, fertilizing, and replacement (as needed).

- Quantify cost savings in energy use reduction for heating and cooling and storm water management; recognize non-quantifiable benefits in reduced heat-island effect, air quality improvements, noise reduction, and aesthetics.
- Promote achievements using kiosks, displays, brochures, and/or presentations.

## V. REFERENCES

Airport Business.com "Collaboration Creates Green Roof at O'Hare," May 26, 2010, www.airportbusiness.com/web/online/Top-News-Headlines/Collaboration-Creates-Green-Roof-at-OHare/1$36832 (accessed July 1, 2010).

The Associated Press. "DC 2nd on Green Roof List, Baltimore 4th." The Washington Examiner, June 1, 2010, www.washingtonexaminer.com/ (accessed Oct. 21, 2010).

Blok, Kayla et al. "Greening New Heights – A Feasibility Analysis of a Green Roof on the Life Sciences Center," May 2009, http://environmental.science.dal.ca/Files/ENVS_3502_projects_2/ green_roof_final_report.pdf (accessed Oct. 21, 2010).

The City of Phoenix, Planning Commission. "Green Roofs White Paper," January 3, 2006, http://phoenix.gov/urbanformproject/wp11.pdf (accessed July 2, 2010).

The Encyclopedia of Earth. "Green Roofs," updated August 29, 2008, www.eoearth.org/article/Green_roofs (accessed July 6, 2010).

Factory Mutual Insurance Company, FM Global Property Loss Prevention Data Sheets. "1-35: Green Roof Systems," January 2007.

Federal Energy Management Program, U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (based on analysis by A. Dürr, 1995). "Green Roofs," Federal Technology Alert, DOE/EE-0299, pg. 8, August 2004.

GreenRoofs.com, The International Greenroof and Greenwall Project Database. "O'Hare International Airport Building 607 New South Vault and New Duct Bank Buildings," May 2009, www.greenroofs.com/projects/pview.php?id=913 (accessed July 6, 2010).

Lazarus, Mary Ann, Hellmuth, Obata and Kassabaum (HOK). "O'Hare Modernization Program Green Roofs Presentation," January 16, 2007.

Rink, Grace. "Sustainable Construction Practices, OMP Case Studies: North Airport Traffic Control Tower," Airports Going Green Conference Presentation, Aug. 5, 2009.

# Integrating Sustainability into Highway Projects: Sustainability Indicators and Assessment Tool for Michigan Roads

Krishna P. Dhakal,[1] and Jun S. Oh[2]

[1] Graduate Student, Department of Civil and Construction Engineering, Western Michigan University, 1903 W. Michigan Ave., Kalamazoo, MI 49008-5316, PH (269) 276-3216; FAX (269) 276-3211; e-mail: krishnaprasad.dhaka@wmich.edu
[2] Associate Professor, Department of Civil and Construction Engineering, Western Michigan University, 1903 W. Michigan Ave., Kalamazoo, MI 49008-5316, PH (269) 276-3216; FAX (269) 276-3211; e-mail: jun.oh@wmich.edu

## ABSTRACT

Studies have shown that nation's conventional highways are unsustainable in many aspects. Highways are facing severe challenges, such as deteriorating condition, congestion, energy supply and shortfall of funding for maintenance and capacity expansion to meet increasing demand. At the same time, they consume huge amount of exhaustible natural minerals and energy, and generate waste and emission which deteriorate global environment contributing to climate changes and reduction of the earth's sustaining capacity. Therefore, sustainable highway construction, maintenance and operations have become a prime concern these days. The study examines various research findings and best management practices in road construction industry including some established rating tools for building and some newly proposed tools for road. Finally, it identifies Indicators for sustainable roads for Michigan and develops a sustainability assessment tool to evaluate those indicators that can be expressed as a single comparable index as an end product.

## INTRODUCTION

Highway system is an inevitable component for present mobility and economic development, but the existing highway system causes many issues owing to their environmental impacts, financial inefficiency, natural resource exploitation, and adverse social repercussions. Many innovations, management practices and statutory provisions adopted during past decades have made considerable achievements; however, the improvements are insufficient and the road infrastructures continue facing unabated threats like congestion, deteriorating conditions, green house gas emissions, pollution, depletion of exhaustible resources and financial scarcity. Therefore, exploring a comprehensive, integrated and radical approach for incorporating sustainability measures into highway has been a major challenge for today's transportation professionals, engineers and researchers.

Previous studies reveal the multiplicity of aspects that render the existing highway systems unsustainable. The transportation sector is responsible for 22% of global

energy consumption and 25% of fossil fuel burning producing 30% of global air pollution and greenhouse gases (AASHTO, 2010). Highway emission causes increased mortality from respiratory diseases (Dora and Philips, 2002) and traffic noise threatens community quietness. Thousands of people and animals are either injured or killed, in road accidents in a year. Social disparity regarding investment, natural resource exploitation, and in sharing benefits and detriments has been an issue of socio-political conflict. In addition, the present road system suffers from economic loss in terms of lost hours as a consequence of congestion. For example, congestion in US roads caused 4.2 billion hours of delay resulting in 2.9 billion gallons of wasted fuel and nearly $80 billion of congestion costs in 2007 (USDOT 2008). To regain the functional ability of the system for nation's economy, US roads require a substantial investment. Paradoxically, many state Departments of Transportation (DOTs), including Michigan, are curtailing annual programs due to the funding crisis. For achieving sustainable highways, these multiple issues should be addressed in an integrated fashion. Greenroads (Muench et al., 2010) is one such approach which provides ample knowledgebase on road sustainability issues and develops a tool for rating projects. Greenroads is basically a motivation tool, but lacks in measuring and tracking the level of sustainability and its progress.

This paper revisits the multidimensional aspects of road sustainability, defines themes for sustainable highways and explores measures to incorporate those themes into Michigan highways. Finally the paper develops an integrated rating tool by assigning weights to each indicator in proportion to their impact on various sustainability themes. The tool can be utilized for implementing sustainability measures, measuring sustainability performance level, and tracking progress.

## SUSTAINABLE ROADS: THEMES AND DEFINITION

Highway sustainability is the adoption of concept of sustainable development in the context of highways. The concept of sustainable development concentrates on two basic themes- human needs and resource limits, and advocates protecting nature's capacity to sustain life. In the widely accepted definition, Brundtland et al. (1987) advocates preservation of resources for future generations and defines sustainable development as "development, which meets the need of the present without compromising the ability of the future generation to meet its own needs."

### Conventional Three-Themes Model

The conventional definitions of sustainable development emphasize on economy, society and environment as three themes of sustainability. Transport Canada (1997, 2000) states "The goal of sustainable transportation is to ensure that environmental, social and economic considerations are factored into decisions affecting transportation activity." According to Transportation Research Board (1997), "Specifically, it is about how environmental, economic, and social systems interact to their mutual advantage or disadvantage at various space based scales of operation."

**Development of Five-Themes Model**

Since the construction, maintenance and operation of highways require huge financial and natural resources, they play vital role in sustainability performance. The three-theme model includes financial component under economic theme and natural resource component under environmental theme leading to insufficient emphasis on financial and natural resources. On the basis of following reasons, we purpose to separate these components from their respective parent themes and develop a five-theme model as shown in Figure 1.

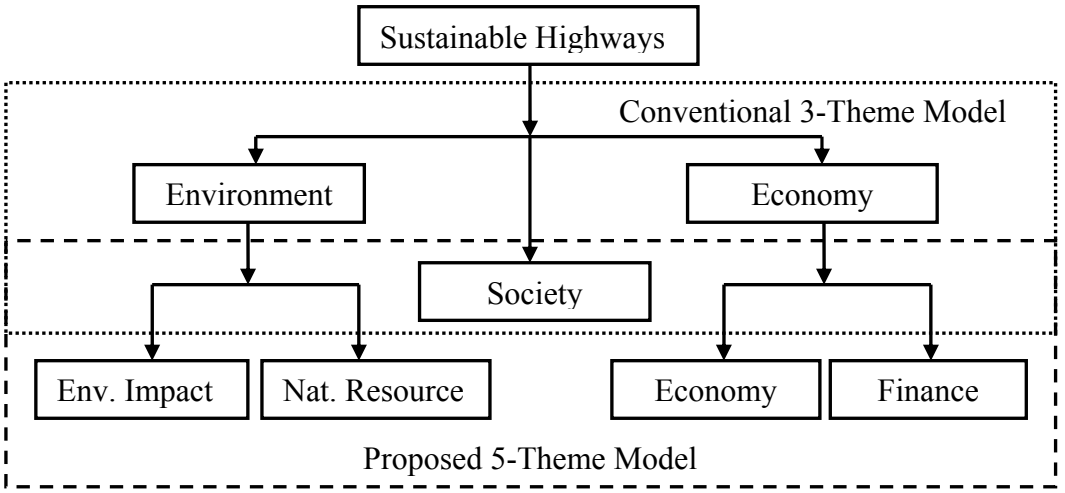



**Figure 1 Development of 5- Theme Model**

Reasons for separating the themes:

i. Economy and Finance
   a. Finance is input for facility construction and maintenance whereas economy refers to the desired output of the facility.
   b. Finance is project specific whereas the economy implicates society.
   c. The 3-theme model overlooks the impact of financial scarcity on sustainability giving emphasis on economic productivity and mobility.

ii. Environmental Impact and Natural resources
   a. Natural resource (Material) is input for the road infrastructure whereas, the environmental impact that includes emissions, waste, pollution and ecological impact is the output of the system.
   b. The separation as a theme helps gain more academic and political emphasis on reducing the use of natural minerals.

Based on the five-theme model, we have developed a more comprehensive definition of sustainable roads as: *The road that protects and enhances physical and biotic natural environment by adapting itself to natural law, makes prudent utilization of natural materials, fulfils optimum social aspirations, supports vibrant economic activities, and generates sufficient financial resource to sustain, improve and update itself.*

## MEASURING AND IMPLEMENTING SUSTAINABILITY

Sustainability is usually evaluated using indicators which are useful for identifying trends, predicting problems, setting targets, evaluating solutions, and measuring progress (Litman, 2010). Since many indicators are required to cover the overall aspect, an index is developed as a single quantitative value. An "index" is a quantitative aggregation of many indicators and can provide a simplified, coherent, multidimensional view of a system (Mayer, 2008). This paper determines indicators, assigns weights and develops index to measure the level of sustainability.

### Criteria for Indicator Selection

Though highway sustainability issues are largely universal, the road exhibits various location specific characteristics. Therefore, the context specific local parameters such as climate, public priorities and resource availability etc. should be included along with universal parameters as basic criteria for determining indicators. This paper considers the various aspects while defining indicators including: sustainability themes and objectives, measurability, data collectability, climatic conditions, socio-economic structure, resource availability, biodiversity, public policy and state priorities.

### Determination of Scale and Relative Weights

Determining the impact level of each indicator and assigning legitimate values for quantification are complex and may not be very objective. Sustainability assessment is not a pure, neutral and objective science, but it rather reflects these implicit judgments (Becker, 1997). For weights of Michigan highways, we assigned a weight of 100 for an ideally sustainable highway. The weight of 100 was then distributed among the indicator categories, which were further distributed among each of the constituent indicators based on potential impact on sustainability. Greenroads (Muench et al., 2010) does not set any standard value for perfectly sustainable road, which makes the tool incapable of determining the actual level of sustainability. For example, how sustainable a segment of roadis unknown. Furthermore, 11 items are required , which means that the tool is incapable of tracking progress achieved by adopting federal and state regulatory actions, most of which have been included under the project requirement. When comparing the sustainability performance of roads built in 1970s with the road built recently, the Greenroads rating tool cannot be used because the road built in 1970s is likely to be disqualified for the evaluation because of the project requirements.

The absence of project requirements in the proposed tool makes the tool capable of rating any highway and hence progress tracking can be done. In addition, sustainability levels of each project can be measured by expressing as a percentage of standard value of 100. For example, a highway with score 60 is 60% sustainable.

## INDICATOR CATAGORIES AND IDICATORS FOR MICHIGAN ROADS

Although sustainability is a global issue, studies reveal that sustainability attributes for roads are largely dependent on regional factors like resource availability, climatic conditions, economic condition, public policies, etc. To determine categories and constituent indicators, we investigated Michigan specific issues including existing road conditions, financial crisis, Michigan's biodiversity, lakes, wetland, climatic conditions, Michigan's economic pattern changes, etc. The weight of each category reflects the level of impact on sustainability while the weight of each indicator reflects the potential of effectiveness in mitigating impacts.

### Planning

Planning is the foundation for every aspect of highway throughout its life cycle from concept generation to facility demolition, and it influences each of the five themes of sustainability. The effectiveness of highways is interdependent with socioeconomic activities and is shaped by the region's land use pattern. In Michigan, out of 122,000 miles of the total road, 9,711 miles are under State; 89,499 miles are under 83 counties; 2,102 miles are under federal government; and 20,667 miles are under 535 incorporated cities and villages (PSCI, 2003). There are more than 1,800 units of Michigan's local government (Michigan Land use Leadership Council, 2003). In addition, some roads share jurisdictions with other states and some with Canada. This scenario demands an effective coordination among various authorities for achieving sustainable highway system through consistency and coordination among visions, policies and plans of all the concerned authorities. Realizing importance of planning as a proactive role player, this study regarded planning as one of major categories and assigned a weight of 11 for its five indicators as shown in Table 1.

### Public Participation

People are the prime investors and ultimate users of the road facility, and face the adverse impacts that result from facility construction and operation. The planners and designers cannot design solution without knowing the people's needs and aspiration. Therefore, public participation is an essential factor in planning, investment decisions, alignment selection, and design of the facility. We adopted public participation as one of categories and assigned a weight of 10 for its seven indicators.

### Mobility and Safety

Increased mobility with desired safety is the prime objective for any highway. Study on Michigan roads reveals multitude of mobility and safety problems with economic, social and environmental implications. The percentage of congested urban highways in Michigan increased from 23% in 2000 to 39% in 2005, and congestion of urban roads was ranked as the eighth worst among fifty states (ASCEMI, 2009). In 2006, a traffic crash occurred every 5 seconds, someone involved in a traffic related injury every 12 seconds, and someone died by a traffic crash every 12 minutes (USDOT, 2008) in the US. Improving safety in highways throughout the nation has become a

primary concern. Therefore, we adopted mobility and safety as one of categories with a total weight of 16 for seven constituent indicators.

**Environmental Health**

Michigan possesses a diverse biodiversity. It is a wonder water land surrounded by fresh water of the Great Lakes and is enriched by some 36,000 miles of stream, more than 11,000 lakes and ponds (DNRE), and thousands of wetlands. These natural features provide food storage and habitat for wildlife, and allow summer discharge contributing to retaining water table. Therefore, their protection is necessary. Since highway construction threats these properties through filling and alteration of hydrology, the mitigation measures are to be applied for road sustainability. We adopted environmental health as one of major categories with highest priority and assigned a total score of 20 for its 13 indicators in order to protect and enhance biodiversity by reducing air and water pollution.

**Access and Equity**

Highways are constructed with the expense of common material and public dollars, and impact communities and the common environment. Therefore, the road facility should be used equitably by allowing equitable accessibility to all people. Conventionally, the high income group use the major portion of facility while the poor have to face the environmental impact in the form of disease because of their incapability of investment in healthcares. Similarly, the physically impaired and aging people cannot use the facility as the young and healthy people. We adopted access and equity as an important component with the intent of providing access to every aspect of highways from job opportunities in highway related activities to facility use, assigning a weight of 12, distributed proportionately over its six indicators.

**Natural Resource**

Road construction and maintenance consume a vast amount of natural resources, such as land, water, Boulders, Gravel, Asphalt, steel, etc. In the US, road construction consumes 530-793 million metric tons of crushed rock, gravel and sand each year for aggregate (Industrial Resource Council). On the other hand, the equipments required for material extraction, production and transportation as well as facility construction consumes huge amount of non-renewable energy. Therefore, for sustainable highways, we have to control resource use by minimizing the use of materials, and reusing and recycling the materials. Realizing this, we adopted minimization of natural resource use as one of categories and assigned a weight of 12 for 6 constituent indicators.

**Financial Competency**

Michigan is facing historical financial scarcity to maintain aging road infrastructures. In Michigan, 38 percent of roads are in poor condition, 25 percent of bridges are structurally deficient or functionally obsolete, storm water management system is poor, and water quality of discharge is low (ASCI MI, 2009). Due to shortfall of funding Michigan roads are deteriorating faster than they can be repaired or replaced (Michigan Transportation Asset Management Council, 2008, 2007). Funding crisis causes poor maintenance and results in low economic productivity, which consequently leads again to financial scarcity and repeats the same vicious cycle. It is impossible to maintain highway facilities without sustainable funding sources. Therefore, we adopted financial health as one of sustainability components and assigned a weight of 10 for 7 constituent indicators.

**Innovation and Quality**

Applications of established standards as well as continuous research and innovation are necessary to increase the product quality and performance. US envisions an innovative transportation system that incorporates efficient, integrated, cost-effective, and sustainable solutions leading to dramatic improvements (USDOT, 2008). Michigan needs innovative solutions to its transportation problems and the state should partner and/or support research to encourage this innovation (MDOT, 2006). For promoting innovations and quality we adopted innovations and quality as one of categories with a weight of 11.

**Sustainability Assessment Index for Michigan Highways**

Based on extensive exploration, this study developed a list of indices for assessing sustainability of the Michigan highways. As shown in Table 1 Sustainability Assessment Index for Michigan
, the list includes 57 indices in 8 different categories. Each item is also representing one or more of five sustainability aspects.

**Table 1 Sustainability Assessment Index for Michigan**

| Category | Indicators | Themes | Score |
|---|---|---|---|
| Planning | Consistency with State/ Regional Plans | E,S,En,F,R | 2 |
| (11 points) | Integration with land use Plan | E,S,En,F,N | 3 |
| | Environmental Management Plan | E,S,En | 2 |
| | Asset Management/Site Maintenance Plan | E,S,En, | 2 |
| | Access Management Plan | E,S,En | 2 |
| Participation | Public outreach policy formulation | S | 1 |
| (10 points) | Public involvement in need assessment | E,S,En | 2 |
| | Public Involvement in Planning Decision | E,S,En | 2 |
| | Involvement in Financial Planning | S,F | 1 |
| | Participation in alignment selection | E,S,En,R | 2 |
| | Involvement in design decision | S,En | 1 |
| | Public involvement in facility upkeep | S,En,F | 1 |
| Mobility/Safety | Context sensitive & consistent design | E,S,En | 1 |

| Category | Indicators | Themes | Score |
|---|---|---|---|
| (14 points) | Optimized signal control | E,S,En | 2 |
| | Optimized conflict points at intersection | E,S,En | 2 |
| | Delay reduction from base year average | E,S,En,R | 3 |
| | Mobility at construction zone | E,S,En | 2 |
| | Safety audit and installation of safety | E,S | 2 |
| | Safety measures for construction workers | S,En | 2 |
| En. Impact | Buffer between water body and road edge | En | 2 |
| (20 points) | Pre-construction hydrological conditions | En,S | 3 |
| | Minimize disturbance in water bodies | E,En | 1 |
| | Preserve habitat and ecological | E,En | 3 |
| | Mitigate light pollution | En | 1 |
| | Use low emission material | S,En | 1 |
| | Use low emission equipment for | S,En | 1 |
| | Dust control during construction | S,En | 1 |
| | Noise control measures | S,En | 1 |
| | Low vibrating equipment | S,En | 1 |
| | Construction waste management plan | En, R | 2 |
| | Low impact snow removal mater./process | En,F | 1 |
| | Water efficient/context sensitive landscape | S,En,R | 2 |
| Access& Equity | Access for all to work | E,S | 1 |
| (12 points) | Access to information | S | 2 |
| | Transit and HOV Access for all | S,E | 3 |
| | Pedestrian path & dedicated access at | S | 2 |
| | Bicycle path & dedicated access at | S | 2 |
| | Facility to disabled and aging people | S | 2 |
| Natural Res | Competent material selection by LCA | R | 3 |
| (12 points) | Reduced Construction footprint | R,En | 2 |
| | Earthwork Balance | R | 2 |
| | Reduced water use | R | 1 |
| | Reused Material | R,En | 2 |
| | Recycled Material | R,En | 2 |
| Financial Comp | Sustainable capital investment | F | 3 |
| (10 points) | Funding plan for Maintenance | F | 2 |
| | Funding for administrative expense | F | 1 |
| | Funding plan for research & development | F,S,E | 1 |
| | Cost effective alignment | F | 1 |
| | Cost effective design alternative | F | 1 |
| | Cost effective material & process | F | 1 |
| Inno. & Quality | Quality management plan implementation | E,En,S,R,F | 3 |
| (11 points) | Application of Lean technique | En,R, | 2 |
| | Innovative project delivery system | En,S,R | 1 |
| | Innovative high performance pavement | E,S,En,R,F | 2 |
| | Innovative rapid construction technique | E,S,En | 1 |
| | Innovative ITS application | E,S,En,R | 2 |
| **Total** | | | **100** |

Note: E= Economy, S=Society, En= Environment, F= Finance, R=Natural resourse use, Res.=Resource, Inno.= Innovation; Comp=competency; presence of the acronym in the Theme field indicates the influence of the indicator on that theme.

## CONCLUSION

The paper presented a list of sustainability indices for Michigan highways, which can be used as an innovative tool to help incorporate sustainability into various aspects of Michigan roads. A total of 57 indicators grouped under eight categories were listed in order to measure sustainability during planning, design, construction, maintenance and improvement. The elements were weighted based on their relative impacts using the findings from literature reviews. A scale of 100 levels was used to evaluate each element's impact on sustainability and the weights were aggregated to a single measure expressed in a scale of 100 levels. The tool, in its most basic form, can be used as an integrated checklist of the sustainability components. It can be used not only as a decision support system for decision makers, but also as an implementing tool for construction professionals and as an evaluation tool for an operating agency. The list includes general elements for all agencies and functional class of the road. Since the impact of the different elements may vary depending on the functional classification of the road, it is recommended that separate tool be developed for freeways, and other classes of the road. Further research is required to find more realistic impacts of each sub category on each of the five sustainability themes.

## REFERENCES:

American Association of State Highway and Transportation Officials (AASHTO), "Sustainability," http://environment.transportation.org/ environmental_issues/ sustainability> (Feb 9, 2010)

American Society of Civil Engineers Michigan Section, ASCE MI (2009), "2009 Michigan Infrastructure Report Card," <http://sections.asce.org/michigan/ michiganreportcard/_files/ASCEMI-Infrastructure-Report.pdf> (March 15, 2010)

Becker, B (1997), "Sustainability Assessment: A Review of Values, Concepts, and Methodological Approaches," The Secretariat of the Consultative group on International Agricultural Research, 1997. <http://www.worldbank.org/ html/cgiar/ publications/issues/issues10.pdf.>

Brundtland et al., (1987), "Our Common Future," World Commission on environment and Development, Oxford: Oxford University Press, p.43

Committee for a Study on Transportation and a Sustainable Environment (1997), "Toward A Sustainable Future: Addressing the Long-Term Effects of Motor Vehicle Transportation on Climate and Ecology," National Academy Press, Transportation Research Board, 1997, <www.trb.org.>

Dora C. and M. Phillips (2002), "Transport, Environment and Health," WHO Regional Publications, European Series, No. 89, World Health Organization, http://www.euro.who.int/document/e72015.pdf

Industrial Resource Council, "Promoting Sustainable Use of Industrial Materials", < http://www.industrialresourcescouncil.org/AboutUs/tabid/ 360/Default.aspx> (Oct 20, 2010)

Litman T. (2010), "Well Measured: Developing Indicators for Sustainable and Livable Transportation Planning," Victoria Transport Policy Institute, BC, Canada, 2010. http://www.vtpi.org/wellmeas.pdf.

Mayer A.L. (2010), "Strength and Weakness of Common Sustainability Indices for Multidimensional Systems," Environmental International 34 (2008) 277-291, 2010 www.elsevier.com/envint.

MDOT State Long Range Plan 2005-2030 (2006), "Meeting Summary," Economic Advisory Group & Stakeholder Workshops; Lancing, Detroit, Grand Rapids, & Grayling, June 22-28, 2006

Michigan Department of Natural resources and Environment, DNRE, "Inland Lakes and Streams," http://www.michigan.gov/deq/0,1607,7-135-3313_3681---,00.html (Oct 20, 2010)

Michigan Land Use Leadership Council, "Michigan's Land Michigan's Future," Final Report prepared for Michigan Governor and Michigan Legislature, Aug.15. 2003. http://www.michiganlanduse.org/ MLULC_FINAL_REPORT_0803.pdf.

Michigan Transportation Asset Management Council (2008, 2007), "Driving Our Economy Forward, County Road Association of Michigan" <http://www.micountyroads.org/econ_broch.pdf >

Muench, S.T., Anderson, J.L., Hatfield, J.P., Koester, J.R., & Söderlund, M. et al. (2010), "Greenroads Rating System v1.0", (J.L. Anderson and S.T. Muench,Eds.), Seattle, WA: Uni.of Washington.<http://www.greenroads.us>

Public Sector Consultants Inc. (2003), "Michigan in Brief," <http:// www.michiganinbrief.org/edition 07/chapter 5/Highway.html>, (March 12, 2010)

Transport Canada (1997, 2000), "Sustainable Development Strategy and Towards Sustainable Transportation," www.tc.gc.ca/envaffairs/english/sustain.htm.

US Department of Transportation, Research and Innovative Technology Administration (2008). "Transportation Vision for 2030"

# A Case Study of Transit-Oriented Transportation Development and Land Use in China

Enhui Yang[1], Yanjun Qiu[2], and W. Virgil Ping[3]

[1]Ph.D.-Student, Department of Highway and Railway Engineering, School of Civil Engineering, Southwest Jiaotong University, Chengdu, Sichuan 610031, China; Currently Visiting Scholar, Department of Civil and Environmental Engineering, FAMU-FSU College of Engineering, Florida State University, Tallahassee, FL 32310, USA; PH 1-850-567-2090; FAX 1-850-410-6142; email: yenhui@fsu.edu

[2]Professor, Department of Highway and Railway Engineering, School of Civil Engineering, Southwest Jiaotong University, Chengdu, Sichuan 610031, China; email: Publicqiu@vip.163.com

[3]Corresponding Author, Professor, Department of Civil and Environmental Engineering, FAMU-FSU College of Engineering, Florida State University, 2525 Pottsdamer Street, Tallahassee, FL 32310, USA; PH 1-850-410-6129; FAX 1-850-410-6142; email: ping@eng.fsu.edu

## ABSTRACT

In recent years, with strong economic development and rapid urbanization in China, traffic congestion has become a common phenomenon in urban areas. Focusing on the high-density development of land use in China, the urban population has become highly concentrated and the number of vehicles has also increased exponentially. Traffic congestion problems have become more and more serious in most of the major cities in China. To mitigate the problems, researchers and planners have learned to use a coordinated approach balancing the transportation development and land use. Transit-oriented development (TOD) refers to residential and commercial districts located surrounding a transit station or corridor with high quality service, good walkability, good parking management, and other design features that facilitate transit use and maximize overall accessibility; it is the new fast growing trend in creating vibrant and livable communities. The TOD has become one of the leading urban planning models in China.

In this paper, one of the residential and commercial districts in Chengdu, which is the largest city in the southwest of China, was selected for the model study in coordinating the transportation development and land use. The study was based on the TOD model and the genetic algorithm (GA) approach using the MATLAB software. It was shown that the optimal solution significantly increased the density and consistency of the land use when compared with the current construction plan. The traffic flow on the major road of this community was just around 67% of its predicted traffic capacity. The feasibility and effectiveness of the TOD model was verified and recommendations were given for managing the transportation development and land use.

## INTRODUCTION

Naturally, within a community there is a direct interaction between the type and intensity of land use and transportation facilities provided. One of the primary objectives of coordinating any land use and transportation system is to ensure that there is an efficient balance between land use activity and transportation capability. Transportation can be visualized as the consequence of the fact that different types of land uses in the city are spatially separated. At the same time, enhanced mobility can also contribute to increased separation of land use. This symbiotic relationship between transportation and land use produces the movement and traffic flow patterns seen in urban areas. The accessibility of places has a major impact on land values, and the location of a place within the transportation network determines its accessibility. Thus, in the long run, the transportation system, and the traffic it carries, shapes the land use pattern (Khisty and Lall, 1998).

Unfortunately, for most cities in China, transportation and land-use decisions have all too often been regarded as distinctly separate issues in analysis, planning, designing, and evaluation. In addition, more people and higher percentages of the total population are now living in the cities.

Since 1978 when the central government started its open-door policies for economic reform, China has undergone rapid urbanization, experiencing a significant rise in per capita GDP and increased vehicle ownership. From 1949 to 2009, the total number of designated cities grew from 132 to 655. The percent of urbanization increased from 7.3% to 45.68%. This was accompanied by the growth of the total urban population. While growth in population and extent of urbanized areas are a rather expected outcome of economic development, concerns have been raised over the spatial pattern of the growth.

Transit-oriented development (TOD) is a strategy to integrate transit and land use by focusing development around the transit station. A typical TOD has a rail or bus station at its center, surrounded by relatively high-density development, with progressively lower-density development spreading outwards one-quarter to one-half mile, which represents pedestrian scale distances (Cervero et al., 2004). TOD is a particular category of smart growth, incorporating new efficient urbanism and location development. It can increase accessibility and transportation options through land use clustering and mix, and nonmotorized transportation improvements. It also reduces the distance required for car trips, allows a greater portion of trips to be made by walking and cycling, and allows some households to reduce their car ownership, which together can result in large reductions in vehicle travel. This reduces total transportation costs and helps create a more livable community. TOD has become the most popular model for the management of transportation development and land use.

This paper attempts to present a case study about coordinating transportation development and land use through the implementation of the TOD model. One of the residential and commercial districts in Chengdu, which is the largest city in the southwest of China, was selected for the model study.

## BACKGROUND

Peter and Doug (1989) developed a cohesive vision for the foundation of change in his book *The Pedestrian Pocket*. The research provided a model of a simple cluster of housing, retail space and offices within a quarter-mile walking radius of a transit system. It was the first time that the TOD model was proposed. Since then, many studies have been done about the TOD model applied to transportation development and land use. A host of urban design philosophies (new urbanism, transit-oriented development, traditional town planning) have gained popularity. Many case studies have been done in several cities and regions in the United States, such as Boston, New Jersey, Washington (D.C.), Chicago, the San Francisco Bay Area, and Southern California (Cervero et al., 2004).

Fehr and Peers Associates (1992) found substantially higher rates of foot and transit travel in traditional communities than conventional suburban subdivisions of the San Francisco Bay Area, though factors like differing transit service levels were not controlled for. The famous Land Use-Transportation-Air Quality (LUTRAQ) study, conducted by Parsons Brinckerhoff Quade and Douglas, Inc. (1993) for the Portland, Oregon region, was one of the more ambitious efforts to date to gauge the travel impacts of "pedestrian friendliness". Frank and Pivo (1994) found that mixed-use neighborhoods were most strongly correlated with walking to work, but rather surprisingly they had no influence on mode choice for shopping trips. Cervero and Landis (1996) reported that residential densities exerted a stronger influence on commuting mode choices than levels of land-use mixture, except for walking and bicycle commutes. For non-motorized commuting, the presence or absence of neighborhood shops was a better predictor of mode choice than residential densities. Cervero and Kockelman (1997) presented a study about 50 neighborhoods in the San Francisco Bay Area. That paper tested its proposition by examining how the 3Ds (density, diversity and design) affect trip rates and mode choice of residents. The research found that density, land-use diversity, and pedestrian-oriented designs generally reduced trip rates and encouraged non-auto travel in statistically significant ways.

Cervero et al. (2004) develop a model for predicting the effects of increased residential and commercial density, and improved walkability around a station on transit ridership. They conducted a large-scale study of TOD in California, including 26 residential developments. They found that TODs throughout California are providing a significant ridership bonus. Over

one-quarter of the TOD residents regularly commuted on transit, compared to 5.4% in the surrounding cities. Marc, et al. (2004) describe methods of evaluating TOD, taking into account urban form, pedestrian accessibility, transit usage, and socio-demographic change before and after transit-oriented development in two U.S. cities. They found that many transit stations were not optimally located to maximize pedestrian access, and that automobile-oriented streets (wide, with heavy and fast traffic) could create a significant barrier for walking. Borg and Orsini (2008) presented a study about the integrating land-use and transport by using TOD in small island states (Malta); the Maltese case study proved that the effect of TOD on communities and its possibilities applied strategically at all levels of city fabric and texture.

Yang et al. (2008) focused on developing an appropriate transit model in the small and medium-sized cities (SMSC) in China, and selected the city of Taicang for a case study. The results showed that the combined transit system could lead to an intensive land use pattern and gave better service to the low-density area in SMSC in China. Thomas and Deakin (2008) reported a case study focusing on the city of Jinan, capital of Shandong province in China. The author conducted preliminary land use analysis on a proposed BRT (Bus Rapid Transit) corridor, Lishan Road, incorporating Smart Growth principles as a framework for creating walkable neighborhoods. The paper presented that roads in Jinan, particularly those in the inner city, need to be pedestrian-friendly. This was even more critical within the context of implementing a BRT network.

## INTRODUCTION OF THE COMMUNITY

Chengdu, the capital city of Sichuan province, is a city with a 2,300-year history. It is the one of 24 historical and cultural cities in China. Chengdu currently has a population of about 11 million. The city's average population density is about 900 persons per square kilometer; even greater, the population density in the downtown area is about 2,300 persons per square kilometer. It is the 4$^{th}$ biggest city in China in terms of population. It is also known as the financial center of southwest China. Chengdu's GDP has risen steadily over the past ten years along with its vehicle population, particularly its small car population. By the end of 2007, there were about 1.8 million vehicles, and about 1 million electrical bicycles (Chengdu Statistical Yearbook, 2008). The traffic congestion problems have become more and more serious in the center of Chengdu during the recent years.

In this paper, one of the living communities of Chengdu is chosen for the case study. As shown in the Figure 1, the Yingbin Road is the main corridor in this community. There are several other main roads around it. The community is about 6 kilometers from downtown Chengdu. It is located in the national high technology development zone, which is approved by the Chinese Science Ministry. The study community is about 3 square kilometers, separated into 23 sections by the roads. Through the investigation, the data of current land use properties, population density and transportation were acquired. Taking into account the developed transportation system, intensification of the resident population and the greatly potential economic development in this area, the TOD can be applied here.

## TOD PLANNING

Many research results showed that in the TOD system, in order for residents to choose pedestrian travel, the suitable walking distance to the main road is about 400 to 800 meters. Compared to developed countries, in China there are lower rates of private car owners, higher population density and fewer public transport vehicles; therefore, the acceptable pedestrian distance should be longer (Li, 2008). In this study, the small 800m×800m area of the study community is chosen for the transportation development and land use planning by using the TOD model. As shown in the Figure 2, the study area is divided into 1024 (32×32) units; each unit has an area of about 625 square meters. Each unit can be developed as a separate construction unit. The top left corner unit is $u_{1,1}$, and the lower right corner unit is $u_{32,32}$. According to the land use planning in this region, the land use types ($k$ values) can be summarized as follows: 1-low density residential; 2-medium density residential; 3-high density residential; 4-policy protecting house; 5-small commercial and

service area; 6-large commercial center; 7-offices; 8-public open space; 9-non-development land; 10-industries; 11-old village house in urban area; 12-road; 13-outside of development regions.

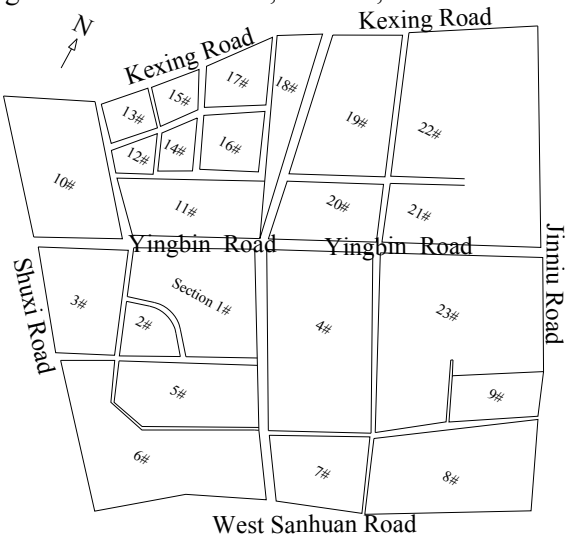



**Figure 1. Study Community**

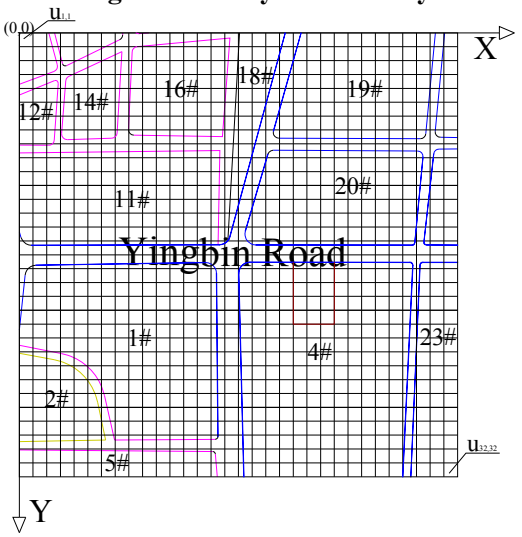



**Figure 2. Coordinates of the study area**

**Model theory**

The compact form of land use should be more effective for use of the social infrastructure. It can improve the flow of traffic throughout the area, and there will be more people to share the transportation facilities; it will also enhance the value of the surrounding land. Of course, land use that is overly-compact is also unacceptable. According to the research presented by Li (2008), the compactness index ($C_{ij}$) can be defined as the number of developed land units around the target units, as shown in the Figure 3. For the purpose of the compact form of land use, the maximum $C_{ij}$ should be collected in the community.

$$\max C_{ij} = \max \sum_{i} \sum_{j} \left( \sum_{i'=i-1}^{i+1} \sum_{j'=j-1}^{j+1} X_{i'j'}^{k} \right) \tag{1}$$

Where, $C_{ij}$ = the compactness index defined for land use

$k$ = the type of the land use

$X_{i'j'}^{k}$ = the dummy variables, if the land use type is $k$ in the unit $u_{i'j'}$, the value is 1, otherwise 0

| | | |
|---|---|---|
| (i-1,j-1) | (i-1,j) | (i-1,j+1) |
| (i,j-1) | $u_{ij}$ | (i,j+1) |
| (i+1,j-1) | (i+1,j) | (i+1,j+1) |

(a) $C_{ij} = 5$

| | | |
|---|---|---|
| (i-1,j-1) | (i-1,j) | (i-1,j+1) |
| (i,j-1) | $u_{ij}$ | (i,j+1) |
| (i+1,j-1) | (i+1,j) | (i+1,j+1) |

(b) $C_{ij} = 2$

**Figure 3. The compactness of land use**

Cervero and Kockelman (1997) found that land-use diversity generally reduced trip rates and encouraged non-auto travel in statistically significant ways. In the developing community, it is important to improve land use diversity. For that purpose, the maximum $X_{ij}^{k}$ should be collected in the community.

$$\max \sum_{i,\ j} (X_{ij}^{k}) \tag{2}$$

Where, $X_{ij}^{k}$ = the dummy variables, if the land use type is $k$ in the unit $u_{ij}$, the value is 1, otherwise 0

In the TOD model, it is important to have a convenient transportation system for the commercial center in the community, especially the major public transport. For the study community, the major public transportation system is the city bus, so there should be shortest walking distance from the commercial center to the bus station.

$$\min(L_{ij} X_{ij}^{6}) \tag{3}$$

Where, $X_{ij}^{6}$ = the dummy variables, if the land use type is large commercial center in the unit $u_{ij}$, the value is 1, otherwise 0

$L_{ij}$ = the distance from the center of unit $u_{ij}$ to the bus station

The compactness and diversity of land use may lead to a decline in quality of life for the citizens. Therefore, it is reasonable to allocate the suitable land use type. Li (2008) presented an indirect index ($C^{kk'}$) to quantify the compatibility according to the noise, pollution, traffic, social habits and Chinese psychological factors; the value is from 0 to 10. The minimal conflict value should be collected in the community.

$$\min(\sum_{i,i'=1}^{32} \sum_{j,j'=1}^{32} X_{ij}^{k} X_{i'j'}^{k'} C^{kk'}) \tag{4}$$

Where, $C^{kk'}$ = the value for evaluating the conflict of the adjacent unit, [0, 10]

In the TOD system, it is most important to encourage households to use the public transportation as much as possible, and reduce their car ownership. It is reasonable to maximize the proportion of the public transportation use.

$$\max \sum_{i,j} \sum_{k=1}^{4} r^{k} d_{ij}^{k} O_{ij}^{k} + \sum_{i,j} \sum_{k=5}^{7} r^{k} d_{ij}^{k} D_{ij}^{k} \tag{5}$$

Where, $r^{k}$ = the rate of public transport

$d_{ij}^{k}$ = the density of land use

$O_{ij}^{k}$ = the total generated traffic in the residential unit $u_{ij}$ ($k$ =1, 2, 3, 4)

$D_{ij}^{k}$ = the total attracted traffic in the commercial unit $u_{ij}$ ($k$ = 5, 6, 7)

According to the actual characteristics of Chengdu, it is reasonable to determine the amount of pollution generated by land development and the corresponding treatment costs. Coordinating land use and transportation should minimize the environmental impact. Li (2008)

presented the average emissions and treatment costs for the different land use types, as shown in the Table 1.

$$\min(\sum_{k=1}^{7}(\sum_{i=1}^{32}\sum_{j=1}^{32}P_k^r c^k d_{ij}^k X_{ij}^k)) \quad (6)$$

Where, $c^k$ = the emissions of waste in the unit $u_{ij}$ (Ton/year)

$P_k^r$ = the treatment costs (RMB /Ton)

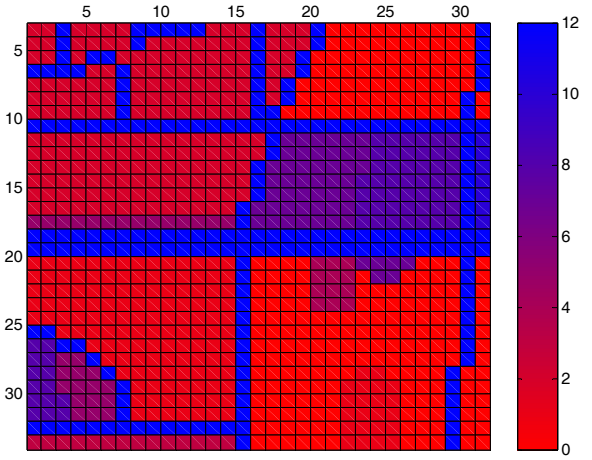
**Table 1. The emissions and treatment costs for different land use types.**

| Land use type | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $CO_2$ | Emissions(Ton/year) | 15.0 | 7.53 | 5.0 | 3.0 | 1.5 | 1.5 | 3.5 |
| | Treatment costs (RMB/Ton) | 308 | | | | | | |
| Wastewater | Emissions (Ton/year) | 400.0 | 210.6 | 140.0 | 140.0 | 2.38 | 2.38 | 4.76 |
| | Treatment costs (RMB /Ton) | 0.9 | | | | 1.2 | | 1.1 |
| Solid waste | Emissions (Ton/year) | 3.0 | 1.296 | 0.88 | 0.88 | 0.25 | 0.25 | 0.28 |
| | Treatment costs (RMB /Ton) | 125 | | | | | | |

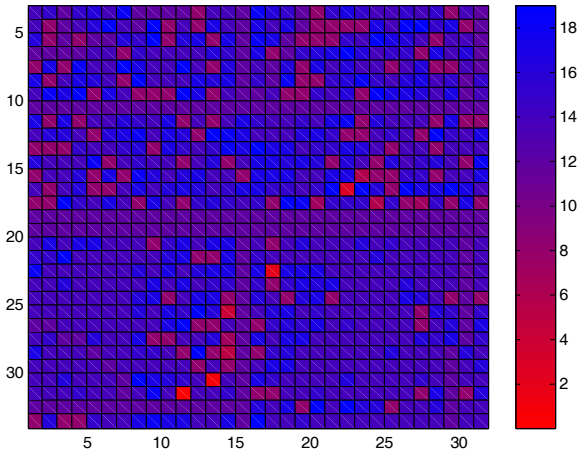
**Optimization solution**

According to the actual situation in Chengdu, the constraints are identified for the construction density, mixed land use restrictions and the minimal rate of public transport. For the optimization solution, there are two major values, land use type ($X_{ij}^k$) and the land use density ($d_{ij}^k$), to be decided for every unit. The first value, land use type ($X_{ij}^k$), can be obtained from the optimization solution of the objective functions (Formula (1), (2), (3), (4)). Then, the land use density ($d_{ij}^k$) can be obtained from the optimization solution of the objective functions (Formula (5), (6)). It is not possible to obtain the optimal at the same time, for the two objectives conflict. The method of weighted mean is adopted to solve that problem here.

According to the four objective functions and constraints, the optimal solution for the land use type is obtained by using the genetic algorithm (GA) approach and MATLAB software. Through the iteration, the objective functions get a fast convergence. The comparison of the current land use type with the optimal solution is shown in the Figure 4. This solution should be appropriately adjusted in the actual planning.




(a) The current land use type

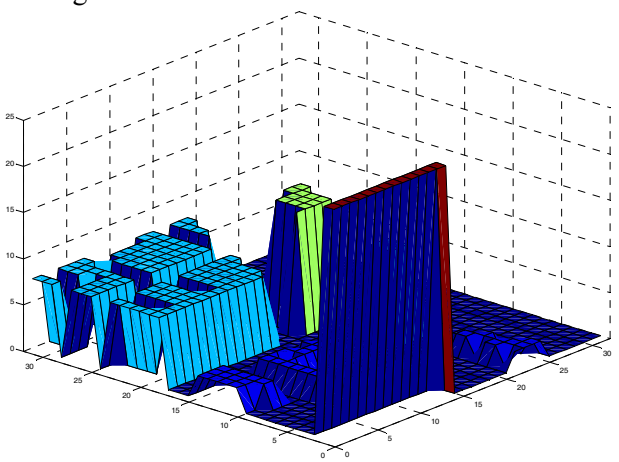



(b) The optimal solution

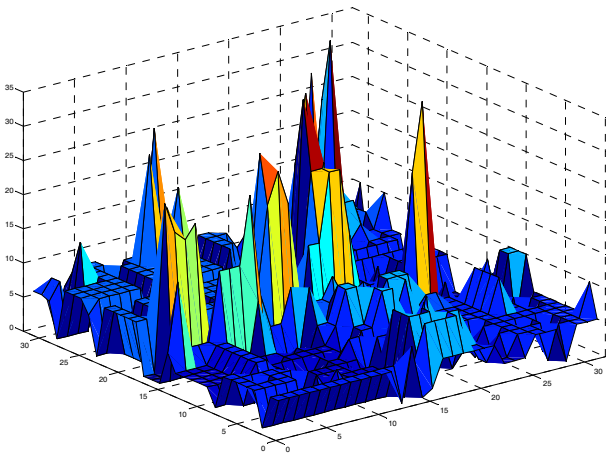
**Figure 4. The land use type**

The optimal solution for the land use density is obtained from the last two objective functions and the method of weighted mean, by using the genetic algorithm (GA) approach and MATLAB software. When running to the 2000 generation of the feasible solutions in the programme, there is a stable convergence. There are 5,754 public transportation passengers every day, and the cost of pollution treatment is about RMB 9.34 million for per year. The existing bus system capacity is only 1,000 persons per day in the Yingbin Road; therefore, it is reasonable to increase the capacity of the bus lines. The comparison of the current land use density with the optimal solution is shown in the Figure 5.

According to further verification of transportation coordination in the study community, it is shown that the future traffic in the Yingbin road will be about 4700 pcu/hour (current traffic volume is about 1200 pcu/hour, added to the increasing traffic which about 3500 pcu/hour from the new land use optimal solution), but the current traffic capability of Yingbin road is about 7000 pcu/hour, which means that it will be operating at just about 67 percent of the capability. The traffic will be smooth, and the TOD planning is feasible.



(a) The current land use density



(b) The optimal solution

**Figure 5. The land use density**

## CONCLUSION

Based on the above, a typical case study about the Transit-oriented transportation development and land use planning is presented in this paper. The conclusions are as follows.

The compactness, diversity and convenience of the public transport system are the major objectives for applying the TOD model in the transportation development and land use planning. It is also important to coordinate the land use development with the natural environment, reducing the environmental pollution caused by land development as much as possible.

Compared with the optimal solution of land use type in this study community, the current land use type is too simple, and the community is arranged into independent pieces, which obviously increases the travel distance. This arrangement is not suitable for making full use of the transportation services and coordinating the relationship between the land use and transport.

From the optimal solution of the land use density, it is an obvious solution to increase the density of commercial development on both sides of the Yingbin road and near the bus station, which is suitably arranged in the middle section of the major road. This will improve the accessibility of the business center and reduce the walking distance.

Further verification of transportation shows that the traffic on the Yingbin road will be smooth. The TOD planning is demonstrated to be feasible and effective for this study community.

## REFERENCES


Borg, M., and Orsini, R. (2008). Transit oriented development - Integrating land-use and transport in small island states. *WIT Transactions on the Built Environment*, 101, 457-466.

Cervero, R., Murphy, S., and Ferrell, C. (2004). Transit Oriented Development in the United States: Experiences, Challenges, and Prospects. *TCRP Report 102*.

Cervero, R., and Kockelman, K. (1997). Travel Demand and the 3D's: Density, Diversity, and Design. *Transportation Research D*, 2(3), 199-219.

Cervero, R., and Landis J. (1996). Mixed land uses and commuting: evidence from the American housing survey. *Transportation Research A*, 30(5), 361-377.

Chengdu Statistics Bureau (2008). Chengdu Statistical Yearbook 2008. China Statistics Press. Beijing, China.

Frank, L., and Pivo, G. (1994). The impacts of mixed use and density on the utilization of three modes of travel: the single occupant vehicle, transit, and walking. *Transportation Research Record 1466*, 44-52.

Fehr and Peers Associates (1992). Metropolitan Transportation Commission bay area trip rate survey analysis. *Metropolitan Transportation Commission*, Oakland, CA.

Khisy, C. J., and Lall, B. K. (1998). Transportation Engineering. *Prentice Hall*, New Jersey.

Li, Y. (2008). Sustainable mass rail transit based on Transit-Oriented Development. Harbin Institute of Technology University. Harbin, China.

Marc, S. (2004). Comparing transit-oriented development sites by walkability indicators. *Transportation Research Record 1887*, 34-42.

Peter, C., and Doug, K. (1989). The Pedestrian Pocket Book: A new Suburban Design Strategy. *Princeton Architectural Press*.

Parsons Brinckerhoff Quade and Douglas, Inc. (1993). The Pedestrian Environment: Making the Land Use, Transportation, Air Quality Connection. *LUTRAQ Project*, Vol. 4A, Portland, Oregon.

Thomas, A., and Deakin, E. (2008). Land use challenges to implementing transit-oriented development in China: Case study of Jinan, Shandong Province. *Transportation Research Record 2077*, 80-86.

Yang, C., Wang, W., Lu, J., and Wan, Q. (2008). A research on adaptive transit combined with transit-oriented development for small and medium-sized cities. *International Conference on Intelligent Computation Technology and Automation*. Hunan, China. 305-310.

# Insights on the determinants of walk trips using large scale travel survey data

C. Morency[1], M. Trépanier[2], V. Tremblay[3] and E. Poliquin[4]

[1]Catherine Morency, Ph.D., P.Eng., Assistant professor (Corresponding author), Department of civil, geological and mining engineering, email: cmorency@polymtl.ca, [2]Martin Trépanier, Ph.D., P.Eng., Associate professor, Department of Mathematics and Industrial Engineering, email: mtrepanier@polymtl.ca, [3]Valérie Tremblay, M.A.Sc., Department of civil, geological and mining engineering, email: vt_tremblay@hotmail.com, [4]Eric Poliquin, B.Eng., B.A., master's student, Department of civil, geological and mining engineering
Ecole Polytechnique de Montréal, C.P.6079, Succ.Centre-Ville, Montréal QC H3C 3A7, Canada, PH (514) 340-4711 ext.4210, FAX (514) 340-5763, email: eric.martel.poliquin@polymtl.ca

## ABSTRACT

The purpose of this research is to contribute to the body of knowledge on the determinants of walk trips in urban areas using individual data from a large-scale travel survey data. The main purpose of this research is to identify features that can have an influence on the individual decision to walk to travel short distances (less than 1.6 km).

Benefiting from the availability of microdata from a large-scale travel survey (2003, sample of some 60,000 households), a study of short trips is conducted for people living in more than sixty districts of the Montreal Area. Walk trip shares are estimated as well as a set of variable describing urban structure (population density, tortuosity of the trip on the road network, intersection density). Using these variables as well as the individual and households attributes, a model describing mode choice for short trips, is estimated at the individual level.

## INTRODUCTION

The share of walking among urban trips has continuously declined through the years. Well aware of the multiple benefits of active transportation, both for human health and the environment, urban planners have tried to reverse this trend. Convincing people to walk for their short trips is not an easy task. Having more insight into the factors influencing this behavioral choice could help focus on the best strategies and interventions.

The objective of this paper is to contribute to the current discussions and researches on the determinants of walking in urban areas. It benefits from the availability of trip microdata from a large-scale travel survey held in the Greater Montreal Area (GMA). It focuses on the mode choice of individuals for short trips, less than 1,600 meters (1 mile), looking for variables that could help better understand the observed behaviors. A model is estimated at the individual level to see what attributes have a significant impact on the decision to walk.

The paper is organized as follows. First, an overview of the body of knowledge regarding the determinants of walking is provided. Then, the study area, namely the GMA, is described using various key facts mostly outputted from travel surveys. A descriptive analysis of the main factors *a priori* considered to affect the

decision to walk is presented. Then, building from this analysis, a discrete choice model (trip level) is estimated to observe the links between the decision to walk for short trips and various attributes. Results are then discussed.

## BACKGROUND

In a majority of North American cities, urban environment fosters the use of the private car at the detriment of walking (Frank and Engelke, 2000). This trend has even started to reach older cities, for example in Europe, that were formerly recognized as walking paradise. The omnipresence of the car in the transportation scheme has become an issue world around. It is now well acknowledged that the high dependence on private cars, for short and long travel, increases road congestion, pollution, soil erosion, ground loss and destruction of wildlife habitats, among others (Morency et al., 2009).

In addition, the benefits of walking are recognized by many health professionals. Indeed, walking is very important for public health; it helps keep muscles healthy, it decreases the risks of depression and anxiety as well as other psychological troubles (Frank and Engelke, 2000, Sallis et al., 2005). Moreover, this activity is satisfactory.

### The importance of active modes

A study carried out in 1995 in the United States showed that 86% of the trips are made by car while walking accounts for only 5% (FHWA, 1997). Knowing that physical inactivity is responsible for about 200,000 deaths per year in the United States only, a so low a level of walking is critical.

As mentioned before, North American cities are car-oriented. As many specialists agree, one cannot hope that inactive people will change behaviors and increase their level of physical-activity in a non-friendly environment (Demers, 2008, Loukopoulos et al., 2005). Walking trips are determined by travel distance, road safety, route attractiveness and network connectivity. Elements such as presence of trees, public telephones or urban art can also attract pedestrians, notwithstanding their destination (Southworth, 2005). A convivial environment, the presence of cycle paths or adequate pavements can also encourage active transportation as well as increased awareness of the population on the possible benefits (Sunseri and Walton, 2007). Moreover, proximity to pedestrian infrastructures as well as their general accessibility are also important in encouraging people to walk (Sallis et al., 1990). Litman (2003) showed that districts with such designs are more accessible, increase quality of life, improve public health, contribute to the local economic development and are more equitable.

### Barriers to active transportation

Several studies support the idea that there are two groups of elements that discourage people to walk: personal and environmental barriers. On the one hand, personal barriers are subjective considerations that influence the way people perceive and choose among available transportation alternatives. On the other hand, environmental barriers are related to the infrastructures and services availability and structure. King et al. (2003) report that a strong relation exists between density of activity locations in

the residence' neighborhood and the level of physical activity of people. It suggests that living within walking distance of attractive destinations such as parks, groceries or pedestrian paths can have a significant effect on the daily behavior of people. Using data from a stated preference survey, Riley reports that the decision to walk is more influenced by parking costs than by gas prices (Riley, 2008). Few other studies focus specifically on walking trips; some deal with walking as part of the mode choice process, among other modes (Habib et al., 2009), or concentrate solely on walking for recreational trips (Bhat and Gossen, 2004). In all cases, Logit models are often used to predict mode choice.

Safety is also identified as an essential value to support walking in neighborhoods. People who do not perceive their district as being safe tend to make fewer physical activities near their home place (Riley, 2008, Pucher, 2003). In Sailor County, California, the "Safe Routes to School" program was set up to identify and develop fixed paths for children to allow them to go to school using an active mode. In the second year of the program, the participating schools observed a 64% increase of walk trips and a 114% increase of bicycle trips (Staunton et al., 2003).

## STUDY AREA AND KEY FACTS

This research relies on trip microdata gathered during the 2003 Origin-Destination travel survey data conducted in the GMA (details on the surveys available at: http://www.cimtu.qc.ca/index.asp). In the region, these types of survey have been conducted for more than forty years, approximately every five years. In the fall of 2003, about 70,000 households were surveyed (sampling rate of approximately 5%). In addition to information regarding the households and the people, detailed trip diary data are collected for people aged 5 years and older (Trépanier et al, 2008). For the purpose of this research, the short trips (those with a distance of 1.6 kilometer or less) are selected for the analysis. Distances are measured on the road network, using a path calculator developed for the user information trip calculator (Trépanier and Chapleau, 2002). 58,000 short trips (sample) are included in the analysis.

Key facts on the region's districts are provided in Table 1 (based on aggregated values for the districts). Most of the features present high variability and high ranges of values often linked to the spatial location of the districts. The upcoming analysis will confirm some of these spatial relations. Features such as car ownership, household income, main occupation and age are anticipated to be relevant in the decision to walk or not to perform short trips. Other variables related to the structure of the trip and district will also be examined.

**Table 1. Key facts of the analysis districts**

| | Mean | Standard deviation | Minimum | Maximum |
|---|---|---|---|---|
| **District features** | | | | |
| Area ($km^2$) | 54.24 | 73.92 | 0.84 | 486.49 |
| Population density (people/$km^2$) | 14,695 | 11,630 | 402 | 54,000 |
| Household density (HH/$km^2$) | 970 | 1215 | 111 | 6,487 |
| **Household features** | | | | |
| HH size | 2.47 | 0.32 | 1.56 | 3.08 |
| HH car ownership | 1.43 | 0.37 | 0.44 | 2.00 |
| % HH with no car | 41.10% | 7.52% | 25.64% | 58.44% |
| % low income HH (< 20k$) | 12.45% | 7.01% | 2.49% | 31.11% |
| % HH with 1 person | 25.00% | 11.11% | 10.28% | 60.28% |
| % French speaking | 84.43% | 18.98% | 30.38% | 100.00% |
| **People features** | | | | |
| Average age | 37,24 | 3,06 | 30,17 | 47,65 |
| % 16 years and older with driving license | 82,72% | 9,10% | 59,29% | 98,02% |
| % 65 years and older | 11,94% | 4,90% | 2,83% | 32,48% |
| % full-time workers | 42,12% | 3,83% | 29,08% | 50,15% |
| **Travel behaviors** | | | | |
| Trip rate (5 y.o. and older) | 2,30 | 0,13 | 1,94 | 2,58 |
| Modal share of walking for short trips | 29,76% | 15,57% | 0,68% | 75,47% |

## WALKING IN MONTREAL

In Montreal, the modal share of walking has been declining for several years. In 2003, the average share of walking for short trips (1.6 km and less) was less than 30%.

### Relation with distance to CBD

Hence, as for all other modes, this share varies spatially. Actually, there is a strong relation between the modal share of walking for short trips and the distance to CBD (from home), a proxy variable often used to summarize spatial trends. As can be seen in Figure 1, modal share of walking for short trips can be as high as 75% for central districts and slowly declines to less than 10% in suburban areas.

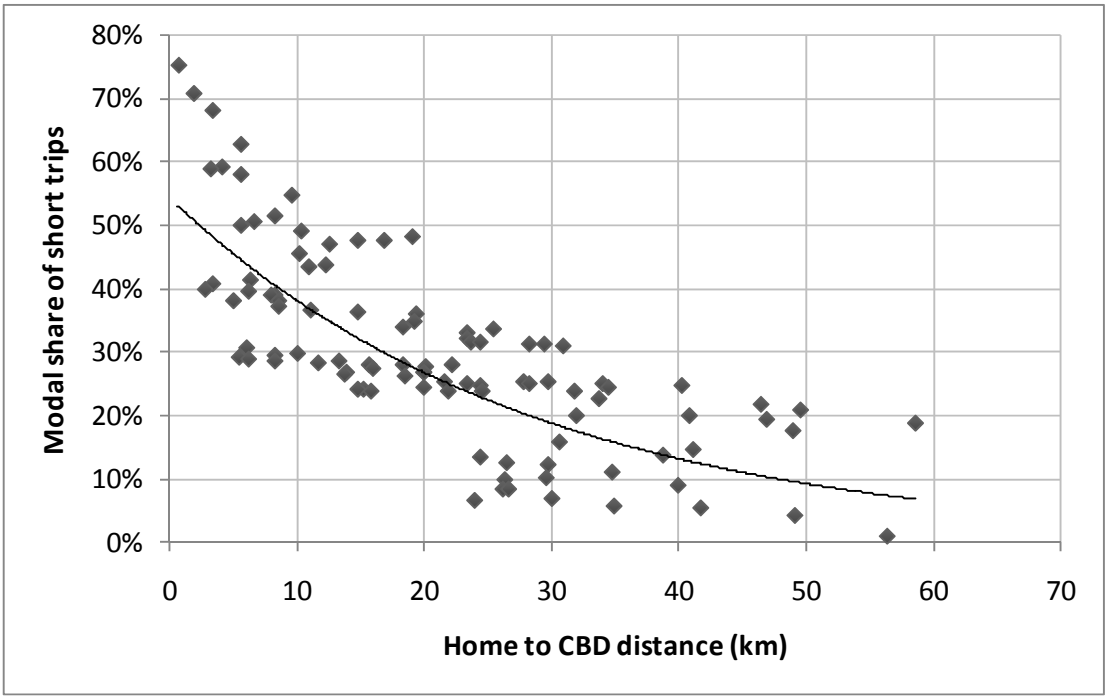



**Figure 1. Modal share of walking for short trips (< 1.6 km), as a function of home to CBD distance, in the various districts of the GMA**

**Relation with population density**

There is a positive correlation between population density and the share of walking for short trips (Pearson coefficient of 0.70). Obviously, walking increases as population density increases but the relation is more ambiguous in low density districts where values are highly variable. Actually, for districts of density less than 1,000 people per square kilometers (47 districts), share of walking ranges from 0.9% to 47.2% (average of 21.6% with coefficient of variation of 50.0%).

**Relation with age**

Other factors are affecting behaviors. Age has an important impact on the decision to walk for short trips. Figure 2 presents the distribution of modal share of walking observed in each of the 100 zones, according to age group. It also presents the mean modal share of walking in the GMA. Obviously, people too young to drive have higher shares of walk trips.

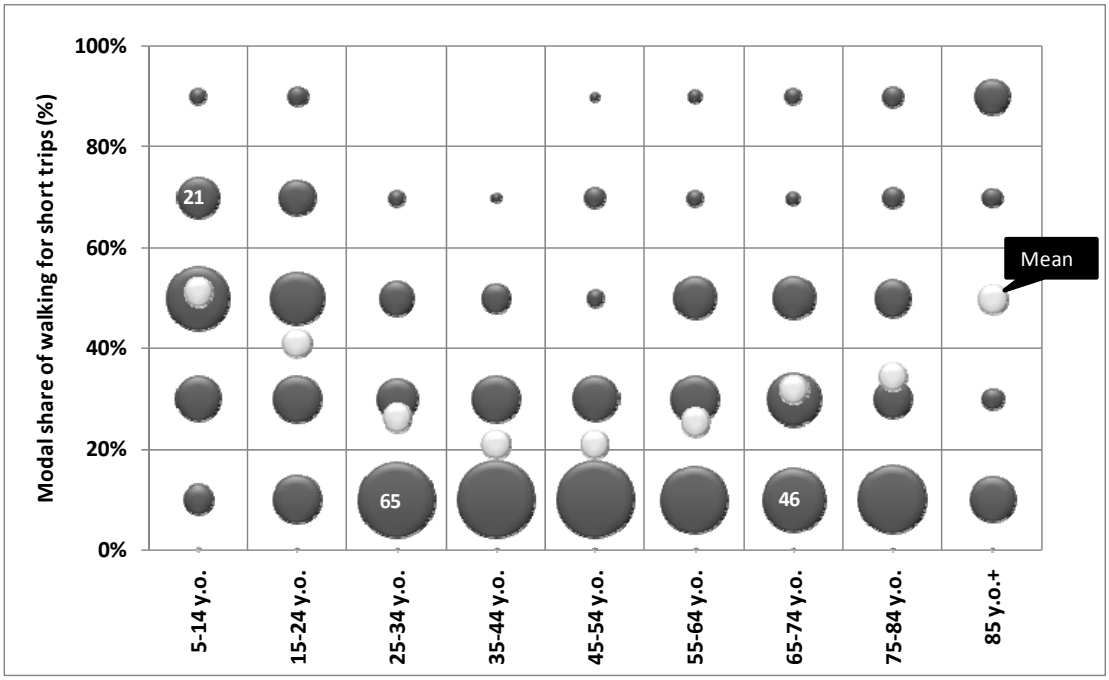



**Figure 2. Distribution of modal share of walking for short trips (< 1.6 km), by population segments (bubble size: number of districts)**

**Relation with household income**

Household average income also seems to be related to the walk share of short trips. Low-income households (less than 20 k$ CAD per year) have the higher share of walking for short trips (average of 40.0%). The share then decreases and stabilizes at around 27% for trips made by people belonging to households owning some 40-100 k$ per year. Share is slightly lower for high income households (21.8%). Still, the variability of this average modal share is the highest for low-income households.

**Street grid pattern**

The relation between the structure of the road network and modal choice is also examined. Often, road patterns in the suburbs are in the shape of crescent and half-moon to slow down cars. This type of road network stands at the opposite of older road networks, organized in more efficient grids that reduce travel distances and increase path alternatives. The suburban-type road network complicates all kinds of trips. Various indicators are used to describe the features of each trip or zone:

- **Tortuosity of the trip on the road network**: This represents the ratio between the network distance that would be traveled by car, and the Euclidean distance (straight line between origin and destination points). On the district level, the average tortuosity index on 1.46 for the region, but it varies from 1.2 (CBD) to 2.3 (suburbs).
- **Intersection density:** The intersection density the ratio between the number of single intersections and the area of the sector.

The influence of these variables, amongst others, is studied using a discrete choice model estimated at the trip level.

## DISCRETE CHOICE MODEL

A simple discrete choice model is estimated at the trip level to observe the influence of various variables in the decision to walk or not when travelling a short distance. Actually, a binomial Logit model is appropriate, at the person level, when there are only two possible outcomes: in our case, walking or not for a specific trip.

### Application of the model

The model developed is based on the assumption that a person will make the choice that maximizes his utility. This means that an individual will choose alternative *i* if and only if:

$$U_{ni} \geq U_{nj} \; , \forall \, j \neq i$$

where *U* represents the utility for the alternative *i* and *n* represents a particular person.

Let us consider a case with two alternatives belonging to a set of choices where $y_{ni}$ is the observed choice. We will have:

$$y_{ni} = \begin{cases} 1 & \text{when } U_{in} \geq U_{jn}, \quad \forall \, j \neq i \\ 0 & \text{otherwise} \end{cases}$$

A set of explanatory variables were inputted in the model in order to capture potential interactions. These variables relate to household income, age group, district of home residence, trip distance and possession of a driving license. Most of the variables were coded as dummy variables to facilitate the interpretation of results. Various combinations of variables were tested, namely those related to the districts because some of them present perfect colinearity or high correlation.

### Results of the model

The results of the most interesting model are presented in Table 2. All the independent variables are significant at the 99% level and overall, the model is capable to describe about 22% of the inherent variability of the sample. Results show for instance that:

- The probability to walk for a short trip is 1.22 times higher when living in a low-income household, compared to people living in households with incomes of 60 k$ and more;
- Having access to no car within the household increase the probability of walking for short trips by 3.22;
- Young people have higher probability of walking for short trips and the elderly are those who have the lowest probability, all else being equal;
- The probability to walk for people with a driving license is 0.32 times the probability of those not having one;
- As anticipated, the probability to walk decreases with trip length. It is more than eight times higher for trips less than 400 meters, more than four times higher for trips between 400 and 800 meters and almost twice as high for trips of 800 to 1200 meters when compared to trips of 1200 meters and more (reference group);
- The impact of the home spatial location is obvious: probability to walk for short trips decreases as distance from CBD increases, being 2.5 more important for trips made by people living at less than 5 km from the CBD.

- Probability to walk for short trips is hence lower when the tortuosity of the trip on the road network is low, suggesting that a high tortuosity may discourage drivers while they may be walking shortcuts only available to pedestrian.
- Finally, intersection density (that was highly correlated to both population and household density) is positively linked to the probability to walk for short trips, suggesting and increasing easiness to travel and a higher level of connectivity of the available routes.

**Table 2. Results of the Logit discrete choice model**

| | | | | | |
|---|---|---|---|---|---|
| | Number of observations | 58304 | | | |
| | Wald chi2(17) | 10948.1 | Log likelihood | | -30882 |
| | Prob>chi2 | 0.000 | Pseudo R2 | | 0.2261 |
| **pied** | **Variable description** | **Odds Ratio** | **z** | **P>\|z\|** | **Mean value** |
| Incl20k | HH income <20k$ | 1.2243 | 5.97 | 0.000 | 0.1468 |
| Inc2060k | HH income >=20k$ and <60k$ | 1.1452 | 6.07 | 0.000 | 0.4042 |
| *Inc60kmore* | *HH income >=60k$* | *ref* | | | 0.4489 |
| HH0car | HH with no car | 3.2235 | 35.44 | 0.000 | 0.1692 |
| *HH 1carm* | *HH with at least one car* | *ref* | | | 0.8308 |
| P3554yo | 35 to 54 years old | 0.6841 | -13.47 | 0.000 | 0.2922 |
| P5564yo | 55 to 64 years old | 0.8461 | -4.03 | 0.000 | 0.0879 |
| P65yom | 65 years and older | 0.5840 | -13.67 | 0.000 | 0.0981 |
| *P534yo* | *5 to 34 years old* | *ref* | | | 0.5218 |
| Drilic | With driving license | 0.3246 | -44.67 | 0.000 | 0.5363 |
| *Ndrilic* | *No driving license* | *ref* | | | 0.4637 |
| wd400 | Trip distance of less than 400 meters | 8.4134 | 60.12 | 0.000 | 0.1604 |
| wd4_800 | Trip distance of 400-800 meters | 4.2454 | 48.4 | 0.000 | 0.3011 |
| wd8_1200 | Trip distance of 800-1200 meters | 1.9755 | 22.15 | 0.000 | 0.2913 |
| *wd12_1600* | *Trip distance of 1200-1600 meters* | *ref* | | | 0.2472 |
| dcbd_05 | Home-CBD distance of less than 5 km | 2.5609 | 22.46 | 0.000 | 0.1787 |
| dcbd_510 | Home-CBD distance of 5-10 km | 1.4582 | 11.7 | 0.000 | 0.2749 |
| dcbd_1015 | Home-CBD distance of 10-15 km | 1.1378 | 4.1 | 0.000 | 0.1767 |
| *dcbd_15m* | *Home-CBD distance of 15 km and more* | *ref* | | | 0.3697 |
| tm_12 | Tortuosity of the trip less than 1.2 | 0.5241 | -21.91 | 0.000 | 0.2137 |
| tm_1213 | Tortuosity of the trip of 1.2-1.3 | 0.5965 | -16.32 | 0.000 | 0.1563 |
| tm_1315 | Tortuosity of the trip of 1.3-1.5 | 0.6754 | -15.19 | 0.000 | 0.2801 |
| *tm_15m* | *Tortuosity of the trip of 1.5 and more* | *ref* | | | 0.3499 |
| densint | Intersections per square kilometers | 1.0062 | 8.85 | 0.000 | 39.4548 |

## CONCLUSION AND PERSPECTIVES

This paper has documented some variables related to the decision to walk using a large sample of short distance trips (< 1.6 km) from a travel survey. Many variables are correlated to such decision, may them be related to the traveler or its belonging household (age, possession of a driving license, household size, car ownership and income), the trip (length and tortuosity) or the home location (intersection density, distance to the CBD). Supported by a descriptive analysis, a logit model has confirmed the strength of the observed relations as well as the differences due to the scale of the independent variables. Actually, the results of this experience show that short distance walk trips are influenced by intersection density (higher density will generate more walk trips), household income and car ownership (no car, more walk), age (younger people walk more), and tortuosity (discourage the use of motorized modes).

This special interest to walk trips can lead to broader perspectives. This research can be generalized and applied to other large household surveys in order to confirm the first results. In our case, the model will be estimated using historical datasets as well as the data of the 2008 OD survey, when it will become available. In addition, other variables like traffic index and public transit service index could be derived to better take these dimensions into account. More variables related to the structure of the household will also be added (namely the presence of children), the trip (purpose for instance) as well as the linkage of the trip within activity chains. Other modeling approaches will also be examined such as model with spatial expansion factors and interaction variables to further the analysis and capture the impact of spatial location. However, the current study is a step towards stronger collaboration between public health and transportation researchers, aimed to promote physical activities and the use of walk in daily commute.

## ACKNOWLEDGMENTS

The authors wish to thank the Metropolitan Transportation Agency, the Ministry of Transport of Quebec and their partners for providing access to data from the Montreal household surveys, and the National Science and Engineering Research Council of Canada for its financial support. They also wish to acknowledge collaborators from PasseportSante.net who provided access to data for research purposes. Thanks to Marie Demers for its judicious advises.

## REFERENCES

Bhat, C.R., Gossen, R. (2004). A mixed multinomial logit model analysis of weekend recreational episode type choice, Transportation Research Part B: Methodological, Vol. 38, No. 9, November 2004, pp. 767-787.

Demers, M. (2008). *Walk for Your Life*. Vital Health Publishing, 2006, 244 p.

Federal Highway Administration. (1997) *Our Nation's Travel: 1995 NPTS Early Results Report*. Washington, D.C.: U.S. Department of Transportation.

Frank, L.D. and P. Engelke. (2000). How Land Use and Transportation Systems Impact Public Health: A Literature Review of the Relationship Between Physical Activity and Built Form, *ACES: Active Community Environments Initiative Working Paper #1*.

Habib, K.M.N., Day, N. and E.J. Miller. An investigation of commuting trip timing and mode choice in the Greater Toronto Area: Application of a joint discrete-continuous model, *Transportation Research Part A: Policy and Practice*, Vol. 43, No. 7, August 2009, pp. 639-653.

King W.C., Belle, S., Killingsworth, R., Fenton, M. and Kriska, A.M. (2003) The relationship between convenience of destinations and walking levels in older women. *American Journal of Health Promotion*, Vol. 18, No. 1, September-October 2003, pp. 74-82.

Litman, T. A. (2003). *Economic Value of Walkability*, *82th Annual Conference of the Transportation Research Board, Washington*, D.C., CD-ROM, 2003.

Loukopoulos, P., Jakobsson, C., Gärling, T., Schneider, C.M., Fujii., S. (2005). Public attitudes towards policy measures for reducing private car use: evidence from a study in Sweden. *Environmental Science & Policy*, Vol. 8, No. 1, February 2005, pp. 57-66.

Morency, C., Roorda, M. and Demers, M. (2009). Steps in Reserve: Comparing Latent Walk Trips in Toronto and Montreal, Transportation Research Record: Journal of the Transportation Research Board, Issue 2140, pp. 111-119.

Pucher, J. (2003). Promoting Safe Walking and Cycling to Improve Public Health: Lessons From The Netherlands and Germany. *Public Health Matters*, Vol. 93, September 2003, pp. 1509-1516.

Riley, T.J. (2008). The propensity for motorists to walk for short trips: Evidence from West Edinburgh, *Transportation Research Part A: Policy and Practice*, Vol. 42, No. 4, May 2008, pp. 620-628.

Sallis, J., Hovell, M., Hofstetter, C.R., Elder, J., Hackley, M., Caspersen, C., and K. Powell. (1990). Distance Between Homes and exercise Facilities Related to Frequency of Exercise Among San Diego Residents, *Public Health Reports*, Vol. 105, No. 2, March-April 1990, pp. 179-85.

Sallis, J.F., Frank, L.D., Saelens, B.E. and M.K. Kraft. (2005). Active transportation and physical activity: opportunities for collaboration on transportation and public health research. *Transportation Research Part A: Policy and Practice*, Vol. 38, No. 4, May 2004, pp. 249-268.

Southworth, M. (2005). Designing the Walkable City. (2005) *Journal of urban planning and development*, Vol. 131, No. 4, December 2005, pp. 246-257.

Staunton, C.E., Hubsmith, M.D. and W. Kallins. (2003) Promoting Safe Walking and Biking to School: The Marin County Success Story. *American Journal of Public Health*, Vol. 93, No. 9, 2003, pp. 1431-1434.

Sunseri, S., Walton, D. (2007). Impediments to Walking as a Mode Choice Land Transport NZ Research. *Land Transport New Zealand Research Report 329*, 46 p.

Trépanier M., Chapleau R. (2002). Transit itinerary calculation on the Web based on a transit user information system, *Journal of Public Transportation*, Vol. 5, No. 3, 2002.

Trépanier M., Chapleau R., Morency, C.(2008) Tools and methods for a transportation household survey, *Urban and Regional Information Systems Association Journal*, Vol. 20, no. 1, 2008, pp. 35-43.

# Public Private Partnership: A Prospective in the Development of Sustainable Transport Infrastructures and Services in Bangladesh

Md. Shoaib Chowdhury, PhD, PE, MASCE [1]

[1]Adjunct Associate Professor, Department of Civil Engineering, The City College of New York, 160 Convent Avenue, New York, NY 10031; PH (212)-650-7829; email: mchowdhury@ccny.cuny.edu

## ABSTRACT

Bangladesh with limited revenues and development budgets falls far of a short of funding to meet current and future transportation needs. Recently, Bangladesh government has taken a fresh public private partnership (PPP) policy initiative to encourage private sector investors in the development of public infrastructures. In light of such a policy, this paper examines the importance of strengthening policies, institutions particularly in handling design/build contract procurements, developing transparent risk allocation mechanisms, and managing projects and programs for successful policy executions. Government institutions must also define overall goals and objectives of projects and explore PPP opportunity in the early development stage of the project to take advantage of not only private investment but also assistance in the planning, designing and building of projects with appropriate skilled manpower and technology for quick and successful implementations of projects.

## INTRODUCTION

Providing needed transport infrastructures and services is the key for sustained economic growth and development in the era of free market and global economy. However, Bangladesh government has been underfunding in the transport infrastructure sector for decades mainly due to significant shortfall of fiscal resources and gradual reduction of support from international development partners. As a result, physical infrastructures in general and transport infrastructures in particular have been remained major obstacles in the growth and development of business and economy. While demand for transportation has been steadily increasing with the growth of population and urbanization, the funding gap for investment in the infrastructure sector has been widening further. Bangladesh has yet to develop integrated intermodal national transportation infrastructures and systems linking major urban areas. In addition, rapidly growing unplanned urbanization of Dhaka, the capital of Bangladesh, has been created enormous pressure to the concerned authorities to provide needed mobility not only for the economic growth but also for its very existence.

As the government has been failing to keep up the development of infrastructures with the pace of rapid urbanization and growth, a realization has come that the public sector alone cannot provide needed infrastructures for sustained economic growth and development without the support of private sector. As such, Bangladesh government has been showing increasing interest in the development of transport infrastructures including light and heavy rails, urban and regional roads, deep-sea and land ports, inland water

transport terminals etc through the Public Private Partnership (PPP) programs (*Ministry of Finance, 2009*).

## FUNDAMENTAL'S OF PPP AND RELATED DEVELOPMENTS

As defined in the draft Bangladesh Public-Private Partnership Policy and Guidelines (PPPPG) (*Prime Minister's Office, 2009*): "A PPP is a contractual agreement between a public and private entity, whereby the private entity performs a part of government organization's service delivery functions, and assumes the associated risks for a significant period of time. In return the private entity receives a benefit/financial remuneration." Other definitions of PPP such as the one used by the Asian Development Bank (ADB) (*ADB, 2008*) is that "PPPs present a framework that—while engaging the private sector—acknowledge and structure the role for government in ensuring that social obligations are met and successful sector reforms and public investments achieved."

Regardless of how PPP is defined, the major purpose of PPP is to accelerate the development of infrastructures and services for optimal societal benefits by jointly sharing the risks and responsibilities, and providing finance and complementary skills needed between a public and one or more private partners at a least expensive manner. It is a long term business commitment between public sector and private investors by sharing financial, legal, technical and economic risks. A contractual agreement by clearly defining the roles and responsibilities between a government entity and one or more private entities mandates/facilitates such an opportunity to engage the private sector in the development of public infrastructures and services. Private sector generally shares the risks and responsibilities including financing, planning, managing, designing, building, owning, operating and/or transferring infrastructures for mutual benefits of all stakeholders including the users for an extended period (typically for concession type projects between 15 and 30 years).

To undertake PPP projects, first of all, there must be a policy under a legal framework. Secondly, an institutional framework as well as capacities/skills would be needed in the areas of technical, and administration and management of procurements and contracts of all forms of PPP projects. Right policy and institutional setting are critical not only to demonstrate the willingness of government to engage but also to build the confidence of private sector to invest. If executed properly, a PPP project should deliver a high quality service to the users as private sector quality management, innovation, technology and technical skills would complement public sector in the rapid implementation of the project.

A handful of literatures thoroughly discussed potential benefits of PPP projects. A report (*FHWA, 2007*) prepared for Office of Policy and Governmental Affairs of USDOT identified potential benefits of transport related PPP projects including (a) Additional Resources and Capacity, (b) Accelerated Project Delivery, (c) Reduced Costs and Increased Efficiency, (d) Transfer of Selected Risks to the Private Sector, (e) Greater Access to Technology and Innovation, and (f) Increased Accountability for Performance. There are also potential risks in implementing PPP projects such as public and political un-acceptance, mismanagement of programs and projects, financial and revenue risks, corruptions resulting from political and lobbying groups' interference, and unstable socio-economic and political states.

Throughout the world, PPPs are being widely used to expedite the development of transport infrastructures and services including toll roads, public transport (light and heavy rails), ports and intermodal terminals (land, sea and air) (*Lawther, 2000; Renda and Schrefler, 2006*) with a variety of financing and project delivery options. As identified by ADB (*ADB, 2008*), PPP projects could be of many forms including service contracts, management contracts, lease contracts, Design-build (DB), build–operate–transfer (BOT), Design-build-finance-operate (DBFO), build- own-operate (BOO), Build-Own-Operate-Transfer (BOOT), design-build-operate (DBO), rehabilitate-operate-transfer (ROT), concessions, and joint ventures. However, it is found that service contract (33.33%), BOT (26.67%) and concession type (17.14%) projects has been dominating in Asia.

World Bank data showed that total investments in the infrastructure sectors of developing nations with private sector participation were about $754 billion between 1990 and 2001 (*Quium, 2003*). The transport sector share was about 18 percent. In Asia, Transport sector investments were mainly in the seaport, airport, road and rail sectors. Most of those investments were attracted to China, Malaysia, the Republic of Korea, the Philippines, Indonesia, India and Thailand. Recent data showed that India is leading the developing nations in the flow of private investment (*WB, 2008a*). Another study reported that Asian countries alone planned and funded 137 projects with a total worth of $83.9 billion; and funded and completed 72 projects with a total worth of $44.5 billion between the years 1985-2004 on road related (road, bridge and tunnel) projects (*FHWA, 2007*).

## GOVERNMENT PLANS AND POLICIES FOR PPP DEVELOPMENTS

Bangladesh with a moderate GDP growth of 5.8% per annum during the last decade has been facing a steady resource gap in the range of 4 to 5% of GDP (*Shah, 2009*). Yet, she has ambitious plan to raise the GDP growth to 8% by 2014 with a projected investment deficit of US$28 billion between fiscal years 2010-2014 (*Ministry of Finance, 2009*). Government believes that the planned GDP target can be achieved through the development of PPPs in various sectors including Power and Energy, Transport Infrastructure, Pure Drinking Water and Sewerage, Information Technology, Air Transport and Tourism, Industry, Education and Research, Health and Family Welfare and Housing, etc (*Ministry of Finance, 2009*). However, current trends in infrastructure investments do not support that government can able to achieve a GDP growth of 8 percent by 2014.

PPP concept is relatively new in Bangladesh. In 1989, with the establishment of Board of Investment (BOI), Bangladesh Government has begun to facilitate private investment in the country. Thereafter, in 1996, a policy was adopted to promote private sector investment in the power/electricity generation. In 1997, Infrastructure Development Company Limited (IDCOL), a state owned nonbank infrastructure financial institute was formed to finance PPP projects. In 1999, Infrastructure Investment Facilitation Center (IIFC) was established under the Economic Relations Division (ERD) of the Ministry of Finance to assist government agencies with PPP proposal preparation and technical support. Later in 2004, Bangladesh Private Sector Infrastructure Guidelines (PSIG) were developed; however, there had still been lacking adequate legal, institutional

and policy supports, as well as financial plan on how projects under PPP program would be supported.

In 2009, a draft PPPPG was prepared by the government to stimulate the investment in the infrastructure sectors. It was for the first time that government addressed the importance of a legislative framework on how to enter into agreements with private sectors, outlined institutional frameworks and needs for management, procurement and technical skills development as well as capacity building in handling PPP program. It also addressed the need for a business policy including exit policy and dispute settlement mechanism. Emphasis was also given to attract local entrepreneurs to participate in the infrastructure projects. The formalization process of PPPPG through the approval of Cabinet and Parliament has been progressing now.

Under the current PPPPG, all PPP projects will be administered centrally. In this regard, a PPP cell as a wing of BOI has been established that will act as secretariat of Public-Private Infrastructure Committee (PPICOM), administer and implement business plans, and coordinate with various line ministries and executive agencies. A PPP advisory council (PPPAC) with Prime Minister the chair would advice on the overall policy reform and give guidance to executing agencies including line ministries, PPICOM and Cabinet Committee on Economic Affairs (CCEA). Planning Commission (PC), a wing of Ministry of Planning, will be responsible for creating project pipelines and Finance Division (FD) under the ministry of finance will be responsible for managing and administrating the PPP budget. In association with PC, PPICOM through PPP cell would play key role in the development of new laws, regulations and policies, as well as approving, funding, monitoring and evaluating PPP projects. The three key team players in the PPP development including PPICOM/BOI/PPP Cell, FD and PC are termed as "the PPP triangle".

As defined in the PPPPG, PPP project can evolve either through solicited or unsolicited proposals. Solicited projects are identified by government and unsolicited projects are initiated by private sectors for government evaluation.

## LOCAL EXPERIENCE, PPP FUND AND IMPLEMENTATION CAPABILITIES

Government has been given enormous emphasis on the development of PPPs in the energy, power, and transport sectors including ports to boost economic growth. However, there have been quite a few successes in implementing PPP projects mainly due to lack of comprehensive PPP policy supported by legal and institutional frameworks, skills and capabilities, and poor initiatives. Some successes were noticed in the power, telecom, energy and land-port development projects (*Ministry of Finance, 2009*). However, there remains a very slow progress in implementing PPP projects in the Transport sector. After failures of several attempts in the past, recently Dhaka City Corporation (DCC) has been signed an agreement to build Gulistan-Jatrabari flyover, a nine kilometers long and four-lane flyover in the Capital Dhaka, on a build-own-operate-transfer (BOOT) basis. Other transport related projects that are in the government's priority list include Dhaka-Chittagong access control highway, sky-train encompassing Dhaka, Dhaka underground metro rail, elevated expressways at and around Dhaka, the second Padma multipurpose bridge and deep sea port in Chittagong.

A separate PPP fund has been created under the Annual Development Program (ADP) in the last two annual budgets (e.g., 2009-10 and 20010-11 budgets). The primary purpose of such an allocation of fund was to assist in the development of PPP projects (Pre-feasibility assistance fund), and provide subsidies (viability gap fund) for projects those are strategically important but financially risky and long term financial supports (Infrastructure Investment fund) for the PPP projects. Another new fund, called Bangladesh Infrastructure Finance Fund (BIFF), has been appeared in the 2010-11 budget under the PPP program to meet the huge long term funding needs for PPP projects. However, due to slow progress in formalizing PPPPG as well as inadequate administrative and technical capabilities of executing bodies to implement PPP projects, the allocated $ 360 million (TK 2500 Crore) PPP fund in the 2009-10 budget has not been used.

PPPPG has addressed some of the past weaknesses such as legal, institutional and policy supports, as well as financial plan on how projects under PPP program would be supported; however, substantial work is still needed in the detailed developments of policies, implementation guidelines, and institutional capacity building for administering and managing PPP program. Although, the World Bank has been supported and continuing to support the government in this regard, however, there still remains weak capacity in formulating policies, administering procurements, negotiating contracts, and managing projects (*WB, 2002, 2008b* ). There is also shortage of technical skills in handling large and complex projects. Government credibility in combating corruptions and demonstrating transparency and accountability in the contract administration and project financial matter is also poor.

Bangladesh has made some progress though in standardizing public procurement process. In 2002, World Bank helped Bangladesh in standardizing public procurement process (*Ministry of Planning, 2002*) through the Public Procurement Reform Project. The same year, government has established Central Procurement Technical Unit (CPTU) under the ministry of planning for implementing public procurements. In 2003, Government has issued public procurement regulations. Later in 2006, the parliament has passed a public procurement act (PPA-2006) and, in 2008 issued a new set of public procurement rules (PPR-2008) (*Ministry of Planning, 2002*). However, government agencies have still been showing poor performance in managing procurements. The lack of adequate skills in handling PPP procurement has been hindering the success of PPP program. Other issues is that although competitive bidding is required on all projects involving multilateral agencies such as WB and ADB, however, not all projects funded by government has to meet such requirements. Past experience showed that unsolicited bid could be a source of corruption and also leads to a poor transparency and accountability. Political and lobbying groups' interference as well as unnecessary delay in public procurement process is also common phenomenon.

## RECOMMENDATIONS FOR A SUCCESSFUL PPP PROGRAM

As discussed previously, government agencies have been lacking skills needed to develop comprehensive policies, procedures and guidelines as well as managing and administering PPP projects. In this section, recommendations for a successful PPP program have been made in critical areas in light of PPPPG. These recommendations are

based on author's knowledge on the subject matter as well as extensive review of literatures as listed in the reference section.

**Legal and Regulatory Frameworks and Policy guidelines**

PPPPG conceptualizes the PPP program; however, detailed developments are yet to materialize. Therefore, a comprehensive PPP act by a parliamentary approval as well as guidelines for interpreting policies and procedures are needed in the successful executions of projects. Policies and procedures also need to be updated periodically through identification of gaps and weaknesses.

**Institutional Development and Capacity building**

PPICOM/PPP Cell, FD and PC (The PPP triangle) have to be given adequate administrative and budgetary powers as well as functional capabilities to operate as influence free, independent bodies to fulfill their respective obligations. All executing agencies and line ministries have to be empowered with technical and managerial capabilities. Administrating PPP programs would require highly specialized team with skills in different areas including but not limited to law, finance, business administration and management, and engineering. Through appropriate policies, government should focus on creating such experts in both public and private sectors. Government may seek technical assistance from international agencies such as ADB and WB as well as from qualified consultants to rapidly building PPP team.

**Creation of Project Pipelines**

PC in association with executing agencies must develop project identification and prioritization criteria through a comprehensive and continuous integrated transportation planning process to ensure best utilization of limited resources. For each project in the pipeline, clear technical and business scopes have to be developed showing tangible and intangible socio-economic benefits supported by valid data. This early information is very critical for both lenders and investors. An adequate fund under the PPP budget has to be allocated for early scoping and feasibility studies of all projects included in the pipeline. Projects that are self sustained and improve the quality of public life such as transit terminal with associated commercial and residential developments has to be considered to offset some of the financial risk involve in transportation investment by giving non transportation development. A comprehensive environmental law also has to be passed to justify that PPP projects supports sustainable development.

**Procurement and Contract**

The success of a PPP project greatly depends on the successful development and management of procurement and contractual agreement. Improper procurement may lead to a project failure or significant damage to it. Under the PPP program, management of procurement and contract are much more complex than that of traditional one, yet such a task is relatively new to the public administrators. Therefore, it is the critically needed skilled area that must be improved to translate the PPP policy into a reality.

Although, selecting a type of procurement depends on project needs, however, it is not recommended noncompetitive bidding resulting from an unsolicited proposal. If non-competitive bidding cannot be avoided, transparent criteria must be developed by

clearly stating under which circumstances such a bid could be acceptable while reducing the risk of corruption. Government may also follow examples of other countries to ensure transparent procurement and financial transaction such as Philippines in an effort to combat corruptions (*Dumol, 2000*).

Negotiation

The agreement must adequately address scope of business, roles and responsibilities, performance expectations, exit policies and conflict resolution mechanisms, and stakeholders (all parties including users) interests to avoid confusion. Again, it requires strengthening technical, commercial, legal and contractual, and negotiation skills. A poorly documented contract and unskilled negotiation could give excessive leverage and benefits to private sector beyond that is rationally deserved and appropriate, which could result in a political failure for mismanagement of public assets/funds. From that prospective, government responsibility and accountability is even greater with PPP project than that of traditional one.

Scoping

Past experience suggests that weak contractual document may create scoping crisis and force to re-scoping, renegotiate or even postponed the project, which could be very costly. Thus, government should learn lessons from their own as well as other countries mistakes, and develop a comprehensive contract document and conflict resolution mechanism (*Gómez-Ibáñez et al, 2004; Ke et al 2009*). It is also necessary to clearly define the goals and objectives of projects. Setting standard criterion to measure objectives and evaluate the performance towards achieving goals are important.

Risk Management and Conflict Resolution

Accurate prediction of project risks, costs, and benefits is critical for a successful implementation. It also requires clearly addressing exit policy upon completion of concession period or in the event of a serious disagreement during the concession period. There are many forms of risks that a project has to consider such as public acceptance, management, financial and project cost (cost overrun), delay in design and construction, uncertainty in profit, currency exchange rate as well as change of government and laws, political and civil unrest etc. Therefore, it requires developing standard tools for managing such risks as well as appropriate techniques for identifying and mitigating it. As one of the most vulnerable nations to climate change with a weak environmental policy, future environmental disaster and/or policy change may also increase the risk that government must address. Therefore, it requires keeping a provision on how to manage and share a risk in the event of a disaster such as earthquake, and flood.

PPP is not only a long term business agreement but also a long term commitment to serve public at a reasonable cost. Therefore, to avoid any future conflict between the government and private provider that may jeopardize public interest, contractual agreement must address how to resolve potential future conflicts in case of a new event. This could be explained by an example - if in the future, a parallel road project is implemented and found that it has diverted some of the traffic of the existing facility to

the new one, a provision to identify and negotiate how to fairly distribute the revenue of diverted traffic would be required.

**Socio-political Risk**

Another barrier in attracting such a long term private investment is the weak political system and culture. It is often found that the decisions and commitments of outgoing administration are not honored by incoming administration. Therefore, there is a risk of litigation if a contractual agreement is not supported by legal requirements. Appropriate legislation in support of a contractual agreement is required to ensure that legal commitment made by one administration would be an obligation to honor by another. Ensuring sustained political and economic stability, that is yet to establish, is also required to create investor's confidence.

**Business support**

By creating a favorable business environment such as level playing ground for all participants in the procurement process, government can able to create investors confidence and thus, competition. Providing efficient services in the approval process for permits and licenses are also required.

Finance

Government should assist investors to get long term low interest loans for financing PPP projects on a priority basis through BIFF, IDCOL and other sources. As BIFF, and IDCOL funds are very limited and all projects may not be qualified for the funds, financing projects through local, bilateral or multilateral banks, as well as nonbank sources such as Non-resident Bangladeshi (NRB) remittance funds (a long term investment fund that may be generated by attracting NRBs in a similar concept of foreign currency account) and long term bond funds (long term institutional investors such as insurance companies and pension fund managers may be targeted for this fund) has to be considered. However, to take advantage of NRB remittance and bond funds, government has to develop legal frameworks.

Incentives

Under the provision of viability gap fund, projects that are not commercially viable but economically justifiable, government considers providing subsidy to cover a reasonable amount of project cost. However, how such projects will be identified and the amount to be subsidized has to be formulated rationally. Exemption of tax and duty free import of related machineries and equipments should be allowed for an extended period.

In addition to viability gap fund, minimum rate of return guaranty on strategic projects, and profit sharing opportunities between private and government sectors on highly profitable projects should also be considered. Terms and conditions for rate of return guaranty and profit sharing should be formulated carefully so that benefits and risks are evenly distributed. Government may also explore the possibility of getting climate change and related funds from international sources as grants and aids that could be used to provide subsidies for PPP projects that are potential for improving air quality and saving energy.

Risk Allocation

Government should consider sharing financial risk on projects that appears unsustainable from business prospective but potential for promoting socioeconomic and sustainable developments. Such a risk could be to take over the facility on a fixed price basis or providing subsidy given the project fails to generate profit for an extended period of time from the revenue generated through user charge/toll/price etc.

**Public Participation and Stakeholder Communication**

For successful implementation of projects, public support is critical. However, there is no formal provision in the PPP guidelines about such participation. Government should take initiatives to conduct brainstorming and special sessions to receive ideas and inputs from the public as well as from private investors. Such an effort will create strong communication channels between stakeholders. Involving citizens including business leaders, civil society, policy experts, media and other appropriate focus group early in the project development stage would not only help identifying issues and needs, but also help getting their supports for the project.

**Project Delivery Options**

Under the PPP program, the innovative Design-Build (DB) contract as oppose to traditional Design-Bid-Build (DBD) contract likely will reduce project delivery time and cost as design and build phases can progress concurrently under a single contract. Such a project can be managed efficiently while minimizing constructability/technical risks.

**Service quality and User Charge**

Social welfare cannot be maximized if private sector gets the monopoly and impose excessive charge to users. In this regard, government has yet to develop a rational formula and mechanism to determine user charge/toll. A rational approach has to be followed in determining tolls (user charge) to be charged at a given time during the concession period. Government may also consider shadow toll to relieve pricing burden on users. It also requires ensuring minimum service quality that may be measured by setting service evaluation criteria.

## CONCLUSIONS

This paper addresses the prospect of PPP development in Bangladesh in light of governments' recent initiatives and practices. The weaknesses in current practices are identified and recommendations in the critical areas are made for sustainable PPP development. It is found critical to develop institutional capabilities in formulating policies and procedures, managing and administering contracts, negotiating and conflict resolutions, and implementing PPP projects. In addition, project pipelines have to be created with clearly defining technical and business scopes, and opportunities of each project to attract investors. Standardizations of procurement rules and processes, and project delivery options are necessary for efficient implementation of projects and curbing corruptions. Standard mechanisms and techniques have to be developed to allocate and manage project risks. Government should also seek for innovative financing options such as NRB remittance funds and long term bond funds to support PPP projects.

It is expected that with proper policy, planning and incentive offers, government can able to take the opportunity to accelerate GDP and sustained economic growths by rapidly expanding needed infrastructures and services. By best utilizing the lessons learned from local and international experiences, government may further strengthen PPP program provided such lessons are incorporated through appropriate amendments to policies, procedures, and regulations.

## REFERENCES

ADB (2008). "Public-Private Partnership (PPP) Handbook." *Asian Development Bank*, Mandaluyong City, Philippines

Dumol, M. (2000). "The Manila Water Concession: A Key Government Official's Diary of the World's Largest Water Privatization." *The World Bank*, Washington, D.C.

FHWA (2007). "User Guidebook on Implementing Public-Private Partnerships for Transportation Infrastructure Projects in the United States." Final Report, *Office of Policy and Governmental Affairs*, FHWA, USDOT

Gómez-Ibáñez A., Lorrain D., and Osius M. (2004). "The Future of Private Infrastructure." *Working Paper,* Harvard University, Cambridge, Massachusetts

Ke Y., Wang S., Chan A. (2009). "Public-Private Partnerships in China's Infrastructure Development: Lessons Learnt." *Proc. of Int'l Conf. on Changing Roles: New Roles and New Challenges*, TU Delft, Faculty of AREH, The Netherlands, pp. 177-188

Lawther W.C. (2000). "Privatizing Toll Roads-A Public-Private Partnership." *Praeger Publishers*, USA

Ministry of Finance (2009). "Invigorating Investment Initiative Through Public Private Partnership." *A Position Paper, Finance Division, Ministry of Finance*, Government of the Peoples' Republic of Bangladesh

Ministry of Planning (2002). "Central Procurement Technical Unit," IMED, Ministry of Planning, *Government of the People's Republic of Bangladesh* (accessed, June 13, 2010, http://www.cptu.gov.bd/default.aspx)

Prime Minister's Office (2009). "Draft Bangladesh Public-Private Partnership Policy and Guidelines." *Government of the People's Republic of Bangladesh*, Dhaka

Quium A. (2003). "Private Sector Participation in the Transport Sector: Trends, Issues and Institutions in the Asia-Pacific Region." *Transport and Communications Bulletin for Asia and the Pacific*, No. 72

Renda A. and Schrefler L. (2006). "Public-Private Partnerships: Models and Trends in the European Union." *European Parliament*, Brussels

Shah S. (2009). "Bangladesh Financial Sector: An Agenda for Further Reforms." *Asian Development Bank*, Mandaluyong City, Philippines

WB (2002). "Bangladesh Country Procurement Assessment Report." *Report No. 24144-BD*, The World Bank, Washington, D.C.

WB (2008a). "Public-Private Infrastructure Advisory Facility (PPIAF)." *Gridlines Note No. 30*, The World Bank, Washington, D.C.

WB (2008b). "Implementation Completion and Results Report: Public Procurement Reform Project for the People's Republic of Bangladesh." *Procurement Services Unit*, South Asia Regional Office, The World Bank

# The Role of Access Management in Sustainable Development

Kristine M. Williams, AICP[1] and Herbert Levinson, PE, NAE[2]

[1] Program Director, Planning & Corridor Management, Center for Urban Transportation Research, University of South Florida, 4202 E Fowler Avenue, CUT100, Tampa, Florida 33620-5375; PH 813-974-9807; FAX 813-974-5168; email: kwilliams@cutr.usf.edu
[2] Icon Mentor, Region 2 University Transportation Research Center, City College, New York and Transportation Consultant, New Haven, Connecticut; PH 203-949-9700; email: hslevinson@aol.com

## Abstract
Efforts to manage highway access are occasionally characterized as inconsistent with livable communities, complete streets and smart growth. One reason is the past emphasis placed on management of vehicular access, which some view as symptomatic of auto-oriented planning and therefore counter to multimodal goals. Another reason is a lack of understanding of the role of access management in advancing sustainable development and supporting walking, bicycling and transit use. This paper explores the role of access management in accomplishing a more sustainable approach to transportation and development planning in the US. Topics include land use and transportation relationships, network planning and activity center strategies, and site design concepts for urban arterial development. Policy changes are suggested.

## Introduction
Access management involves the classification of roadways according to function, application of suitable spacing and design standards by roadway type, and codifying these standards in state and local access codes. State and local governments use access management policies to preserve the functionality of their roadway systems. The policies are implemented through access/development permitting, design standards, and the roadway improvement process. Systemwide access management policies are sometimes complemented by the development of corridor management plans to integrate land use and transportation goals and policies along a major transportation corridor. Guidance is contained in the TRB *Access Management Manual* (TRB 2003) and in numerous federal and state publications, National

Cooperative Highway Research Program reports, and state and local government access management codes.

State and local governments use a variety of techniques to control access to major arterials and other roadways important to state and regional mobility and freight movement. The techniques are designed to reduce travel delay, manage congestion, and minimize the potential for crashes. They include improvements to benefit transit, pedestrians, and bicyclists, as well as different strategies and treatments for urban, suburban, and rural settings. Examples of access management techniques and practices include:

- Increasing spacing between traffic signals to maintain efficient progression;
- Optimizing driveway location, spacing, and design;
- Use of exclusive turning lanes so drivers can wait safely to complete a turn and turning vehicles do not delay through traffic movement;
- Median treatments to reduce left-turn conflicts and guide motorists to safe locations for left turns and U-turns;
- Careful management of driveway access in the vicinity of roundabouts for driver and pedestrian safety;
- Use of service roads and network enhancements to reduce the need for direct driveway access to major arterials;
- Land use policies that promote improved site design and internal connections between adjacent businesses;
- Land use policies that increase non-vehicular access to and across major roadways, such as bicycle/pedestrian only connections between developed areas and to transit stops and midblock crossings at high pedestrian crossing locations; and
- Land use policies that increase local network connectivity to allow vehicles to circulate within corridor activity centers rather than on the highway system.

Despite a broad range of potential policy options and techniques, efforts to manage highway access are occasionally characterized as inconsistent with livable communities, complete streets and smart growth. One reason is the emphasis on management of vehicular access, with less focus on walkability and public transportation. Arterials flanked by fields of parking remain unattractive, and wide multi-lane access points are difficult for pedestrians to use. Another reason is a lack of understanding of the role of access management in advancing sustainable development and walking, bicycling and transit use.

This paper explores the role of access management in accomplishing a more sustainable approach to transportation and development planning in the US. Topics include land use and transportation relationships, network planning and functional hierarchy, and site planning and design concepts. Policy changes are suggested.

**Understanding Access, Accessibility and Mobility**

Three key terms establish the relationship between access management and sustainable development – access, accessibility, and mobility. They have been defined in various ways depending upon the context in which they are applied. The definitions in Table 1 are useful in understanding how to integrate land use and transportation.

Table 1: Definitions of Access, Accessibility and Mobility

| | |
|---|---|
| Access | The ability to enter and exit a property via the transportation network. |
| Accessibility | An area-wide measure of the ease of travel between locations within a defined geographic area (e.g. is the ability to reach a given location from numerous other locations, or the ability to reach a variety of other locations from a given location.) |
| Mobility | The ability of people to make trips to satisfy their needs or desires by various modes of transportation conveniently and within a reasonable amount of time. |
| Source: Adapted from National Highway Institute, *Access Management, Location, and Design*, 1998 | |

Exploring how these terms interrelate within the built environment provides insight into the role of access management in a sustainable system. For example, frequent driveway access along a highway reduces accessibility of a major activity center, as traffic entering and exiting individual sites conflicts with traffic headed to the center. Accessibility at the activity center site may be further constrained by the inability to reach the center from the surrounding neighborhood without driving onto the highway.

Alternatively, access management would reduce driveway access along arterial routes and promote supporting local networks in developed areas. Networks within and connecting to activity centers could include streets, sidewalks, bicycle lanes and off-street, multi-use paths. This scenario would improve mobility, accessibility and access to the center in relation to both the highway and the surrounding area. People and goods could access the activity center via alternative paths and circulate easily and conveniently using a variety of modes. Arterial traffic conflicts would be reduced, leading to improved arterial safety and operations and preserving regional accessibility. The activity center areas could also be more efficiently served by transit–both express transit service and local transit circulators.

Figure 1 illustrates two scenarios that exemplify the issues discussed above. The top half of Figure 1 (left to right) demonstrates how separating land uses into stand alone developments with disconnected local networks increases local traffic circulation on the arterial system and can result in more arterial conflict (access) points. The bottom half of Figure 1 shows how local traffic and conflicts on the arterial are reduced when land uses are organized into activity centers with a unified street network. The bottom example is also more conducive to walking, bicycling and transit use.

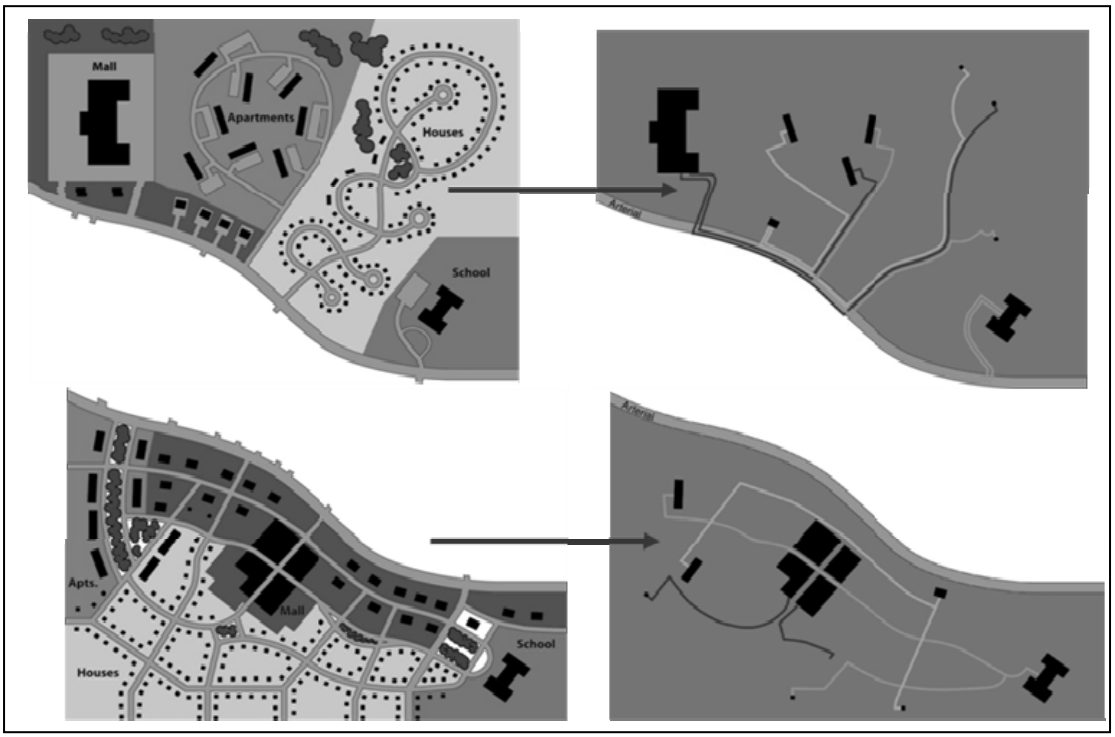



Figure 1: Network Connectivity, Access and Arterial Traffic. Source: Duany Plater-Zyberk and Company as supplemented by Glatting Jackson Kercher Anglin, Inc. (now AECOM).

**The Importance of Centers**

Implicit in the coordination of transportation and land use is recognition that every community needs an operational center that is linked with other parts of the community. Centers can occur at varying scales from the major urban core or downtown, to a district shopping center or neighborhood center. Neighborhoods that include a greater mix of land uses within reasonable proximity not only have greater choice of travel alternatives, they also afford residents greater convenience in meeting daily needs. The availability of convenient travel alternatives translates into a higher quality of life and as such, can be viewed as a key component of livability.

Ideally, central business districts and other major urban activity centers should be highly accessible both regionally and locally via a variety of transportation modes and multiple paths. These paths could include:

a) Freeways, expressways, and other access-controlled major arterial highways, along with regional transit service (e.g. commuter rail, rail rapid transit, light rail, bus rapid transit) to support regional mobility between major activity centers and key points in the community;

b) Regularly spaced arterial and major collector roadways, complemented by local transit service to support mobility within and across urbanized areas; and

c) A dense, connected network of minor collector and local streets, alleys, multi-use paths, sidewalks, and user facilities (e.g. bicycle racks, benches, water fountains, etc.); to support neighborhood mobility within and between local activity centers and surrounding residential areas.

Major activity centers should be located where they can be readily accessible to transit. Zoning ordinances with sustainability in mind could require transit oriented developments as well as large office and institutional developments (e.g. over 300,000 square feet) to locate within 400 to 600 feet of an express transit stop. Zoning also could limit the amount of commercial land and encourage this land to be clustered at key nodal points to reinforce activity center development patterns.

Alternatively, the tendency to develop large residential subdivisions at the urban fringe, and to focus goods and services onto strips along arterial highways forces residents to make more auto trips, longer trips, and focuses local traffic onto the arterial system. These development decisions preclude transit and walking and generally make travel less convenient. People may have to drive even where they live within walking distance of their destination. This pattern of development is not sustainable. It magnifies demand on the arterial system and increases the need for continuous roadway improvements as discussed below.

**The Cycle of Highway Obsolescence**

Highways cause decentralization of metropolitan areas even as fringe development places new pressures on the highway system. Research by Cervero, for example, demonstrates that freeway improvements induce growth and investment along these corridors, as "real estate development gravitates to improved freeways, and traffic increases spawn road investments over time (Cervero 2003)."

Stover and Koepke (2008) note that land use changes stimulated by the increased accessibility created by a new highway, result in a cycle of highway obsolescence as poorly managed development access results in increased traffic conflicts, crashes, and congestion (Boarnet 2008). The highway must then be widened or reconstructed to restore its safety and capacity at considerable cost and disruption to the public and the environment. This cycle of highway obsolescence is anything but sustainable.

A more sustainable approach would include focusing development into activity centers and around core areas, rather than dispersing it, and engaging in more effective management of the existing system through a variety of techniques that include access management. Boarnet (2008) further advises planners to place less priority on preventing future congestion through improvements to fringe highways that induce additional exurban growth, and higher priority on congestion relief measures, strategic improvements at bottlenecks, and demand management. The emphasis in core areas or activity centers should be less on relieving congestion - a sign of vitality - than on expanding and reinforcing alternative modes, improving walkability through dense and connected networks, and promoting a diverse and compatible mix of land uses.

**Functional Hierarchy, Complete Streets and Context Sensitive Design**

The foundation of access management is a functional hierarchy based on managing mobility and access. In the context of urban areas and complete streets this planning concept remains relevant – although the nature of street types will continue to evolve, as it has throughout our nation's history.

Transportation plans for complete streets will need to maintain a functional hierarchy of design types based on desired operating speeds, carrying capacity and public safety. This will involve different approaches to access design on major corridors intended for longer distance, higher speed travel than on those where local circulation is a priority. The design of complete streets will differ depending on their context and the modes expected on a corridor. Arterial street cross sections should fit within the available right of way. They should be "complete" in that they consider the needs of motorists, pedestrians, and as appropriate cyclists.

Access management is one element of roadway design and as such it must also be sensitive to context. Some guidelines are suggested below:

- Access management techniques, such as medians and driveway controls, significantly increase pedestrian and bicycle safety and enhance corridor aesthetics. Street medians should be wide enough to provide adequate pedestrian refuge. Painted medians are adequate, although not ideal, on 4-lane streets in urban settings. Elsewhere, and along six-lane arterials, physical medians are desirable; they should be at least 6 feet wide to provide an adequate refuge for pedestrians.
- Ten-foot turning lanes are adequate in dense urban settings, elsewhere 12-foot lanes are desirable.
- Bicycle lanes should be provided within the street cross section, but on suburban corridors with frequent driveway access bicycle lanes are often unsafe. Multi-use paths could be provided to connect subdivisions and activity center areas.
- From a pedestrian standpoint, 4-lane streets are preferred as they are easier to cross and constitute less of a barrier to pedestrians.
- Sidewalks should be provided in urban and suburban settings and developments should be required to connect to the sidewalk if buildings are not fronting directly on the street. Provision for future sidewalks is desirable in exurban areas.
- Uniform signal spacing of ½ mile on arterials offers many benefits in terms of progression and flexibility to accommodate differing volumes of traffic over time, but in urban environments it is often difficult to achieve. As population densities increase, street grids become denser and therefore spacing of major junctions is reduced. Therefore, somewhat shorter spacings of signalized intersections and shorter cycle lengths may be required. Cycle lengths will depend on the type of street and the type of environment. Typical progressive speeds are 20 mph in the central business district, 25-30 mph in urban settings, and 30-40 mph in suburban and exurban environments.

- Avoiding very long signal cycles in urban and suburban areas will benefit pedestrians, as well as auto and bus traffic. Obviously, in some situations longer cycle lengths will be necessary.
- Curb-lanes may be used for parking during off-peak periods.

As more modal options are integrated into the transportation system, it will be increasingly important to carefully manage and design access so as to minimize conflicts between modes. The practice of access management will likely evolve beyond the current focus on balancing through traffic and development access, to more comprehensively balance the access interface between auto traffic, the pedestrian, and various forms of public transportation.

**Site Planning Concepts**

On a site planning level much can also be done to reconfigure how access is managed. An important aspect of "sustainability" is to translate motorized trips into walking trips. This can be achieved in two basic ways. (1) Creating multi-use activity centers that integrate retail, office, residential, and recreational areas and (2) placing buildings along streets to create a better sense of place, and to improve both pedestrian and bus access. In both cases, the clustering of activities can result in fewer, carefully designed access points, reduce vehicle trips between proximate activities, and encourage pedestrian and transit trips. Arrangements, located alongside streets reduce walking distance (as in crossing streets) and make it easier to board buses.

Figure 2 shows how building footprints can be transposed, with the buildings located close to the streets, and the parking placed in the rear. This transposition achieves several important objectives: (a) the rearrangement gives a "village" look to the developments, (b) the building groups on each side of both streets are within easy walking distance of each other, (c) buses operating along both streets can conveniently serve the various buildings and (d) vehicle storage space can be improved.

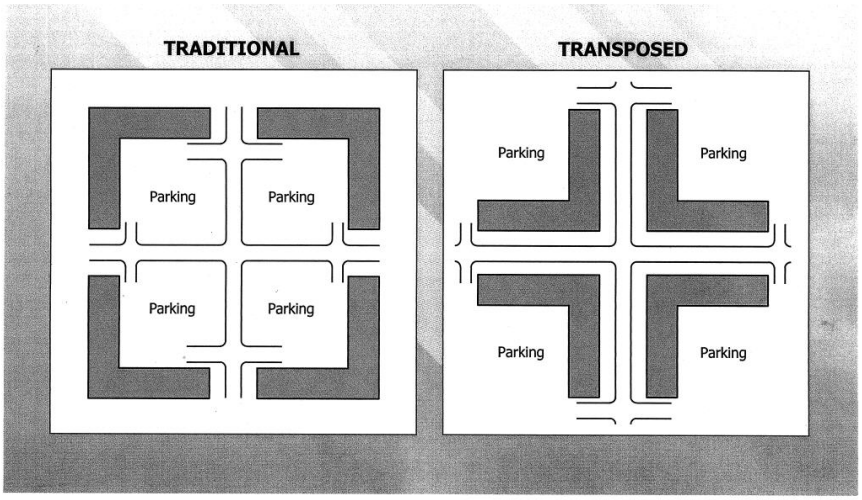



Figure 2: Modifying strip development to improve transit access.

Figure 3 shows how a large mixed-use development can be made more transit and pedestrian friendly. A major off-street transit-way with a centrally located station penetrates the heart of the development. A system of pedestrian-ways links the developments with the transit station, surrounding streets, and outlying commercial development along an arterial street. A landscaped buffer separates the development from the nearby freeway.

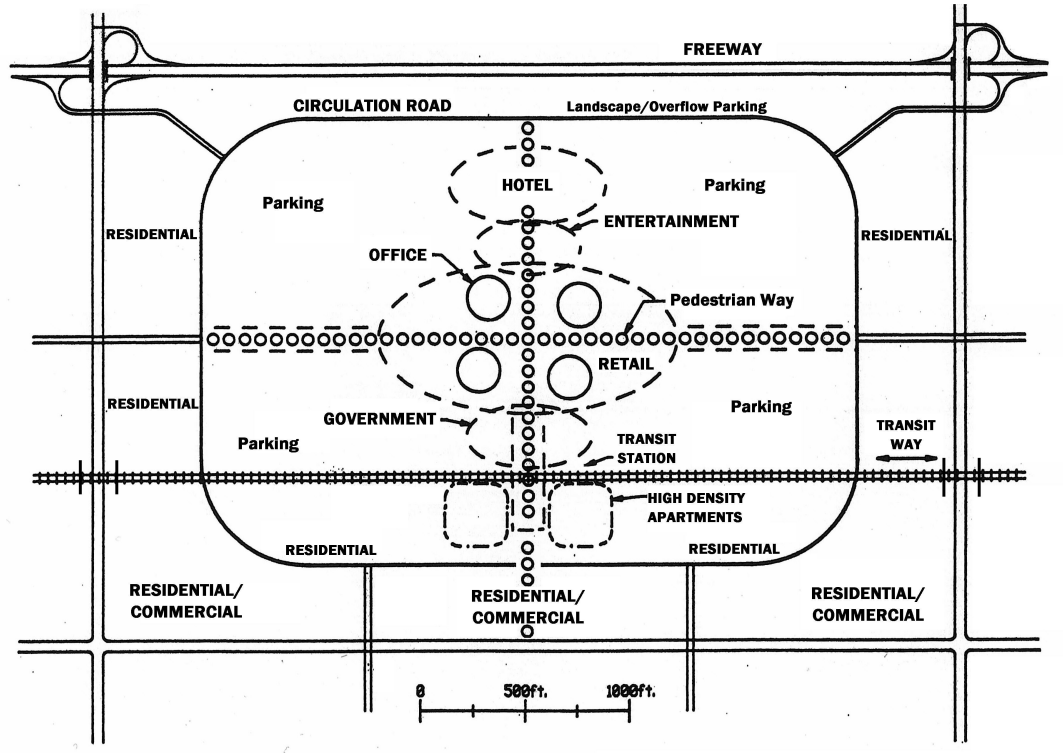



Figure 3: Modifying mixed use development to support transit and pedestrians. Source: Adapted from NCHRP Report 348 (Koepke and Levinson 1992).

Figure 4 shows how a typical auto-oriented development can accommodate commercial frontage along an arterial roadway. The first westbound access point from the artery only allows a left turn entrance; the second access point permits all movements, but could be simplified to prohibit the left turn entering movement.

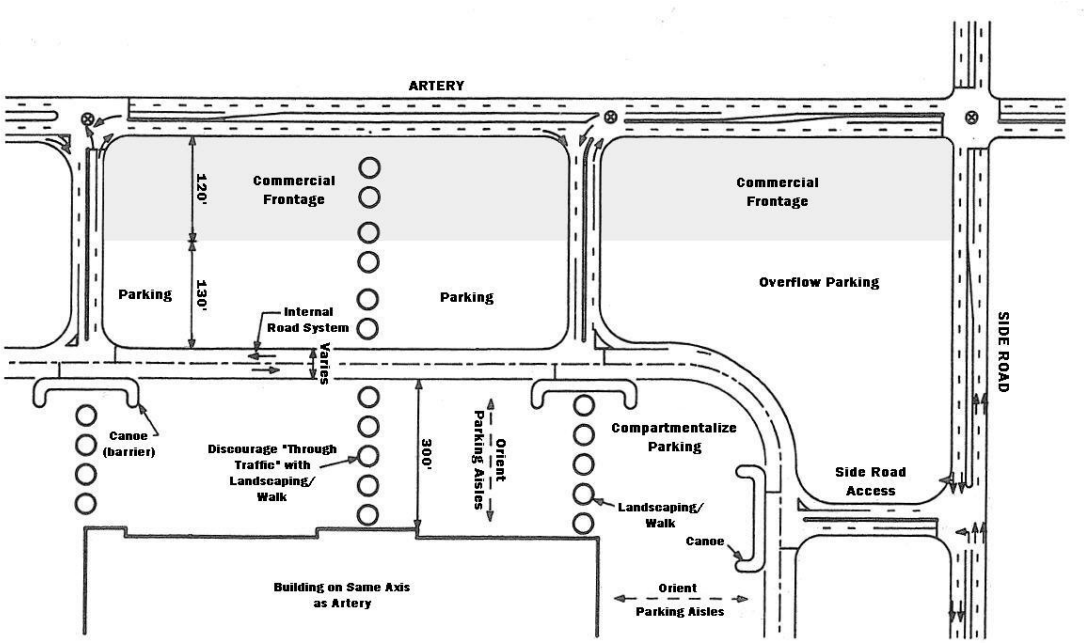



Figure 4: Retrofitting auto-oriented development to accommodate commercial frontage. Source: Adapted from NCHRP Report 348 (Koepke and Levinson 1992).

**Conclusion**

Access management provides a framework for coordinating transportation and land use planning and decision making. However, it cannot fully effect smart growth and sustainable development - they must be carefully integrated with other complementary strategies in local government plans and regulations.

Several important policy actions are needed to foster sustainable mobility and development. State, county, and local agencies should develop access management codes. The focus should be on better corridor management. Corridor overlay zones provide an important way to retrofit existing roads. Transportation impact assessment procedures should be multi-modal. They should address three basic concerns. Can people reach developments conveniently and safely on foot, by public transport and by car. In urban settings, access management can be complemented with form-based codes and buildings fronting on roadways with sidewalks to improve walkability, as well as the overall roadside image.

State transportation agencies commonly note that a key impediment to access management is the lack of authority to require local governments to implement the necessary corridor land use and network development solutions. This separation of powers constrains the effectiveness of state access management programs in advancing sustainable development, but it is not a shortcoming of access management per se. State transportation agencies should, however, more directly advance and reinforce sustainable development through state access management policies, context sensitive design, and comprehensive corridor management planning. Local governments, which have authority to undertake land use and network planning along highway corridors, must in turn more effectively implement these changes.

Creating sustainable transportation corridors in the years ahead calls for a bolder and broader vision of mobility and livability. It will require new perspectives, policies, and controls that address the complementary needs of transportation and land development. Investing in automobile infrastructure and in alternatives to the automobile are clearly both essential. The point is not highways versus other alternatives, but rather to focus growth into areas that can support modal alternatives and avoid adverse development impacts on highways that provide regional mobility.

**References**

Boarnet, M. (2008). "Transportation Infrastructure and Sustainable Development: New Planning Approaches for Urban Growth," *Access*, Fall, No. 33, 27-33.

Cervero, R. (2003). "Road Expansion, Urban Growth and Induced Travel – A Path Analysis," *Journal of the American Planning Association*, Spring, 69(2), 145-163.

Koepke, F., and Levinson, H. (1992). *NCHRP Report 348: Access Management Guidelines for Activity Centers*, Transportation Research Board.

Stover, V.G., and Koepke, F. (2008). *Transportation and Land Development*, 2nd edition, Institute of Transportation Engineers.

Transportation Research Board (TRB). (2003). *TRB Access Management Manual*, Committee on Access Management.

# Sustainable Design Guideline Development for Infrastructure Projects

Susanne DesRoches, LEED AP[1] and Philip Cremin, P.E., M. ASCE[2]

[1]Sustainable Design Manager, Port Authority of NY & NJ, 2 Gateway Center – 16th Floor, Newark, NJ 07102
Tel: (973) 792-3559; FAX (973) 792-4303; Email: sdesroches@panynj.gov
[2]Assistant Chief Civil Engineer, Port Authority of NY & NJ, 2 Gateway Center – 16th Floor, Newark, NJ 07102
Tel: (973) 792-4326; FAX (973) 792-4303; Email: pcremin@panynj.gov

## Abstract

The Port Authority of NY & NJ (PANYNJ) is a bi-state agency that builds, operates, and maintains infrastructure critical to the New York/New Jersey region's trade and transportation network. These facilities include America's busiest airport system, marine terminals and ports, the PATH rail transit system, six tunnels and bridges between New York and New Jersey, the Port Authority Bus Terminal in Manhattan, and the World Trade Center. For more than eight decades, the Port Authority has worked to improve the quality of life for the more than 17 million people who live and work in New York and New Jersey—a region that supports 8.6 million jobs with an estimated gross regional product of more than $929 billion dollars. Part of working to improve the quality of life for those who live, work, or visit our region is to build, operate, and maintain PANYNJ facilities in a sustainable fashion.

As regional transportation agencies work towards meeting sustainability and environmental goals, the US Green Building Council's Leadership in Energy and Environmental Design (LEED®) has become the most common sustainable design standard for facility design, construction, and operations. The LEED model functions well for building-centric projects, but transportation agencies often build, operate and maintain a variety of facility types, including roads, subways, railroads, ports, and airports. Since 2007, the PANYNJ has had Sustainable Design Guidelines with performance metrics that align with those of the LEED Green Building Rating System. These have worked well in guiding staff in new building construction and renovation projects, but they have not provided adequate sustainable design guidance to the majority of the PANYNJ's infrastructure projects. Recognizing a need for more comprehensive guidelines, the PANYNJ's Engineering Department devised additional guidelines specifically addressing sustainable design for transportation infrastructure. This paper presents an overview of the development of these guidelines, with particular emphasis on the case studies examined.

## Introduction

The year-long guideline development process involved an interdisciplinary and cross-agency team. Because the guidelines would need to take into consideration the variety of facilities and projects the PANYNJ builds and operates, the members of this team were selected from across all disciplines and facilities to ensure that all user perspectives were represented. The team relied on agency initiatives as case studies to develop sustainable design strategies for site design, water and energy conservation, material resources, construction environment and landscape maintenance practices.

## Development Process Utilizing PANYNJ Case Studies

Guideline development began formally in April 2009 with a day-long charrette involving an interdisciplinary and cross-agency team. A matrix of sustainable design strategies from outside agency guidelines was used as a tool for guiding the discussion at the day-long charrette. These strategies were compiled into similar topic areas analyzed for possible use in the guidelines.

Many PANYNJ infrastructure projects using sustainable design initiatives were already underway. These projects were instrumental both in the initial development of the guidelines and as standalone case studies in the final document. By including existing projects as case studies in the guidelines, not only would these initiatives be documented at the project level, but implementing these initiatives in future projects would be easier, since staff would have them as models and reference.

To identify possible case studies, the team evaluated all existing PANYNJ contracts, looking for those that incorporated sustainable infrastructure initiatives. The team also identified several potential initiatives that were not currently in use but that could be incorporated into upcoming projects as test cases to determine their potential applicability to PANYNJ projects. These initiatives and their applicability to the development of the guidelines are described in detail here.

## Case Studies: Projects Incorporating Sustainable Design Initiatives

### Relocation of Redneck Avenue

Figure 1: Plan View of Existing Redneck Avenue Roadway Alignment at Teterboro Airport, Teterboro, New Jersey

The FAA mandated that the PANYNJ implement improvements to the runway safety areas at all PA airports. One improvement was to install an Materials Arresting System (EMAS) at the south end of Runway 1-19 at Teterboro Airport in Bergen County, NJ. The site was located in an environmentally sensitive area, a floodplain with significant wetland areas. In order to make room for the arrestor bed, an existing county road and major utilities had to be relocated. The New Jersey Department of Environmental Protection had rejected the initial roadway relocation design because it would have significantly impact wetlands.

The PANYNJ recognized that an integrated team approach involving all stakeholders was needed to create an environmentally sensitive design that would also satisfy the FAA. The PA called a meeting at the airport involving all facility staff, the FAA, the NJDEP, the utility companies, the Army Corps of Engineers, and the PANYNJ engineering department. A consensus was reached on how to proceed to meet the mandated safety improvements while minimizing the environmental impact on the existing wetlands.

A new roadway alignment was developed that incorporated an S-curve with super elevation to avoid the forested wetlands. Steel sheeting was introduced to minimize impact on the affected area by reducing the grading footprint. Extensive cut and fill studies were performed to minimize the amount of net fill needed and to reduce the amount of material that would need to be trucked in. In this design, existing soil was reused on-site to the maximum extent possible and extensive stormwater best practices were introduced, including underground detention to alleviate downstream impact. An earthen berm was removed and replaced with a green wall—the first such application at a PANYNJ facility. This revised design met all FAA and Bergen County criteria and resulted in a 50% reduction in the amount of wetland acreage impacted from the original design. This project served as an important case study in the guidelines because it showed how an integrated team approach led to an environmentally sensitive design and successful project implementation.

**JFK Airport Rehabilitation of Runway 13R-31L**

Figure 2: Concrete Pavement on Runway 13R-31L John F. Kennedy International Airport

At nearly three miles long, runway 13R-31L (the "Bay Runway") at John F. Kennedy International Airport is the second longest runway in the country. This runway was over sixty years old and in need of major rehabilitation. In 2008, a study was conducted, including a full life-cycle cost analysis, to evaluate both asphalt and concrete pavements options. Extensive consideration was given to the facility impacts over the life of the pavement and the closures that would be required to accommodate each option. Concrete was selected based on the life-cycle analysis. While it had a higher up-front cost impact and would require a full, extended closure of the runway, the concrete pavement should not require major rehabilitation for at least forty years. This reduces not only facility disruption but also the need for new asphalt. This strategy has the added environmental benefit of increasing the solar reflectivity over approximately seventy acres (See Figure 2.).

Numerous sustainable initiatives were incorporated into the runway design: An on-site concrete plant was used to eliminate the need to truck in concrete. The existing base pavement of the runway was incorporated into the design to reduce significantly both the amount of pavement that needed to be removed and the time required for construction. The top six inches of asphalt were removed and reused for shoulder and erosion pavements as well as for a haul road. An infiltration trench was included to reduce the amount of water collected in the storm drainage system, and recycled HDPE was used for the subdrain system.

After the design was completed, the grading of the infield areas was evaluated using AutoCad Civil 3D. Based on this evaluation, changes were made that resulted in a

cost savings of $2 million and a reduction of 30,000 cubic yards of material that would have to be disposed of off-site.

This project provided several strategies included in the guidelines, such as balancing earthwork, using recycled materials on-site, using durable materials to enhance pavement life-cycles, and using materials to mitigate the heat island effect.

**Stewart International Airport Pervious Asphalt Parking Lot**

Figure 3: Permeable Asphalt at the Stewart International Airport, Orange County, New York

Recently, the PANYNJ constructed one new parking lot and significantly expanded another at Stewart International Airport using pervious asphalt pavement (Figure 3). The total paved area is approximately ten acres. These parking lots were designed for 100% rainwater infiltration: The water drains directly into the subgrade, not the storm drainage system. Some of the water is also collected and sent to a storage area for the irrigation purposes. The irrigation system also uses solar powered pumps. Strategies used in these parking lot designs were incorporated into several credits in the guidelines including full stormwater infiltration on site (called *rainwater neutrality*[i]), collection of rainwater to reduce the amount of potable water used for irrigation, and include on-site renewable energy systems.

**Stewart International Airport Sanitary Sewer Rehabilitation**

The existing sanitary sewer system at Stewart Airport exhibited significant groundwater infiltration and was overloading the town's treatment plant. To correct these problems, a contract was put together to replace most of the system. Pipe bursting technology and cured in place liner systems were used to minimize disruption to the facility. This project was used in the infrastructure guidelines as a

case study for implementing trenchless technologies, which minimize environmental impacts and lessen facility disruption when replacing subsurface utilities.

## Test Cases: New Sustainable Design Initiatives

As part of the guideline development process, several initiatives were identified that were not currently being used by the agency. These included the use of warm mix asphalt (WMA) and recycled asphalt. These initiatives were then field tested in pilot projects to determine whether and how they might be used in other PANYNJ projects. Warm mix asphalt (WMA) was specified on three contracts:

1. Parking Lot P6 at Newark Liberty International Airport
2. An administration building parking lot also at Newark Liberty
3. A major rehabilitation of the restricted service road at JFK.

A related initiative—use of pervious concrete—was also identified as an initiative that needed to be tested before being incorporated into the guidelines.

## Final Guideline Development

Through a six-month iterative vetting process, a total of forty-nine strategies were developed (See Exhibit 1). This initial list of strategies was developed into credits, each having its own requirements with a range of possible points and required documentation. The credits and the credit requirements were evaluated to determine their applicability to PANYNJ projects. Certain credits were dropped or merged with other credits. In addition, some new credits were added based on further investigation of the latest trends in the industry.

With this list of credits finalized, an implementation process was created to determine how each strategy would be applied on a project-by-project basis. An analysis was conducted on all contracts put out to bid over the previous two-year period. All of these contracts were classified as either infrastructure or building projects. Infrastructure projects were then broken down into fourteen types plus three supplemental categories. This classification system will enabled staff to focus on only the credits that were applicable to a specific project type.

Once the fourteen infrastructure project types were identified, sample projects were evaluated to determine which credits were applicable to which project types. Where possible, more than one project was evaluated for each project type to give the most accurate assessment. As a result of this exercise, the team was able to identify a list of potential credits for each project type. Since each credit has a range of possible points depending on the level of sustainability, each project was scored to determine the range of achievement for each project type. An interactive Excel checklist was created for use on each project to identify applicable credits. (See Exhibit 2.)

## Next Steps

The PANYNJ is currently in a soft launch phase. During this phase, executive staff throughout the agency will be briefed and the guidelines will be rolled out on a trial basis. Several projects under construction are already implementing the new sustainable initiatives for field testing purposes. The intent of these guidelines is to optimize design and to ensure that all potential sustainable initiatives have been reviewed. Concerns have been raised about potential impacts to costs and schedules, but as the case studies have shown, many of these initiatives have the potential not only to reduce cost and improve schedules but also to minimize facility disruption. At the end of the year, the pilot projects and the new initiatives will be evaluated, and the guidelines will be finalized for official implementation.

**Acknowledgements**

We would like to acknowledge the contributions of the following:
Croxton Collaborative Architects, P.C.
475 Fifth Avenue, 22nd Floor
New York, NY 10017

# CREDIT LIST

**SITE SECTION**

IS-1 Utilize an Integrated Team Approach
IS-2 Prepare a Site Assessment
IS-3 Maximize Use of Previously Developed Sites
IS-4 Maximize Use of Known Contaminated Sites
IS-5 Protect the Ecological Health of Wetlands and Floodplains
IS-6 Protect and Maintain Absorbent Landscapes
IS-7 Utilize Pervious Pavement
IS-8 Utilize Appropriate Vegetation
IS-9 Use Turfgrass Appropriately
IS-10 Amend and Reuse Existing Soils
IS-11 Balance Earthwork
IS-12 Coordinate Utility Work
IS-13 Utilize Trenchless Technology
IS-14 Mitigate Heat Island Effect
IS-15 Minimize Light Pollution
IS-16 Optimize Public Environments – Bicycles and Pedestrians
IS-17 Optimize Traffic Safety
IS-18 Optimize Roadway Alignment Selection
IS-19 Expand or Enhance Intermodal Connectivity
IS-20 Use Transportation System Management
IS-21 Use Transportation Technologies

**WATER SECTION**

IW-1 Implement Stormwater Best Management Practice Strategies
IW-2 Implement Rainwater Neutrality
IW-3 Reduce Use of Potable Water for Irrigation
IW-4 Utilize End Use Metering – Water

**ENERGY SECTION**

IE-1 Optimize Energy Performance
IE-2 Commission Electrical and Mechanical Systems
IE-3 Utilize End Use Metering – Energy
IE-4 Use On-Site Renewable Energy
IE-5 Protect Ozone Layer
IE-6 Provide Alternative Fueling Stations

**MATERIAL SECTION**

IM-1 Use Recycled Materials
IM-2 Use Local / Regional Materials
IM-3 Reuse Materials
IM-4 Use Durable Materials
IM-5 Use Sustainably Harvested Wood
IM-6 Minimize Use of Toxic and / or Hazardous Materials
IM-7 Enhance Pavement Lifecycle
IM-8 Preventative Pavement Maintenance
IM-9 Utilize Warm-Mix Asphalt Technology

**CONSTRUCTION SECTION**

IC-1 Minimize Pollution from Construction Activity
IC-2 Protect Existing Natural Systems
IC-3 Utilize Transportation Management During Construction
IC-4 Utilize Green Construction Equipment
IC-5 Reduce Noise and Vibration During Construction
IC-6 Implement Construction Waste Management
IC-7 Implement Integrated Pest Management During Construction

**OPERATIONS AND MAINTENANCE SECTION**

IO-1 Implement Sustainable Landscape Maintenance
IO-2 Maintain Soil Quality

Exhibit 1: List of Credits from the Sustainable Design Project Manual for Infrastructure Projects (April 20, 2010)

## PORT AUTHORITY OF NEW YORK & NEW JERSEY
## SUSTAINABLE DESIGN PROJECT MANUAL FOR INFRASTRUCTURE PROJECTS

**GENERAL PROJECT INFORMATION**

PROJECT : Bay Runway (sample project)
LOCATION: JFK
LE/A:
CERTIFIED BY:
DEPARTMENT:
EMAIL ADDRESS:

TARGET RATING:
TARGET NUMBER OF POINTS:

I certify that the information contained in this document is correct and accurate.

SIGNATURE: _____

DATE: _____

**PROJECT TYPE**

**Choose a Project Type**

| | |
|---|---|
| **Airfield New Construction / Reconstruction** | ☑ |
| Airfield Pavement Rehabilitation | ☐ |
| Bridge - New | ☐ |
| Bridge and Tunnel Rehabilitation | ☐ |
| Civil - Work Orders | ☐ |
| Intelligent transportation System | ☐ |
| Marine Structures - Docks, Wharves, Bulkheads, etc. | ☐ |
| Parking Lot New Construction / Reconstruction | ☐ |
| Parking Lot Rehabilitation | ☐ |
| Port Site Work | ☐ |
| Roadway New Construction / Reconstruction | ☐ |
| Roadway Pavement Rehabilitation | ☐ |
| Trackwork | ☐ |
| Utility New Construction | ☐ |
| Utility Rehabilitation | ☐ |

**Choose Type of Additional Work to be Undertaken**

| | |
|---|---|
| **Landscaping** | ☑ |
| **Exterior Lighting** | ☑ |
| MECH / EL / Fire Suppression System Installation | ☐ |

**SITE**

| CREDIT NUMBER | POINTS ACHIEVABLE | CREDIT NAME | PURSUING CREDIT? | POINTS ACHIEVED | INCLUDE CREDIT |
|---|---|---|---|---|---|
| **IS-1** | 2 | Utilize Integrated Team Approach | ☑ | 2 | ☐ |
| **IS-2** | 2 | Prepare a Site Assessment | ☑ | 2 | ☐ |
| **IS-3** | | Maximize Use of Previously Developed Land | ☐ | | ☐ |
| **IS-4** | 3 | Maximize Use of Known Contaminated Sites | ☑ | 3 | ☑ |
| **IS-5** | 2 | Protect the Ecological Health of Wetlands and Floodplains | ☐ | | ☐ |
| **IS-6** | 1 to 3 | Protect and Maintain Absorbent Landscapes | | | ☐ |
| | | 15% of absorbant landscape protected and maintained (1 point) | ☐ | | |
| | | 30% of absorbant landscape protected and maintained (2 points) | ☐ | | |
| | | 45% of absorbant landscape protected and maintained (3 points) | ☐ | | |
| **IS-7** | | Utilize Pervious Pavements | | | ☐ |
| | | 25% of total pavement area utilizes pervious pavement (1 point) | ☐ | | |
| | | 50% of total pavement area utilizes pervious pavement (2 points) | ☐ | | |
| | | 75% of total pavement area utilizes pervious pavement (3 points) | ☐ | | |
| **IS-8** | 2 | Utilize Appropriate Vegetation | ☑ | 2 | ☐ |
| **IS-9** | 1 | Use Turfgrass Appropriately | ☑ | 1 | ☐ |
| **IS-10** | 1 | Amend and Reuse Existing Soils | ☑ | 1 | ☐ |
| **IS-11** | 1 to 3 | Balance Earthwork | | | ☐ |
| | | 25% less removal of material or less new material required (1 point) | ☐ | | |
| | | 50% less removal of material or less new material required (2 points) | ☐ | | |
| | | 75% less removal of material or less new material required (3 points) | ☐ | | |
| **IS-12** | 2 | Coordinate Utility Work | ☐ | | ☐ |
| **IS-13** | 1 to 3 | Utilize Trenchless Technology | | | ☑ |
| | | Rehabilitation of existing pipe (2 points) | ☑ | 2 | |
| | | Replacement of pipe (1 point) | ☐ | | |
| **IS-14** | 1 to 3 | Mitigate Heat Island Effect | | | ☐ |
| | | 50% of site area utilizes heat island effect mitigation strategies (1 point) | ☑ | | |
| | | 75% of site area utilizes heat island effect mitigation strategies (2 points) | ☑ | | |
| | | 95% of site area utilizes heat island effect mitigation strategies (3 points) | ☑ | 3 | |
| **IS-15** | 2 | Minimize Light Pollution | ☑ | 2 | ☐ |
| **IS-16** | | Optimize Public Environments - Bicycles and Pedestrians | | | ☐ |
| | | Pedestrian amenities (1 point) | ☐ | | |
| | | Bicycle amenities (2 points) | ☐ | | |
| **IS-17** | | Optimize Traffic Safety | ☐ | | ☐ |
| **IS-18** | | Optimize Roadway Alignment Selection | ☐ | | ☐ |
| **IS-19** | | Expand or Enhance Intermodal Connection | ☐ | | ☐ |
| **IS-20** | | Use Transportation System Management | ☐ | | ☐ |
| **IS-21** | | Use Transportation Technologies | ☐ | | ☐ |

| | CREDIT NUMBER | POINTS ACHIEVABLE | CREDIT NAME | PURSUING CREDIT Y / N | POINTS ACHIEVED | INCLUDE CREDIT |
|---|---|---|---|---|---|---|
| WATER | IW-1 | 1 to 3 | Implement Stormwater Best Management Practices | | | ☐ |
| | | | Project site in New Jersey (1 point) | ☐ | | |
| | | | Project site in New York (3 points) | ☑ | 3 | |
| | IW-2 | 3 | Implement Rainwater Neutrality | ☐ | | ☐ |
| | IW-3 | 2 | Reduce Use of Potable Water for Irrigation | ☑ | 2 | ☐ |
| | IW-4 | 2 | Utilize End Use Metering - Water | ☑ | 2 | ☐ |
| ENERGY | IE-1 | 2 to 6 | Optimize Energy Performance | | | ☐ |
| | | | 10% reduction (2 point) | ☐ | | |
| | | | 20% reduction (4 points) | ☐ | | |
| | | | 30% reduction (6 points) | ☐ | | |
| | IE-2 | | Commissioning Electrical and Mechanical Systems | ☐ | | ☐ |
| | IE-3 | 2 | Utilize End Use Metering - Energy | ☐ | | ☐ |
| | IE-4 | 2 | Use On-Site Renewable Energy | ☐ | | ☐ |
| | IE-5 | | Protect Ozone Layer | ☐ | | ☐ |
| | IE-6 | | Provide Alternative Fueling Stations | ☐ | | ☐ |
| MATERIAL | IM-1 | 1 to 3 | Use Recycled Materials | | | ☐ |
| | | | Specify one (1) type of material (1 point) | ☐ | | |
| | | | Specify three (3) types of materials (2 points) | ☑ | 2 | |
| | | | Specify five (5) types of materials (3 points) | ☐ | | |
| | IM-2 | 1 | Use Local / Regional Materials | ☑ | 1 | ☐ |
| | IM-3 | 1 | Reuse Materials | ☑ | 1 | ☐ |
| | IM-4 | 1 | Use Durable Materials | ☑ | 1 | ☐ |
| | IM-5 | | Use Sustainably Harvested Wood | ☐ | | ☐ |
| | IM-6 | | Minimize Use of Toxic and / or Hazardous Materials | ☐ | | ☐ |
| | IM-7 | 1 | Enhance Pavement Lifecycle | ☑ | 1 | ☐ |
| | IM-8 | | Utilize Thin Surface Paving | | | ☐ |
| | | | Use of thin surface paving (1 point) | ☐ | | |
| | | | Use of thin asphalt concrete overlay < 1" thick (1 point) | ☐ | | |
| | IM-9 | | Utilize Warm Mix Asphalt Technology | ☐ | | ☐ |
| CONSTRUCTION | IC-1 | 2 | Minimize Pollution from Construction Activity | ☑ | 2 | ☐ |
| | IC-2 | 2 | Protect Existing Natural Systems | ☑ | 2 | ☐ |
| | IC-3 | 1 | Utilize Transportation Management During Construction | ☑ | 1 | ☐ |
| | IC-4 | 1 | Utilize Green Construction Equipment | ☑ | 1 | ☐ |
| | IC-5 | | Reduce Noise and Vibration During Construction | ☐ | | ☐ |
| | IC-6 | 1 to 2 | Implement Construction Waste Management | | | ☐ |
| | | | 75% diversion - all required materials (1point) | ☑ | 1 | |
| | | | 75% diversion - all recommended materials (1 point) | ☐ | | |
| | IC-7 | | Implement Integrated Pest Management During Construction | | | ☐ |
| O + M | IO-1 | 2 | Implement Sustainable Landscape Maintenance | ☐ | | ☐ |
| | IO-2 | 2 | Maintain Soil Quality | ☐ | | ☐ |

| TOTAL POINTS | | RATING | POINTS | | | | |
|---|---|---|---|---|---|---|---|
| TOTAL POINTS ACHIEVABLE: | 68 | CERTIFIED | 31 | to | 40 | POINTS ACHIEVED: | 38 |
| | | GOLD | 41 | to | 50 | | |
| | | PLATINUM | 51 | to | 68 | RATING ACHIEVED: | CERTIFIED |

© 2010

Exhibit 2: Sample project using the Interactive Credit Checklist from the Sustainable Design Project Manual for Infrastructure Projects (April 20, 2010)

[i] Benz, Stephen, *Sustainable Design: A Balancing Act*, Stormwater Solutions, February 2008

# Applicability of Decorative Concrete Overlays as a Context Sensitive Solution for Transportation Infrastructure

Leslie McCarthy[1], Jason Riegler[2], Ryan DaPonte[3], Aleksandra Radlinska[4]

[1] Full Member of ASCE, Assistant Professor, Civil and Environmental Engineering, Villanova University, 800 Lancaster Avenue, Villanova PA 19085 USA. Email: leslie.mccarthy@villanova.edu
[2] Graduate Researcher, Civil and Environmental Engineering, Villanova University, 800 Lancaster Avenue, Villanova PA 19085 USA. Email: jason.riegler@villanova.edu
[3] Undergraduate Research Assistant, Civil and Environmental Engineering, Villanova University, 800 Lancaster Avenue, Villanova PA 19085 USA. Email: ryan.daponte@villanova.edu
[4] Full Member of ASCE, Assistant Professor, Civil and Environmental Engineering, Villanova University, 800 Lancaster Avenue, Villanova PA 19085 USA. Email: aleksandra.radlinska@villanova.edu

## ABSTRACT:

This research seeks to determine the feasibility of decorative concrete overlay (DCO) as an aesthetic treatment for new and existing concrete infrastructure components. The research approach includes exposing DCO-treated concrete samples to various strength and environmental conditions in the laboratory and field. Laboratory testing such as ultraviolet light exposure, flexural fatigue strength, 90-day salt-ponding and 90-day sodium-chloride tank submersion were conducted on two sets of concrete specimens aged 28 days and 125 days. The DCO designs evaluated in the research included a monochromatic-textured brick pattern and a polished multi-tone stone pattern.

After 1500 hours of exposure to cycles of moisture and light, the accelerated ultraviolet light exposure showed no visual signs of surface degradation on DCO in terms of fading, color change, lustre, or cracking. The flexural fatigue testing results proved very effective given that the bond between concrete and DCO remained intact under cyclical loading. Also, the 90-day salt-ponding test displayed expected side effects of rusting under their metal embankment walls. Only sections in direct contact with the metal were permanently stained while all other sections were not affected. The sodium chloride submersion brick-designed specimens left behind minor salt-like residues on portions of the surface. Ultimately, the durability, versatility, and ease of application of DCO allude to its viability for aesthetically treating infrastructure.

## INTRODUCTION

The concept of applying DCO in infrastructure is an innovative solution that supports the greater initiative nationwide for widespread implementation of green engineering through Context Sensitive Solutions (CSS). Decorative concrete overlay presents a potential advancement for CSS by introducing a new technique and product for concrete components of infrastructure. DCO is a polymer-acrylic cementitious microtopping a few millimeters thick, colored to look like stone, brick, wood, etc., and skimmed or sprayed onto existing or new structural concrete surfaces. The objective of this study was to evaluate the durability and aesthetic characteristics of DCO under various environmental conditions in the laboratory. Successful proof-of-concept may benefit both horizontal and vertical applications in infrastructure, such as concrete traffic barriers, noise walls, walkways, retaining walls, and bridge piers or parapets. Two design patterns typical in highway applications were selected for evaluation: a textured monochromatic-brick and a polished multi-tone grey stone. The study sought to investigate the bond and surface degradation characteristics of DCO through both laboratory tests and a field section. Laboratory tests included the flexural fatigue, salt-ponding, and sodium chloride submersion test. Additionally, visual inspection techniques and ultraviolet light testing were used to evaluate the surface degradation of both laboratory specimens and a field section of DCO exposed to real-time weathering.

### Preparation of Test Samples

As described above, it was determined that flexural fatigue strength, ultraviolet light exposure, salt-ponding and sodium chloride tank submersion tests would be conducted to evaluate the robustness of DCO. During sample specimen preparation, the beams and slabs were cast using wooden forms and the standard concrete cylinders were cast using 100×200 mm (4×8 in.) plastic molds. Twelve beams were fabricated to precisely 915×400×140 mm (36×5.5×5.5 in.). In order to test the durability of DCO on vertical components of infrastructure, concrete samples were designed in accordance with Pennsylvania Department of Transportation (PennDOT) specifications for mixtures used in vertical structures (PennDOT Publication 408, 1994). In order to be consistent with PennDOT Section 714 (standards for concrete barriers, median barriers, and sound barrier panels), the design used No. 57 coarse aggregate, standard Type A fine aggregate, and ordinary portland cement. The concrete was designed to have a water to cement ratio of 0.45 and 6% air content. The specimens were mixed using a 0.2-cubic meter (7-cubic foot) concrete mixer and poured into the formwork following ASTM C192 procedure (ASTM, 2007). Additionally, each beam was reinforced with one #5 bar placed directly in the center of the sample. After 24 hours, the samples were removed from the forms and cured in a moist room at a constant temperature of 22.8 °C (73 °F) with a relative humidity of 98% for 7 days. Similar formwork was designed for slab specimens.

**Application of DCO to Test Samples**

A complete surface preparation of the specimens is essential to ensure optimum bond strength between the structural concrete and the decorative overlay (Berger, 1989). Therefore sufficient cleaning and profiling of the concrete was necessary to prevent any bonding failures. There are two conditions that must be met before the coating can achieve a strong bond with the concrete. The first condition is that the surface must be clean of dirt, chemicals, and other contaminants. The second condition is that the surface must be given a mechanical etch or profile. In this work, the first step of surface preparation was to scarify or grind down the surface to remove any dirt, coatings, grease, or sealers. Denatured alcohol was then rubbed onto the surface of the concrete for etching. The final step in sample preparation was cleaning with a high-pressured water jet prior to DCO application on both series of samples to be tested at 28 and 125 days.

The mixture components of the decorative concrete overlay included cement (series 105 polymerized dry component), a liquid modifier (modified acrylic resin), and any desired colorant (Concrete Technology Inc, 2009). The proportion of a complete mixture used was 8 ounces colorant to 544 ounces modifier to 50 pounds of cement. The mixing process occurred immediately prior to application of DCO to the test specimens by trowel, sponge, and/or spray gun. While each application technique can be done either separately or concurrently, beams, cylinders and brick-designed slabs in this study were treated using a trowel and spray gun, while the stone-patterned slabs were applied by the sponge application. A primer coat consisting of three parts water to one part liquid modifier was sprayed onto the concrete surface to ensure enhanced adhesion. Then the first coat of sandstone-colored DCO was applied using a trowel and allowed to dry for approximately two hours (Figure 1a). Afterward, grout lines (Figure 1b) were designed by applying lines of tape along concrete samples before spraying the second coating of red colorant on the slabs. This procedure was used to make the slabs resemble a monochromatic brick pattern as pictured in Figure 1c.

**(a) (b) (c)**

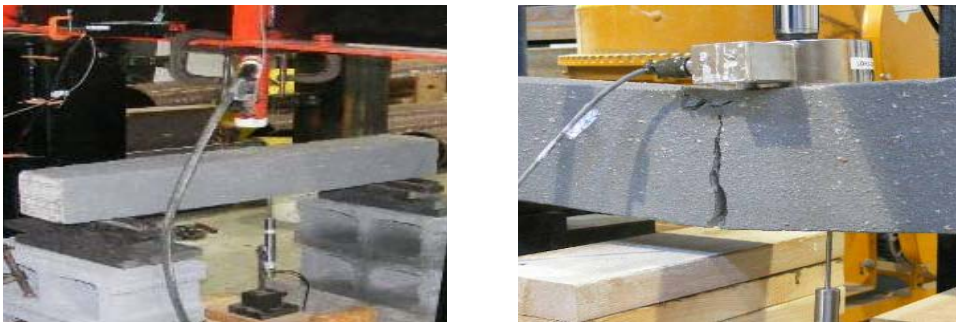
**Figure 1. (a) First coat (sandstone), (b) grout lines, (c) finished brick pattern**

The last and final step in the decorative overlay preparation process was the application of two clear sealant coats, designed to enhance the DCO's durability by providing protection against environmental elements including temperature, sunlight,

and humidity. The sealant coat was also used to prevent penetration of dirt and chemicals and is designed to enhance the color and vibrancy of the DCO. Manufacturers recommend that the clear coat be reapplied every three years to the overlay for the purpose of preventative maintenance (Concrete Technology Inc., 2009).

**Flexural Fatigue Testing**

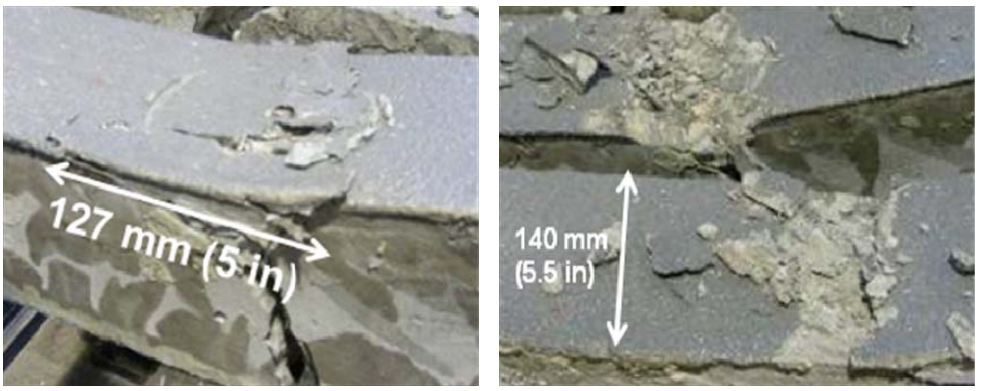
Flexural fatigue testing was performed on concrete beams with a DCO cover to test the effectiveness of the bonding between the concrete and decorative layer under the flexural load. The DCO was applied on three of the four longitudinal faces of the beam, namely at the top and sides. A third-point loading test, ASTM C 78-08 (ASTM, 2008) was used to test the DCO durability by investigating the effect of cyclical loading on the bond of the DCO to structural concrete. The simply-supported reinforced decorative concrete beam specimens were subjected to loads applied in 90.7 kg (200-lb) increments, and then qualitatively examined for cracks and other failures in the overlay before increasing load to the next level. The apparatus, as pictured in Figure 2a and 2b, included a load cell, hydraulic piston, and a deflection gauge located at the center span of the beam. As the amount of force at center-span increased, the decorative concrete beam was monitored for cracking, particularly in the overlay. All twelve beams were loaded until failure.



**(a)** **(b)**

**Figure 2. (a) Flexural fatigue test system, (b) piston applying force to load cell**

Concrete failure, as shown in Figure 2b, led the decorative concrete overlay to bubble and chip off. However, an inspection of the chips that had broken off revealed that the overlay had remained attached to the top layer of the broken concrete pieces. This implies that the force applied to the beam caused spalling to the concrete and not the decorative overlay (concrete fractured rather than the overlay) as seen in Figures 3a and 3b. Under cyclical loading, the reinforced concrete beam failed as expected with cracking and spalling primarily at center span. Directly underneath the hydraulic piston, no cracks were found in the overlay. Ultimately for both the 28-day and 125-day aged specimens, the bond between the DCO and structural concrete endured the force applied from the piston.




(a) (b)

**Figure 3. (a) Spalling on 2nd beam tested, (b) spalling on first 2 beams tested**

## Ultraviolet Light Exposure

The objective of the ultraviolet light (UV) experiment was to expose ten specimens through replication of weathering effects such as direct sunlight, rainfall, and heat exposure. A surface area of approximately 58 $cm^2$ (9 $in^2$) on each specimen was subjected to stages of ultraviolet light and moisture using the Q-SUN Xenon Test Chamber. The chamber's interior lamps simulate damaging wavelengths of sunlight by emitting visible ultraviolet and infrared light (Q-Lab, 2008). A graph illustrating the close comparison between xenon arc light and actual sunlight is presented in Figure 4 (Jones, 2007). The DCO-treated specimens were subjected to 120-minute repetitive cycles following ASTM G155-05 Cycle 1 (ASTM, 2005). Each cycle consisted of 102 minutes of light followed by 18 minutes of light accompanied by continuous water spray, resulting in a total testing period of 1800 hours. Exposure conditions included an irradiance of 0.35 $W/m^2$, a wavelength 340 nm, and a constant temperature of 50 °C. Given the limitation on the amount of sunlight possible in one actual day, an accurate estimation of the time of exposure corresponds to 219.5 days based on the average climate rate of 8.2 hours of sunlight per day in the state of Florida (Florida Climate, 2008). Floridian climate was chosen as the one with largest amount of sunlight, and as such, presenting the harshest climate.

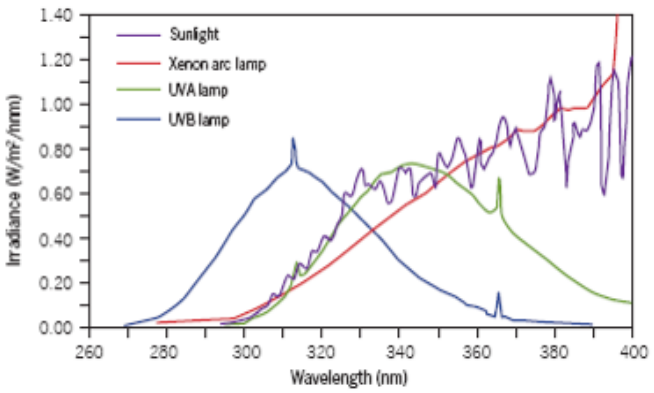



**Figure 4. Spectral output of various UV sources compared to sunlight (Jones, 2007)**

Even after 1800 hours of simulated continuous sunlight, with partial cycles of water spray, none of the specimens showed any visible signs of degradation. In this case, its surface degradation is qualified in terms of fading, color change, loss of glossiness or sheen, or cracking. Figures 5a and 5b display the comparison between the weathered specimens and control specimens for each DCO design tested.


**(a)** **(b)**

**Figure 5. UV-exposure results (a) DCO-brick control vs. weathered brick, (b) DCO-stone control vs. weathered stone**

**Salt-Ponding Test**

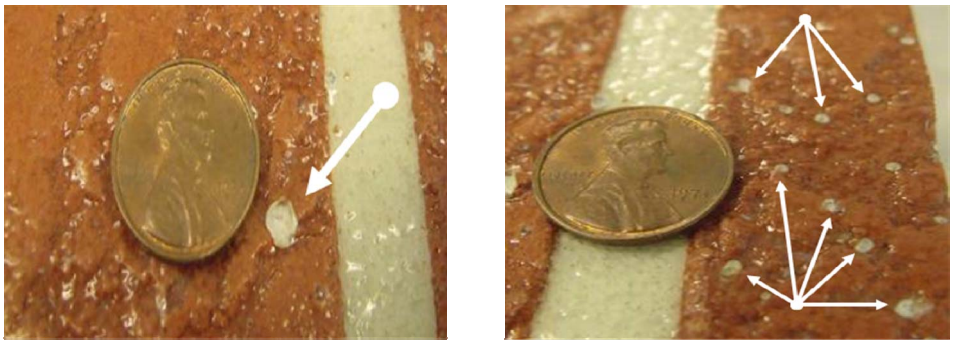
In order to observe the effect of sodium chloride ions on the decorative concrete coating, a 90-day salt-ponding test was conducted in accordance with ASTM C1543-02 (ASTM, 2002). The intent of running this test was to simulate the effects of deicing salts on DCO-treated concrete. In the experiment, a 645 $cm^2$ (100 $in^2$) embankment frame filled with a reservoir of sodium chloride solution was set upon the surface of six laboratory slabs to enable continuous ponding. The framing braces were made of metal and were caulked along the interior edges as shown in Figures 6a and 6b. Testing was performed on six 28-day old slabs split evenly between the monochromatic brick and multi-toned stone patterns. Metal was selected as the frame material to evaluate the effect of DCO's contact with metal posts, guardrail or fencing, such as along sidewalks and trails. Each specimen was coated on all four sides with an epoxy to prevent any lateral moisture from escaping. A 3% reagent-grade sodium chloride (NaCl) solution was placed at a depth of 15 ± 5 mm on the DCO slab and stored in an environmental chamber at a temperature of 23.0 ± 2 °C with a relative humidity of 50 ± 5 %. During storage, the slabs were covered with a polyethylene sheet to prevent the evaporation of water from the solution. After 90 days, the embankment frames were drained and specimens were allowed to dry. Subsequently, each slab was examined by comparison to a control specimen to identify any resulting changes to the physical properties of the overlay.

As a result of continuous exposure to the sodium chloride solution within the metal-framed embankment, a significant amount of rust formed after a period of 90-days, consequently staining the DCO. Figures 6c and 6d show the ponded specimens after both the removal of the embankment frame and washing of the surface.

(a) (b) (c) (d)

**Figure 6. Ponding specimens (a) stone pattern (before test), (b) brick pattern (before test), (c) stone pattern (after test), (b) brick pattern (after test)**

A significant impact of the ponding test was the formation of small circular pockets where the decorative coating had washed away (Figures 7a and 7b). These small craters were found only on the brick specimens and measured between 0.2 and 0.6 cm in diameter. It appears that this damage was the result of air bubbles that had formed during the application of the DCO protective clear coat that then had burst during the chloride intrusion testing of the surface.



(a) (b)

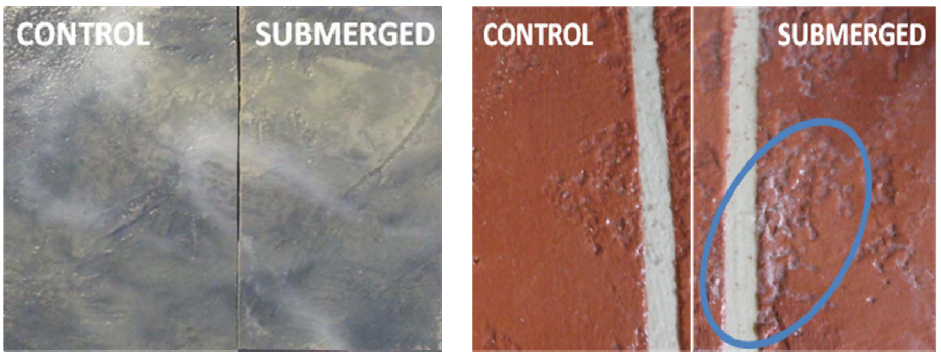
**Figure 7. Ruptured surface craters (a) large, (b) small, evident after conclusion of chloride ponding test**

## Chloride Tank Submersion Test

The sodium chloride tank submersion test was engaged to simulate the effects of DCO-treated concrete infrastructure components (e.g., bridge piles) when submerged in marine saltwater environments. Visual inspections of the decorative concrete coating were made after submerging six slab specimens and five cylindrical specimens in a sodium chloride bath for 90 days. The procedure followed the ASTM C 1556-04, which determines the penetration of chlorides into concrete (ASTM, 2004). In this study, the uncoated areas (i.e., exposed structural concrete) of the specimens were sealed with an epoxy and DCO-surfaced area was left unsealed. The specimens were then submerged into a limewater bath and saturated for 24 hours to prevent absorption during subsequent chloride solution testing. Before testing, each specimen was thoroughly rinsed with tap water and then placed in the sodium chloride solution. After a period of 90 days, each individual slab and cylinder was

examined and compared to a control specimen to record any identifiable changes to the physical properties of the DCO.

Figure 8a compares the stone-patterned specimens where no evidence of salt residue or surface degradation was found. Figure 8b compares brick-patterned samples coated with a white salt-like residue at various locations on the surface of the DCO compared to the control. This effect was not evident on the smoother multi-toned stone pattern, which could mean that monochromatic and/or textured patterns are more susceptible to the saline liquid residue. Other surface degradation characteristics such as fading, flaking, or loss of luster were not found in either DCO design. Variable tones of DCO may be more applicable for concrete structures that will be exposed to sodium chloride runoff or partial submersion in marine environments.




**(a)** **(b)**

**Fig. 8. Control vs. submerged specimen (a) stone (no salt-residue), (b) Control specimen (salt-residue)**

## CONCLUSIONS

The focus of this research was to initiate evaluation of a new potential context sensitive solution tool for state transportation agencies to utilize on new or existing infrastructure. Thus far, the research has shown that decorative concrete overlay (DCO) appears to be a durable and easily accessible tool for aesthetic treatment of today's infrastructure. DCO demonstrated efficiencies in terms of ease and speed of application by three simplistic approaches (trowel, sponge, and/or spray), which could help minimize work zone setup times and impacts to the traveling public. Flexural fatigue, UV light exposure, salt-ponding and sodium chloride tank submersion tests were conducted on both 28-day and 125-day old specimens. Thus far, testing has confirmed that the bond of DCO to structural concrete can withstand extreme load and weathering conditions. In all twelve flexural fatigue tests, the bond of DCO to concrete was only broken along fracture lines, proving that the concrete and not the DCO had cracked during increased loading cycles. The ultraviolet light exposure results showed that DCO can withstand excessive rays of sunlight in extreme conditions, without losing key decorative qualities such as luster and color brilliance. A considerable amount of rust deposited on the DCO surface under metal framing, as a result of the salt-ponding test. In addition, small circular air pockets

caused minute amounts of surface degradation as a result of application of the protective clear coat. Chloride tank submersion testing did not result in any permanent surface deterioration. The successful laboratory results in terms of strength and resistance to extreme weathering prove that DCO is an effective option for the aesthetic treatment of transportation infrastructures.

## RECOMMENDATIONS

Initial robustness testing of DCO points positively to its potential as a CSS tool; however, further research is required to prove its durability in terms of bond strength and full-scale environmental conditions. For example, an investigation into the freeze-thaw durability is necessary to determine the ability of the overlay to withstand temperature cycling. This is particularly important to ascertain the use of DCO on infrastructure components in regions that are subjected to a variety of weather extremes. Additionally, results from pull-off adhesion strength testing would give important insight into the bond strength of the DCO to the concrete surface. During the salt-ponding test, the direct contact of the metal embankment wall with the sodium chloride solution caused oxidation to take place, consequently influencing the aesthetics of the DCO. In order to prevent this from happening, a galvanizing coating must be applied to the surface of the metal attached to the concrete structure. Otherwise, a non-corrosive material such as plastic or aluminum for posts, fencing, etc., is preferred when installed in DCO-treated infrastructure. Further experimental investigation of salt-ponding tests with plastic embankment frames is suggested for future work. Also, the specimens subjected to sodium chloride tank submersion displayed a noticeable saline residue on the surface of the overlay, specifically on the monochromatic brick-textured slabs. Therefore, initial research indicates that any infrastructure components that would be subjected to aqueous marine conditions such as bridge piers may use DCO multi-toned colorant designs with a smooth polished texture to avoid visible leftover salt concentrations during tidal fluctuations or after recession of deicing salt runoff.

## ACKNOWLEDGMENTS

The authors appreciate the generous support of Villanova University, AiRestoration, Dauphin Masonry, Inc., CTI, and Rowan University.

## REFERENCES

ASTM Standard C 1543, 2002, “Determining the Penetration of Chloride Ion into Concrete by Ponding”, ASTM International, West Conshohocken, PA.

ASTM Standard C 1556, 2004, “Determining the Apparent Chloride Diffusion Coefficient of Cementitious Mixtures by Bulk Diffusion”, ASTM International, West Conshohocken, PA.

ASTM Standard G 155, 2005a, “Standard Practice for Operating Xenon Arc Light Apparatus for Exposure of Non-Metallic Materials”, ASTM International, West Conshohocken, PA.

ASTM Standard C192, 2007, “Standard Practice for Making and Curing Concrete Test Specimens in the Laboratory, ASTM International, West Conshohocken, Pa.

ASTM Standard C 78, 2008, “Standard Test Method for Flexural Strength of Concrete (Using Simple Beam with Third-Point Loading)”, ASTM International, West Conshohocken, PA.

Berger, William F. "Surface Preparation for Concrete Flooring." *Journal of Protective Coatings and Linings* (1989). Web.

Concrete Technology Inc. (2009), “CTI-200 Series Pro Seal Sealer,” Concrete Technology Inc., Largo, FL.

Concrete Technology Inc. (2009), “CTI System Series 105 Grout,” Concrete Technology Inc., Largo, FL.

Concrete Technology Inc. (2009), “CTI System Series 110/111 Modifier,” Concrete Technology Inc., Largo, FL.

Deatrick, John, and Shari Schaftlein. *Results of Joint AASHTO/FHWA Context Sensitive Solutions Strategic Planning Process*. Rep. Raleigh: Center for Transportation and the Environment, 2007. Print.

Jones, Mark. "Accelerated Weathering and Durability." *BUILD* Oct.-Nov. 2007: 65-66. *http://www.branz.co.nz/cms_show_download.php?id=548*. Web. 22 June 2010.

Lins, V.F.C., and M.F.A.S. Araujo. "Photodegradation of Hot-Mix Asphalt." *Fuel* 87.15-16 (2008): 3254-261. Print.

Lizarazo-Marriaga, Juan, and Peter Claisse. "Determination of the Concrete Chloride Diffusion Coefficient Based on an Electrochemical Test and an Optimization Model." *Materials Chemistry and Physics* 34.4 (2004): 603-09. Print.

"Q-Sun Xenon Test Chamber." *Q-Lab The Most Trusted Name in Weathering Testing*. 2008. Web. 03 June 2010. <http://www.q-lab.com/Q-Sun.html>.

"Tampa, Florida Climate, Temperature, Average Weather History, Rainfall/ Precipitation, Sunshine." *World Weather and Climate Graphs, Average Climate Charts, Guide to Precipitation, Temperatures, Best, Friendly, Holiday Climate*. 2008. Web. 03 June 2010. <http://www.climatetemp.info/usa/tampa-florida.html>.

United States of America. Commonwealth of Pennsylvania. Department of Transportation. *Publication 408, Section 714 - Precast Concrete Products*. 1994. Web. 24 June 2010.<http://www.dot.state.pa.us/Penndot/reginfo.nsf >.

Yang, C. C., and L. C. Wang. "The Diffusion Characteristics of Concrete with Mineral Admixtures between Salt Ponding Test and Accelerated Chloride Migration Test." *Materials Chemistry and Physics* 85.2-3 (2004): 266-72. Print.

# Creating Fully Sustainable Walkable Communities with PRT (Personal Rapid Transit)

Shannon Sanders McDonald, AIA, S1027arch@aol.com

## ABSTRACT

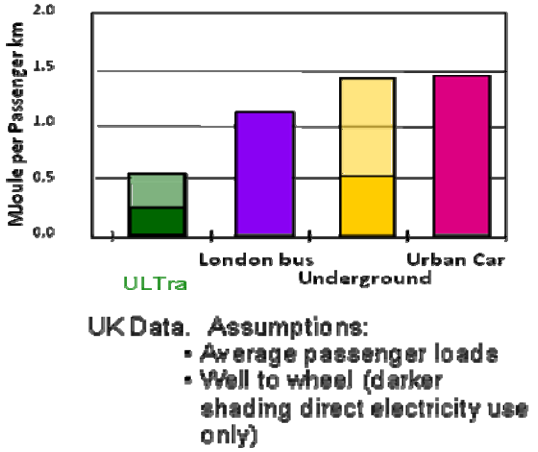
An emerging automated transit system, PRT can provide exactly the solution that has been sought to create fully walkable sustainable communities. This advanced technology allows for the full integration of walk-ability and the mix of other low tech solutions such as the bicycle. PRT (personal rapid transit) is currently either being implemented – such as at Heathrow airport or in the planning stages linking airport and surrounding transit village concepts such as at Sam Mineta airport in San Jose, California. PRT not only brings a greater level of service to the user due to its point to point movement pattern, but smaller driverless automated vehicles can provide linkages to sustainable energy sources and provide smaller more frequent stations for a fully compact walkable community. This paper shows how this critical balance between machine movement patterns, the built environment and the pedestrian can occur can occur with the application of PRT.

## FULL PAPER

A 1970's vision of modern transit technology is currently becoming reality around the world due to advances in computer technology. PRT (Personal Rapid Transit) now sometimes called ATN (Automated Transit Network) is currently available at Heathrow airport with systems in various planning stages in Suncheon City, South Korea; San Jose, CA -Sam Mineta Airport and Mazdar City, Abu Dhabi, United Arab Emirates - and studies for multiple other locations. PRT provides the opportunity to create walkable transit friendly environments by allowing more compact and interconnected ways to move through space due to its small size and new transit moment paradigm – point to point – providing service similar to a cab. This technology has the ability to move more people faster to their destination – even more time-efficiently than a car and more time-efficient than the standard transit loop or line systems; therefore they can encourage and hopefully increase transit use.

This technology offers another important design feature – its ability to interface three-dimensionally both inside and outside of buildings. This capability allows for many right-of way issues to be approached in various ways especially in existing urban fabrics by providing choices both on the ground and with air-rights. PRT can link multiple transit and automobile systems and connect with sustainable energy sources changing transportations' relationship with urban design and architecture. New integrated solutions that address multiple overlapping issues such as energy sources, the environment and persons of disability can now be easily addressed due to its small size and relative scale. PRT provides a lower cost and better energy use as documented for the modern Heathrow system. **(FIGURE 1)** There are many benefits for the application of PRT in our built world such as more affordable land use patterns, cost effective transit, ability to connect all current forms of movement, ability to be constructed in existing and historical environments, simultaneously used as a freight application and cost-effective interconnected station design.

Integrating a new way to move through space is a more sustainable solution that allows for stations within easy walking distance encouraging walking. Initially these systems are being implemented in two areas – local existing environments such as airports and entirely new cities, however multiple studies are underway around the world for implementation into existing cities and towns. This paper will examine the potential linkages for the Southern Polytechnic Campus of 2050 using PRT to create a more fully accessible walkable sustainable campus design.




| | $m/mile |
|---|---|
| Automatic People Movers (“APM”) | 30 - 100 |
| Light Rail | 20 - 40 |
| ULTra | 10 - 15 |

**FIGURE 1:** Chart showing energy usage and cost comparisons. ULTra: PRT Sustainable Transport Source: Martin Lowsen of ULTra

PRT was first implemented in the United States at Morgantown, West Virginia at West Virginia University using a 25 people capacity (now known as group rapid transit) point-to-point driverless vehicles (a system where you go directly from your point of origin to your final destination – no stops in between) called PRT, although not in full original vision. The dedicated elevated track system is still functioning today providing unique transportation choices for the university and the town and is in the process of studying expansion. Although at the time due to its groundbreaking nature the system was developed with extensive cost overruns slowing the further development of this paradigm changing movement technology.

The modern PRT systems, based on the original concept provides an on demand, point to point service; a small transit vehicle (optimal number is around 4 people) offers many positive benefits along with greater capabilities to move more people more time efficiently. The off-line station concept that is crucial to the efficient functioning of the on demand, point-to-point system presents unique opportunities for station design. The system also allows for expansion of use as needed with only the cost of additional vehicles. The original vehicles upfront costs could be limited to a small fleet; expansion as appropriate.

This newly emerging movement system will allow easier access for all ages as station design can now be integrated within buildings at multiple levels. Navigating steep topography and dealing with poor weather as well as connecting multiple other forms of transit to create a seamless system will allow faster and smoother access for all. Linking spaces and places in entirely new ways, the architects’ unique spatial training will become essential in allowing this new technology to be implemented to

its full potential.

There is a direct relationship in the ability to move around our built environment in a timely, accessible, and sustainable way and to the success of that living environment. Providing choices and options for each user can create the greatest success in achieving this goal. Finding long term sustainable energy, environmental, planning and financial solutions that can work for all and across our diverse country is difficult especially if depending on one form of movement – the car. Relying on the private car fueled by gasoline also poses the issues of parking and roads that spread out the distances between buildings in villages, towns and cities. Adding transit brings more cost into the equation while frequently extending travel times due to the combination of travel time of car and transit - discouraging transit use. Walking only is a good option if the center is designed within distances that the user will accept and that also link to multiple flexible transit options, however weather, carrying items, persons of disability and aging population concerns can make this approach not feasible for all. Concurrently addressing the full environmental impact adds even more challenges to finding appropriate solutions. Not incidentally the amount of land consumed for parking cars can be at least 54% of the total land area coverage in a typical commercial development as documented in a study by the public works department for the City of Olympia, WA. Parking for business and housing as well as for multiple other daily needs adds to the complexity for solving full accessibility movement issues. Add into the mix the need for delivery and freight in a timely way so that business can prosper and the complexity is overwhelming and the sustainable solutions hard to find. Many focus their research on one or two aspects; however this is starting to change as it is in the combination of the issues where better solutions can be found. The architect is trained to balance and resolve the multiple complex issues typically also "customizing" for the specific site, owner and location with training in synthesizing this spatial complexity. However, the architect is frequently consulted for the purpose of designing the station or to focus on what the built environment "looks like". While aesthetics is one key concern to an architect, how that is obtained is by finding solutions that balance ***all*** of the design needs.

Creating walkable transit friendly environments requires more compact and interconnected ways to connect the movement systems that we have for new design and movement through space and place.

**Image 1:** Japanese Bicycle Commuters

**Modes of Transportation**

The "standard" modes of transportation choices are: walking, bicycle, motorcycle, car, plane and multiple forms of transit. Walking offers some solutions within certain distances however, will not address the full range of human needs. Bicycling extends this distance and is used in Japan as part of the commuters experience from home to parking/transportation centers. **(IMAGE 1)**

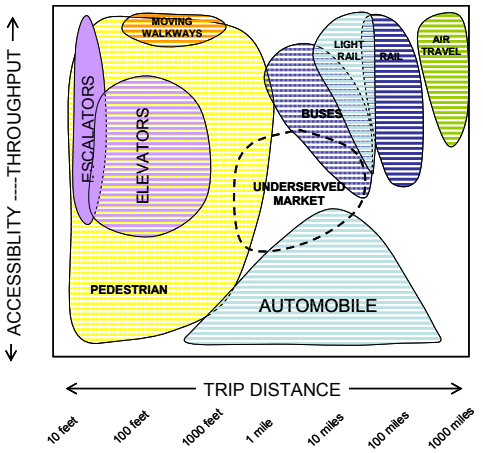



**FIGURE 2:** Conceptual comparison of modes in the Accessibility – Throughput spectrum from Young, Miller, McDonald. Keys to Innovative Transport Development. Presented at the 87 Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.

Motorcycle extends this distance even further. The car can now come in various sizes such as the electric cart all the way to the largest delivery vehicle with various ranges in-between, while transit can involve a bus or trolley to heavy and light rail. Depending upon each situation all forms of transportation can be found to serve different movement distances and needs. Movement needs are complex and interconnected and PRT can allow for these connections and interconnections to occur. **(FIGURE 2)**

**SOUTHERN POLYTECHNIC CAMPUS ANALYSIS 2009**

The architecture students at Southern Polytechnic State University were excited about the idea of designing the SPSU campus of 2050 implementing new ways to move around an expanded campus. They proposed how the University could expand in academics and student enrollment, proposing new courses of study and a physically enlarged campus, while contemplating how education might be delivered. They were given one requirement: to integrate newly emerging movement technologies into the campus plan. They began by analyzing the history, evolution and existing campus conditions.

**University History**

As part of the design process it is important to understand the history of the site and area -- as the past always holds seeds for the future. As SPSU was founded as a technical institute and continues today to emphasize technology education this

University appeared the appropriate place to embrace and utilize PRT. Upon further research the University also has historic and current research with movement technologies such as automobiles, submarines, and aircraft. Currently research with electric vehicles, Formula One cars, and submarines provided the core for embracing a 2050 vision. Solar power and nuclear plant technology are the core energy research areas for the campus. The University also historically had an unusual activity – the bathtub races, where bathtubs were converted into vehicles and then raced on the campus ring road. A great campus to think about new movement technologies! A very brief history:

1948 - SPSU was founded as the Technical Institute
1958 - SPSU became the Southern Technical Institute
1961 - Hoyt McClure/ Acting Director led to the movement of building 8 buildings on 120 acres of land.
2009 - Southern Polytechnic State University expanded and now encompasses more than 230 acres and 35 buildings.

**SPSU Figure/Ground**

The campus is a "typical" ring road campus with the original buildings constructed with the typography in mind. The first buildings were nestled into the hillside and oriented for passive heating and cooling - a very progressive environmental approach at the time. Over the years new development has expanded the ring round outward to, however the new buildings have not been oriented with either the topography or environment as part of the design strategy.

The campus is located with direct highway access to downtown Atlanta and Chattanooga TN. The students were also very interested in the connectivity between their campus and these locations as well as to both Georgia Tech and Georgia State in downtown Atlanta and The University of Georgia in Athens, GA. **(IMAGE 2)**

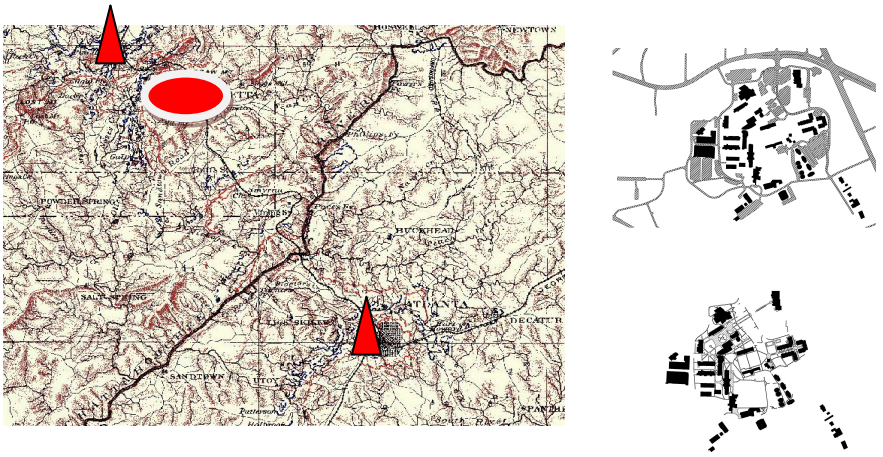


**IMAGE 2:** Topo map of Campus and surrounding area and figure ground with sidewalks and roads

**University Planning and Goals**

This is a commuter campus although on-campus housing is expanding as more traditional students are attending. The campus also provides training and continuing education programs mainly at night. Multiple types of parking therefore is required as part of the mix for the suburban Marietta, GA campus of Southern Polytechnic State University. The campus recently constructed its first parking structure – a typical stand alone structure. The majority of the small lots close to existing buildings remain (the historic pattern); however the plan moving forward is to construct more parking facilities on the ring road.

One student documented the congested area of traffic and they were located at the entry points to the campus- not where any parking structures have been built or are planned. This actually decreases the walkabilty due to the topography and the needs of the users and no attempt was made to address handicap accessibility within these plans. This is a common development strategy for many Universities to place parking on the periphery, however often implemented without any overall planning looking at all the pedestrian and vehicular movement patterns and structure of the campus. (**FIGURE 3)**

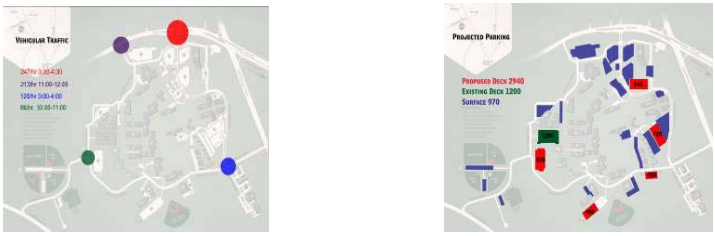



**FIGURE 3:** Southern Polytechnic State University student, Katie McCulloch analyzing vehicular traffic and projected parking.

The University has joined the American College and University Presidents Climate Commitment to sustainability and implementing PRT would be one way to address these goals. One student documented all of the natural existing conditions, sunlight, orientation, weather among others so that the original approach to sustainability could be understood and new buildings could address these concerns. Other students documented the potential student growth of the campus and every student suggested new areas of study that would fit with a technology campus.

**Walking and Inter-Site Movement**

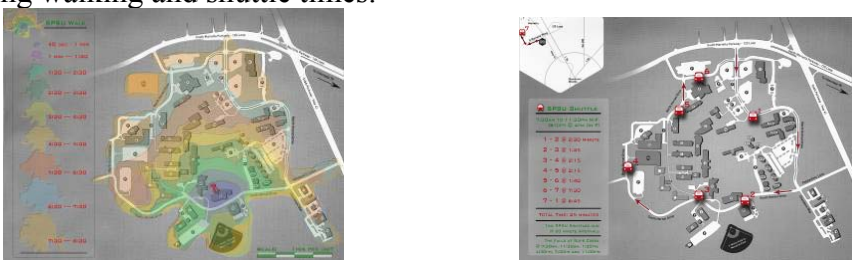
The campus is however quite walkable for the mobile, except if you are carrying many items, walking late at night or dealing with bad weather. The students documented these times as well as the areas of campus that received the most foot traffic. The students also documented the vegetation and walking paths. The new bus circulator was also tracked for the arrival and destination times to compare this to walking. (**FIGURES 4 and 5)**

The one-way loop system at SPSU was rarely used due to the time required to go from one point to the next on the bus shuttle was in almost all cases excessive in

comparison to the walking time through other campus paths. This is not an unknown condition for other existing one way loop transit system and travel times; as adding a shuttle is the first response when enough parking is available but not within walking distance to destinations without any analysis of the actual usefulness of such a system. While the intent was appropriate - the integration with the existing environment with movement patterns was not explored. Creating small strategically placed multi-modal parking facilities can assist in linking sustainable power sources for multiple efficient uses of building and machine. These new technologies are the perfect systems to begin to connect all of the possibilities to create more sustainable power and energy sources due to its new paradigm of smaller vehicles. The weather, topography of the land (hilly versus flat) also contributes to a successful design strategy if a fully accessible environment is to be created.


**FIGURE 4:** Southern Polytechnic State University student, Katie McCulloch analyzing existing pedestrian density and pedestrian traffic areas.

**FIGURE 5:** Southern Polytechnic State University student, Kira Melville analyzing walking and shuttle times.



**Handicap Accessibility**

One student thoroughly studied accessibility on campus; mapping and photographing the entire campus. This campus is basically not accessible and has many needs in meeting this goal. If PRT were to be implemented all of the issues related to solving this complex problem could be addressed. (**FIGURE 6)**

Accessibility is about more than providing for the persons of disability and meeting a set of safety rules; it is about providing mobility for all of us over our lifetime – Universal Design. As at some point in our lives all of us will face challenges in order to be mobile. Transportations' future will be dependent on the ability to provide access for everyone at every stage or point in their lives, especially as our population ages. This comprehensive view of transportation accessibility requires a broad vision for how many different systems can work together in combination with our existing built world to meet this goal, as one specific system can never meet all of our needs. The United States due to its size, breadth and

complexity of living environments from urban to rural and everything in between demands that we embrace and support a comprehensive network that interconnects all of us within our living places, from our starting point to our destination -- for every stage in our lives. Interweaving architectural solutions with transportations technological advances can address the full meaning of accessibility even for people that do not physically require it but their existing living environments do – such as traffic congested edge cities. Building upon the existing transportation systems that

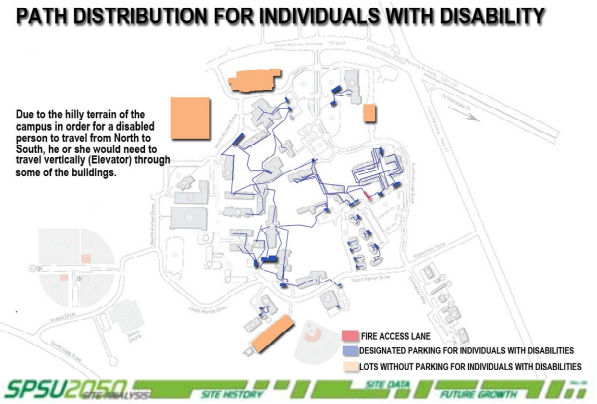



**FIGURE 6:** Southern Polytechnic State University student, Omar Foster analyzing path distribution for individuals with disabilities.

are now evolving to meet today's sustainable challenges and the varied architectural environments that currently exist; if we embrace emerging viable technological solutions that are called "people movers", then we can learn the meaning of bold ideas to meet big challenges – total accessibility for all.

Movement technologies that we can directly access from within a building can provide a safe and accessible way for certain travelers such as the elderly and children to reach their destination safely. A true interweaving of man and machine will provide better access and flexibility while creating wonderful new spaces for nature and man to exist at multiple levels and in surprising ways. Providing these options in conjunction with walkable environments can be created optimizing our living solutions.

Downtown Minneapolis is an excellent example of how multi-levels are used for greater pedestrian access to the city. The "second" level of the city interconnected by the skywalk system is directly links parking facilities and the urban fabric of the city. The parking facilities are spaced appropriately for easily access from outside multiple points. Pedestrians can easily, quickly and safely move around the city as there are no traffic and road issues to interact with. A new way of connecting between places, is also created as a more "natural or flexible way to connect is achievable.

A more "natural" approach to pedestrian movement can be established while providing a safe interaction with machines. A more "natural" approach to movement can be described as a networked system. **(Figure 7)** This networked system opens up new locations for retail, parks and people places that were previously not available within the grid or spoke and wheel organizational patterns for such use. If the

networked system allows and encourages multiple spatial overlaps between levels even more possibilities and linkages are created.

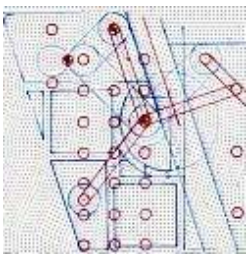


**FIGURE 7:** Montana State University student, Matthew Killham abstract visualization of new ways to provide linkages.

## SOUTHERN POLYTECHNIC CAMPUS DESIGN 2050

When working within the existing campus the current infrastructure of the road systems may handle a PRT type transit but typically would not allow the full development of its potential as a networked system. However, by starting with the existing road infrastructure, a base for the new PRT transit system can be established. PRT can function on the road based system reducing upfront costs, provide less congestion by eliminating individual drivers, and provide easy access to existing connection points while allowing for multiple destination points. This will provide for the user reduced travel time and therefore encourage transit use. Also due to the size of the vehicle and its on-demand capabilities, it can also reduce overall fleet size or at minimum allow for better management based upon demand.

Retrofitting the existing parking facility with charging stations can provide another energy source for these new technologies. Off-line vehicles can be charging as other vehicles are in use. This is the way that early electric car fleets operated. So integrating a small fleet of vehicles can provide the more cost-effective option for transit applications. Each student took different approaches to how all of these issues were combined. Here are just a few visual examples of solutions. Each students unique solution could be developed into a full paper of its own, but the underlying consensus was that these small flexible new movement technologies can allow a new way to move around campus that complements walking, can resolve issues for persons with disabilities, provide a way for freight or delivery to occur on campus and provide options when needed for all of us. Sustainability occurs on many levels in transportation. Energy, time-efficiency, air-quality, land conservation, cost-effectiveness, and reduction of single occupancy vehicle use are the key points to take into consideration in defining transportation sustainability. Therefore it was viewed as a complementary system that would encourage walking. **(FIGURE 8)**

### Final Analysis Review

Each architecture students' specific analysis resulted in individual full power point presentations concerning each individual area of study showing in full detail the results of their work – all of this cannot be presented in this paper.
The important factor is the physical integration of all of the daily life needs for a

**FIGURE 8:** Southern Polytechnic State University student, Kira Melville Transfer station locations.

student, employee or visitor so that walkability and accessibility for all can go hand and hand is currently not available. Parking facilities have been designed successfully linking with mixed use such as at the University of Pennsylvania, Philadelphia, 2002, were a parking facility/grocery store provides for the multiple needs of walkers, bike riders and car users in a University campus setting. These synergies create greater feasibility for multi-tasking, reducing automobile trips and providing integrated solutions for better time management. Integrating mixed-use to develop fully functioning living environments is a good approach to reduce multiple trips by automobile for daily living. In Des Moines, IA, one peripheral parking facility is connections can provide easier access for persons of disability and other user groups such as the elderly. Linking transit with these mixed-use centers further expands the ability to serve wider population needs as even without an automobile a person may be able to function independently in a safe and accessible way connected with the bus system of the city, a day care center, dry cleaning, video rental and other typical daily/weekly living needs. This allows the user to consolidate daily/weekly needs in a time and cost effective way. Allowing these multiple connections can provide easier access for persons of disability and other user groups such as the elderly. Linking transit with these mixed-use centers further expands the ability to serve wider population needs as even without an automobile a person may be able to function independently in a safe and accessible way.

# Paper 460
# FROM CONGESTION TO RELIABILITY
# EXPANDING THE HORIZON

H. S. Levinson[1] and R. Margiotta[2]

[1] Herbert S. Levinson, 5305 Ashlar Village, Wallingford, CT 06492; PH (203) 949-9700; email: hslevinson@aol.com
[2] Richard Margiotta, 1265 Kensington Ave., Knoxville, TN 37922; PH (865) 670-8516; email: rmargiotta@camsys.com

## ABSTRACT

Traffic congestion has many dimensions including intensity, extent, duration and reliability.

- Conventional congestion analyses use Highway Capacity Manual demand-capacity-delay relationships for assessing performance of specific roadway locations or sections. Broader measures include applying differences or ratios between actual and desired travel times and speeds.

- Reliability brings a new set of metrics to roadway performance calculations. While typical measures deal with averages, reliability deals with the ranges and distributions of travel times. It focuses on estimating the travel time rates that are not likely to be exceeded.

- This paper gives some of the key metrics associated with reliability and shows how they differ from conventional congestion analysis. It draws upon and extends recent research efforts.

## INTRODUCTION

Traffic congestion has been long recognized as an important concern throughout urban America. It has been extensively studied in terms of location, intensity, extent, and duration. More recently, emphasis has been placed on a related aspect—travel time reliability. This paper overviews congestion measurement and gives some key metrics associated with estimating reliability.

## BACKGROUND

Much work has been done in estimating roadway performance. The Highway Capacity Manual (HCM 2000) sets forth demand-capacity-delay relationships for assessing service levels for specific freeway, arterial street locations and sections. Various congestion analyses, such as suggested in NCHRP 398, Quantifying Congestion (Lomax 1997) suggest broader measures that include applying differences or ratios between actual and desired speeds. More recently, SHRP 2, Project LO3 Analytical Procedures for Determining the Impacts of Reliability Mitigation

Strategies sets ways to quantify reliability and assess the impacts of various mitigation strategies (Cambridge Systematics 2010).

Reliability brings a new set of metrics to assessing roadway performance. While typical congestion/service level metrics deal with averages, reliability deals with the ranges and distributions of travel times.

This paper sets forth basic congestion measurement concepts. It then presents the key metrics associated with assessing reliability, drawing and building upon the recent research efforts.

## QUANTIFYING CONGESTION

Traffic congestion is defined as the travel time or delay in excess of that normally incurred under light or free-flow travel conditions. Unacceptable congestion is travel time or delay that exceeds an agreed upon norm. This norm can vary by type of transportation facility, geographic location and time of day. It also can be expressed as the decrease in travel speeds during busy traffic conditions as compared with free-flow conditions. It is expressed either as a difference or ratio.

Congestion can be recurrent—occur each day at about the same time in the same place or it can be non-recurrent—as experienced where incidents or inclement weather occur.

Highway Capacity Manual procedures are commonly used for assessing the performance of specific roadway sections or intersections. For freeways, speed-flow relationships are translated into densities (vehicles per lane per mile that are keyed to various levels of service. For signalized intersections, volume-capacity ratios and signal timings and cycle lengths help quantify delays; in both cases, service levels ranging from A to F are defined.

Various congestion measures have emerged to quantify congestion consistently and easily. These measures have several goals (Levinson 1996).

- Provide information on the magnitude and range of congestion in severely congested operating conditions.

- Provide values that are easy to use and understand. The use of a continuous numerical scale remedies some of the problems associated with the discrete letter grades of HCM procedures, where small changes in volumes or delay estimates can change LOS designations where they occur near a boundary.

- Permit local and State agencies to customize unacceptable congestion by location, facility type or travel mode.

- Readily communicate findings to planners, policy makers, and the general community.

A speed reduction index keyed to the percent reduction in free-flow is shown in Figure 1. The index, set forth in NCHRP 398 (Lomax, 1997) represents the percent speed reduction divided by ten. Thus, a 50% speed reduction has an index of 5.

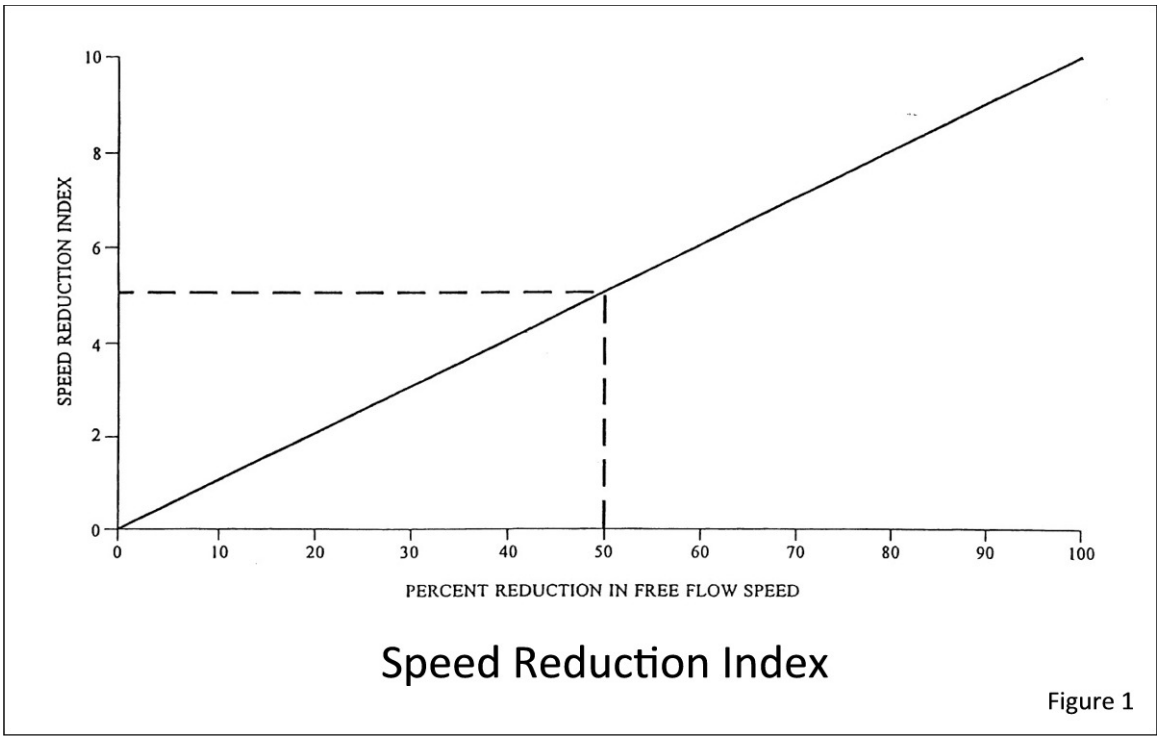



Figure 1

A modification of this concept led to the widely accepted travel time index (TTI). This index is derived from the ratio of actual travel times rates to free-flowing travel rates both expressed in minutes per mile. (Equation 1)

$$TTI = \frac{Minutes/Mile\ Actual}{Minutes/Mile\ Free-Flowing} \quad (1)$$

Commonly, a free-flow rate of one minute per mile is used for freeways. In this case, the TTI becomes the actual delay-rate (minutes per mile experienced).

Most urban freeways have travel time indices of less than 1.5. Few exceed indices of more than 2.5.

## ASSESSING RELIABILITY

Travel time reliability brings a new perspective to roadway performance and congestion analysis. It deals with the distribution and range of travel times along a given section of roadway. It can be defined as how travel times vary over time. It also can be defined in terms of the probability of failure, where failure represents traffic flow breakdown.

Reliability metrics build upon typical statistical measures such as the mean, median, and various percentiles, as well as the standard deviation and variance. The probability of failure can be expressed in percentiles. A 95 percentile travel time rate is associated with a 5 percent failure; an 80 percentile travel time rate is associated with a 20 percent failure.

The reliability metrics recommended in SHRP 2 Project LO3 (Cambridge Systematics 2010) are shown in Table 1. Of these, the 95 percentile travel time index (the same as the as percentile travel time when the free-flow speed is 60 mph) is especially significant. The 80$^{th}$ percentile is useful for assessing effects of improvement strategies.

**Table 1**

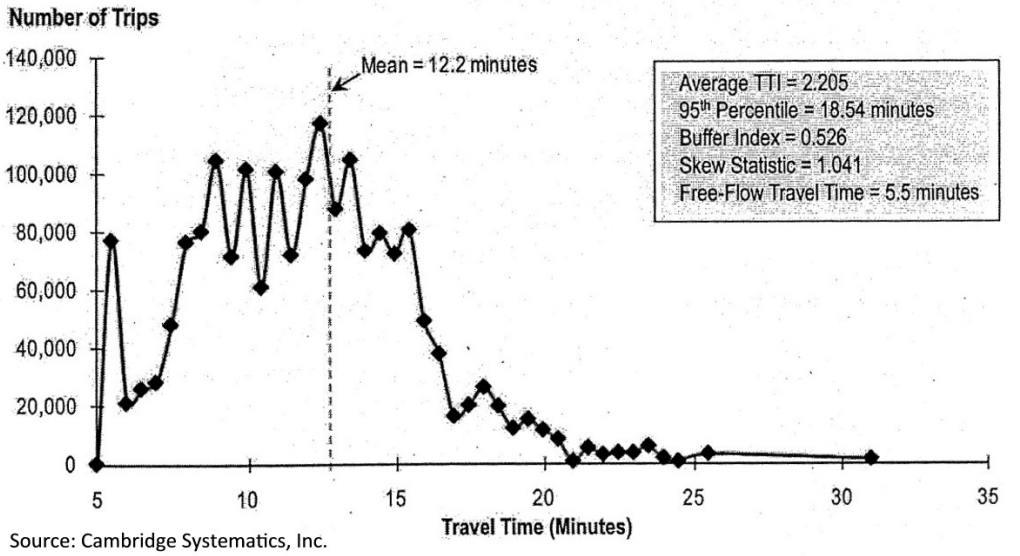
**RECOMMENDED RELIABILITY METRICS**

| Reliability Performance Metric | Definition | Units |
|---|---|---|
| Buffer Index (BI) | The difference between the 95$^{th}$ percentile travel time and the average travel time, normalized by the average travel time.<br>The difference between the 95$^{th}$ percentile travel time and the median travel time, normalized by the median travel time. | Percent |
| Failure/On-Time Measures | Percent of trips with travel times <:<br>• (1.1 * Median Travel Time); and<br>• (1.25 * Median Travel Time).<br>Percent of trips with space mean speed <:<br>• (50 mph, 45 mph, 30 mph). | Percent |
| Planning Time Index | 95$^{th}$ percentile Travel Time Index. | None |
| 80$^{th}$ Percentile Travel Time Index | Self-explanatory. | None |
| Skew Statistic | The ratio of (90$^{th}$ percentile travel time minus the median) divided by (the median minus the 10$^{th}$ percentile). | None |
| Misery Index (Modified) | The average of the highest five percent of travel times divided by the free-flow travel time. | None |

SOURCE: Cambridge Systematics, Inc.

The travel times along sections of uncongested freeways usually are normally distributed. Knowing the mean, standard deviation, and coefficient of variation makes it possible to readily compute percentiles, and reliability metrics. However, as travel time rates increase, there is even greater increase in the measures of dispersion.

The shape of travel time distribution along congested freeway sections is much broader. As shown in Figure 2 for I-75 in Atlanta, the long-tail to the right of the mean suggests a log-normal distribution.




Peak Hour Travel-Time Distribution
Atlanta, I-75 NB
2007

Figure 2

Detailed relationships between the average and both 88 and 95 percentile travel times were developed for freeway, sections of I-20, I-285, I-75, and I-85 in Atlanta. The results are shown in Figure 3.

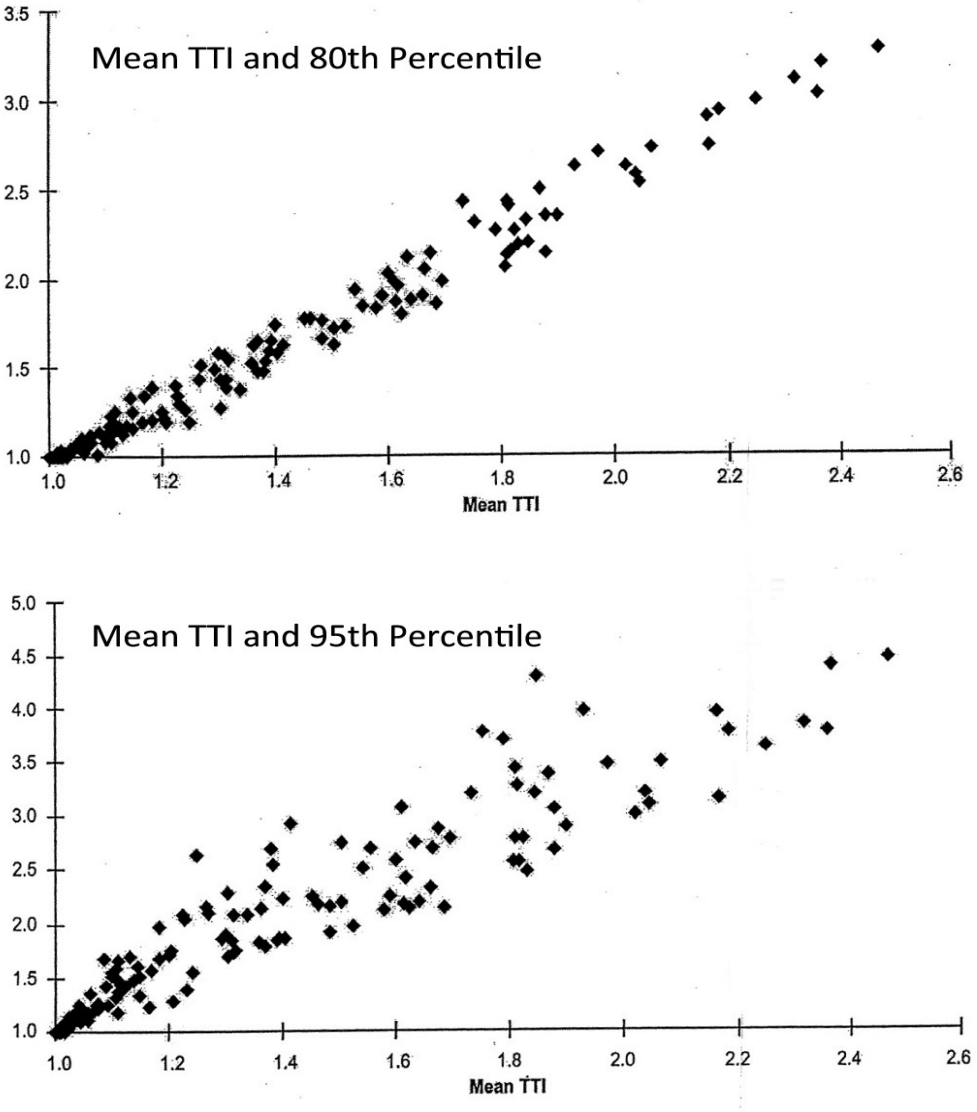



Figure 3

The mean travel time index (that is the same as the mean travel time) ranges from 1 to 2.5. The 80% value ranges from 1 to 3.5. The 95% value ranges from about 1 to 4.5, and has a more dispersed pattern.

The exponential statistical models developed by SHRD 2 Project LO3 for both the 80 and 95 percentile for freeway sections in six cities were as follows:

$$95^{th} \text{ Percentile TTI} \cong \text{Mean TTI}^{1.8834} \cong \text{Mean TTI}^{1.88} \quad \text{Multi-city data} \quad (1)$$

$$80^{th} \text{ Percentile TTI} \cong \text{Mean TTI}^{1.365} \cong \text{Mean TTI}^{1.36} \quad (2)$$

The equations developed for the Atlanta freeways were:

$$93^{rd} \text{ Percentile TTI} \cong \text{Mean TTI}^{1.6950}$$

$$80^{th} \text{ Percentile TTI} \cong \text{Mean TTI}^{1.3162}$$

An independent linear regression analysis was prepared for a sample of peak period observations for 12 Atlanta Freeway sections. Six to seven data points were analyzed for each travel direction for each 15-minute period. There were 83 individual observations. The predictive equation was as follows:

$$95^{th} \text{ Percentile min/mile} = 1.717 \frac{\text{min}}{\text{mile}} - 0.389 \quad r^2 = 0.780 \quad (3)$$

$$\cong 1.7 \text{ min/mile} - 0.4$$

Where freeway traffic is not congested (e.g. TTI 1.0 to 1.2 minutes per mile) the coefficients of travel time variations approximate 15%. The 80 and 95 percentile values can be computed using the normal distribution. The 95 percentile travel time can be estimated as follows:

$$95^{th} \text{ Percentile} = x + 1.645\,\sigma \quad (4)$$

Where $\bar{x}$ = mean minutes per mile

$\sigma$ = standard deviation

$\frac{\sigma}{\bar{x}} = \bar{x}$ (coefficient of variation)

The equation becomes:

95 percentile = x (1 + 1.645 cv) (5)

An assumed 10% coefficient of variation (cv) translates into a 95 percentile of 1.16 times the mean. A 15% cv becomes 1.25.

Table 2 compares the 95% travel time indices obtained from the various models. The equations apply up to a TTI value of 2.5. The linear model is consistently higher than the use of the normal distribution statistics. It is about the same as the exponential model for travel time indices between 1.2 and 1.5 minutes per mile, but lower for higher average rates. The exponential models understate variability for exponential travel time rates less than 1.2 minutes per mile, and therefore should not be used below this threshold.

Table 2 also shows a modified exponential model that falls in between the linear and exponential equations. This model takes the form $ax^b$ where x is the TTI index. It assumes a = 1.2 and b = 1.5. The model has a 95 percentile value that is 1.20 when the TTI index is 1.0 and a value of 4.33 when the TTI index is 2.5.

**Table 2**
**COMPARATIVE 95 PERCENTILE TRAVEL TIME INDICES**

| MODEL | MINUTES/MILE OR TTI | | | | |
|---|---|---|---|---|---|
| | **1.0** | **1.2** | **1.5** | **2.0** | **2.5** |
| Normal Distribution | | | | | |
| 10% cv | 1.16 | 1.39 | 1.74* | 2.32* | 2.90* |
| 15% cv | 1.25 | 1.50 | 1.88* | 2.50* | 3.12* |
| Exponential Model | | | | | |
| Atlanta | 1.00 | 1.26 | 1.99 | 3.24 | 4.73 |
| Multi-City | 1.00 | 1.41 | 2.15 | 3.69 | 5.62 |
| Linear Model | 1.30 | 1.64 | 2.15 | 3.00 | 3.85 |
| Modified Exponential Model | 1.20 | 1.55 | 2.11 | 3.17 | 4.33 |

*Generally not applicable.
SOURCE: Exponential Model, Cambridge Systematics; other models computed.

## IMPLICATIONS AND EXTENSION

The concept of travel time reliability extends the concept of congestion to look at "worst case" conditions. The 95 percentile travel time index value provides a reasonable estimate for "worst case" recurred congestion. Predictive models have been derived for estimating these values for freeways, once the "average" conditions are known.

Ongoing strategic highway research by the Transportation Research Board is analyzing how reliability metrics can be applied to the 2010 Highway Capacity. This research will likely include estimating service levels for various failure rates.

Logical next steps in reliability analyses include estimating worst case conditions for signalized intersections assessing the effects of non-recurrent congestion.

## REFERENCES

Highway Capacity Manual, 2000, Transportation Research Board, Washington, DC.

Cambridge Systematics, Inc. with Texas A&M University, University of Washington, Dowling Associates, Street Smarts, H.S. Levinson and H. Rakim. SHRP 2, Project LOS Analytical Procedures for Determining the Impacts of Reliability Mitigation Strategies, Strategic Highway Research Program 2, Transportation Research Board, Washington, DC, February 2010.

Levinson, H.S. and T.S. Lomax "A Travel Time Based Index for Evaluating Congestion" in Compendium of Technical Papers for the 66th Annual Meeting, Institute of Transportation Engineers, Minneapolis, Minnesota, September 15-19, 1996.

Lomax T., S. Turner and G. Shunk with H.S. Levinson, R. H. Pratt, P. Bay and G. B. Douglas NCHRP Report 398 – Quantifying Congestion, Transportation Research Board, National Research Council, Washington, DC, 1997

# AN EMPIRICAL INVESTIGATION OF MOTORCYCLE OWNERSHIP AND USAGE

Konstantinos Kepaptsoglou, Ph.D., M.ASCE[1], Christina Milioti[2], Zoi Christoforou[3], Matthew G. Karlaftis, Ph.D., M.ASCE[4].

[1]Department of Transportation Planning and Engineering, School of Civil Engineering, National Technical University of Athens, 5 Iroon Polytechniou str, Zografou Campus, Athens, Greece, 15773, tel. +30-210-7723167, fax +30-210-772-2404, e-mail: kkepap@central.ntua.gr
[2]Department of Transportation Planning and Engineering, School of Civil Engineering, National Technical University of Athens, 5 Iroon Polytechniou str, Zografou Campus, Athens, Greece, 15773, tel. +30-210-7723167, fax +30-210-772-2404, e-mail: christinamilioti@hotmail.com
[3]Department of Transportation Planning and Engineering, School of Civil Engineering, National Technical University of Athens, 5 Iroon Polytechniou str, Zografou Campus, Athens, Greece, 15773, tel. +30-210-7723167, fax +30-210-772-2404, e-mail: zoic@civil.ntua.gr
[4]Department of Transportation Planning and Engineering, School of Civil Engineering, National Technical University of Athens, 5 Iroon Polytechniou str, Zografou Campus, Athens, Greece, 15773, tel. +30-210-7721280, fax +30-210-772-2404, e-mail: mgk@central.ntua.gr

## ABSTRACT

Motorcycles are a widespread mode of travel in several parts of the world as, for example, in Southern Europe and Asia. However, motorcycle ownership and usage have been scarcely examined, despite being critical for developing transportation and traffic safety plans and policies in urban areas with high motorcycle presence. This pap dels are developed for this purpose based on questionnaires collected in Athens, Greece. A binary Probit model is used for analyzing motorcycle ownership and a hazard-based duration model for explaining factors affecting motorcycle usage. Results indicated that income, gender, age and er investigates factors affecting motorcycle ownership and use; appropriate mo spatial trip characteristic do have an effect on motorcycle ownership and usage.

## INTRODUCTION

Motorcycles are a widespread mode of travel in several parts of the world as, for example, Southern Europe and Asia. Reasons such as the lower purchase and operating costs for motorcycles, their size and maneuverability which facilitates movement and parking even under poor traffic conditions make powered two-wheelers very attractive, particularly to lower income households, professionals with increased mobility needs and younger individuals. For example, among European countries, Italy, Greece, Bulgaria and the Czech Republic exhibit motorcycle ownership rates of over 150 vehicles per 1000 passengers, while the average EU rate is about 20 vehicles per 1000 inhabitants (DfT, 2007). In Asia, major urban areas such as Ho Chi Minh City, Jakarta, Taipei and Bangkok, over 200-300 motorcycles per 1000 inhabitants are in circulation (Acharya and Morichi, 2007).

Past research on motorcycle transportation has mostly focused on traffic safety (recent papers have been published by Paulozi (2005), the National Highway Safety Administration (2007), Savolainen and Mannering (2008), Gkritza (2009), Li *et al.* (2009), Oluwadiya *et al.* (2009) and Pai *et al.* (2009),) and traffic related issues (for example Spyropoulou and Sermpis, (2007), Yannis *et al.* (2007), Lee *et al.* (2010)). Interestingly, motorcycle ownership and usage has been examined to a far lesser extent, despite being critical for developing transportation and traffic safety plans and policies in urban areas with high two-wheeler usage.

Indeed, while car ownership and use have again been a topic of considerable research in the past (see for instance, Steg *et al.* (2001), Limtanakool *et al.* (2006), Clark (2007, 2009a, 2009b), Matas and Raymond (2008), Yang *et al.* (2008), Nollan (2010) and Van Acker and Witlox (2010), Ettema *et al.* (2010)), research on motorcycle ownership exhibits only a handful of studies. Duffy and Robinson (2004) offered an econometric analysis of motorcycle ownership in the UK and indicated a significant elasticity of demand for motorcycles with resp1ect to income. Tuan and Shimizu (2005) developed dynamic discrete choice models for analyzing motorcycle ownership behavior in Anoi, Vietnam and identified several factors affecting ownership, such as price, profession, income and taxation in motorcycle usage. A paper by Burge *et al.* (2007) investigated policy effects on motorcycle usage, based on extensive survey data from the UK; nested logit models were developed for explaining motorcycle ownership and usage in that country. Metin *et al.* (2007) analyzed ownership and use of motorcycles in Jabotabek, Indonesia, and found that socioeconomic and demographic characteristics have a positive effect on ownership. The survival process of motorcycle ownership (holding time) was examined by Chang and Yeh (2007) by using a split-population duration model. Finally, Chiu *et al.* (2009) developed an integrated model of car/motorcycle ownership, type and usage and used it for estimating energy consumption and emissions.

Investigation of choice behavior with respect to ownership and usage of motorcycles is necessary when that mode has a considerable share in the daily demand patterns for a city. In this context, this paper investigates factors affecting motorcycle ownership and use. Appropriate models are developed for this purpose based on a dataset collected in Athens, Greece. Two models are used in particular: a probit model which is used for analyzing ownership and a hazard-based duration model for explaining the factors that affect usage.

## DATASET

Data were collected through an extensive questionnaire based survey in the Athens metropolitan area. Over 8300 respondents participated in the survey, out of which 1086 owned a motorcycle and 285 were riding a motorcycle at the time of the survey. Gathered data included demographic information (gender, age, marital status, household size and location and income), information on vehicle ownership (number of cars and motorcycles per household). Furthermore, participants were asked about the origin and destination of their trip, their trip purpose (work, shopping, education, entertainment) and number of trips per week. Finally, the transportation mode and time period of each interview were noted. Geocodes were used for reporting household, origin and destination locations for each respondent. This way, it was possible to obtain time and distance estimations for all trips performed (Table 1 reports aggregated data per mode) and also to identify population density for each location. Table 2 summarizes collected information in the survey.

**Table 1. Average travel time and distance per transportation mode**

| Transportation Mode | Average distance (m) | Travel time(min) | |
|---|---|---|---|
| | | Mean | Standard Deviation |
| Walk | 931 | 9 | ±5 |
| Motorcycle | 4,343 | 16 | ±11 |
| Car | 5,003 | 23 | ±13 |
| Metro | 4,604 | 27 | ±13 |
| Buses and tram | 4,049 | 33 | ±14 |
| Taxi | 5,111 | 27 | ±13 |

**Table 2. Summary of collected data**

| Type | Definition – Values | Collection method |
|---|---|---|
| Gender | Male, Female | Survey |
| Age | <19, 20-34, 35-64, >64 | Survey |
| Household size | Number of household members | Survey |
| Household location | Appropriate geocode | Survey |
| Car ownership | Number of cars owned by household | Survey |
| Motorcycle ownership | Number of motorcycles owned by household | Survey |
| Income | low, medium, high and very high | Survey |
| Origin, destination | Appropriate geocode | Survey |
| Trip purpose | Work, education, entertainment, shopping | Survey |
| Transportation mode | Walk, car, motorcycle, metro, bus, tram, taxi | Survey |
| Travel distance | Distance travelled using mode, in meters | Derived |
| Travel time | Travel time, in min | Derived |

## METHODOLOGY

As noted in an earlier section of the paper, two models are considered in the modeling effort. Motorcycle ownership is modeled through a probit model, while for motorcycle usage (in terms of overall travel time when using a motorized two-wheeler), a hazard-based duration model is used.

**Probit Models**

Probit models have been widely used to analyze dependent variables of discrete nature (0,1). The relationship among the dependent and the independent variables does not lead to the estimation of a value for the dependent variable, but to the estimation of a probability that one of the two alternatives will occur (Washington *et al.*, 2010). Here, we use Probit models to estimate the factors that affect ownership of a motorcycle. In this sense, increased probability of a factor would indicate that the latter contributes to motorcycle ownership.

The general specification for a univariate Probit model (for an event n resulting in an outcome i) can be expressed as (Washington *et al.*, 2010):

$$Y_{in} = \beta_i \cdot X_{in} + \varepsilon_{in} \quad (1)$$

where $Y_{in}$ defines an unobserved variable representing the latent utility (or propensity) for alternative $i$, $X_{in}$ is a vector of observed characteristics determining the outcome of the event $n$, $\beta_i$ represents a vector of unknown coefficients to be estimated for the alternative $i$, and $\varepsilon_{in}$ represents a vector of error terms.

If further assumed that $\varepsilon_{in}$ follows the normal distribution and that $i=1,2$, we obtain the specification of the binomial Probit model (Washington *et al.*, 2010):

$$\left[ P_n(1) = P(\beta_1 \cdot X) \right]_{1n} - \beta_2 \cdot X_{2n} \geq -\varepsilon_{1n} + \varepsilon_{2n} \quad (2)$$

Equation (2) estimates the probability of occurrence of alternative 1 for event $n$. The terms $\varepsilon_{1n}$, $\varepsilon_{2n}$ are normally distributed with zero mean and variances $\sigma_{12}$, $\sigma_{22}$. Anybody who own or does not own a motorcycle can be regarded as an event, having as an outcome motorcycle ownership or not. The binomial Probit model of Equation (2) can be used for estimating factors contributing to motorcycle ownership. Under this assumption, equation (2) provides the probability for motorcycle ownership.

**Duration Models**

Hazard based duration models are commonly used for describing the elapsed time or the duration of an event; examples in the field of transportation include modeling of the duration of trip making activities, travel time, and so on (Washington *et al.*, 2010). Following Washington et al. (2010), let $T$ be a nonnegative random variable that represents travel time with a motorcycle. The probability distribution of $T$ can be represented in a number of ways, of which the survival and hazard functions are the most useful. The *survival function* is defined as the *probability* that $T$ is of length *at least* $t$ (i.e. travel time with motorcycle of at least $t$ min) and is given by (Washington et al., 2010):

$$F(t) = P(T \geq t), \quad 0 < t < \infty \quad (3)$$

The notation used here suggests that $F(t)$ is a monotone left continuous function with $F(0) = 1$ and $\lim_{t \to \infty} F(t) = 0$. The probability density function (p.d.f.) of $T$ is (Washington et al., 2010):

$$f(t) = \lim_{\Delta t \to 0^+} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = \frac{-dF(t)}{dt} \quad (4)$$

The *hazard function* specifies the instantaneous failure rate at $T = t$, conditional upon survival to time $t$, and can be defined as follows (Washington et al., 2010):

$$\lambda(t) = \lim_{\Delta t \to 0^+} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{F(t)} \quad (5)$$

It is important to note that hazard functions are extremely useful in practice. They indicate the rate at which perceived waiting time increases after lasting for time $t$ and

for this reason is more interesting than the survival or the c.d.f. functions. Also, from Equation (7) it can be seen that $\lambda(t)$ specifies the *distribution* of *T* since,

$$\lambda(t) = -\frac{d \log F(t)}{dt}$$

by integrating and setting *F(0) = 1*

$$F(t) = \exp\left\{-\int_0^t \lambda(u)du\right\}$$

and the p.d.f. of *T* becomes

$$f(t) = \lambda(t)\exp\left\{-\int_0^t \lambda(u)du\right\} \quad (6)$$

The literature suggests a wide variety of functional forms for the duration distributions such as the exponential, the Weibull, the Lognormal, the inverse normal, the Loglogistic, and others (Washington *et al.*, 2010). Interestingly, these distributions display very different behaviors and the selection of the functional form to be used will have important implications in the practical significance of the results.

## RESULTS AND EMPIRICAL FINDINGS

### Motorcycle Ownership

Probit model results for motorcycle ownership are shown in Table 3. The dependent variable in this case (MOTO), takes the value of 1 for motorcycle owners and 0 otherwise.

**Table 3. Probit model results for motorcycle ownership.**

| Variable Name | Type | Variable Description | Model parameters | |
|---|---|---|---|---|
| | | | Coeff. | t-stats |
| *Dependent Variable* | | | | |
| MOTO | binary | =1 if motorcycle ownership; =0 otherwise | - | - |
| *Explanatory Variables** | | | | |
| Constant term | - | - | -0.856 | -6.35 |
| GENDER | binary | =1 if men; =2 if women | -0.312 | -4.72 |
| AGE2 | dummy | =1 if respondents are between 20-34 years old; =0 otherwise | 0.588 | 8.47 |
| AGE4 | dummy | =1 if respondents are between 35-64 years old; =0 otherwise | -1.973 | -9.14 |
| SINGLE | dummy | =1 if women and single; =0 otherwise | -0.988 | -8.36 |
| CAR_OWNERSHIP | ordinal | Number of cars in the household (1-3) | -0.301 | -6.79 |
| LOWINC | dummy | =1 if income is low; =0 otherwise | 0.256 | 3.147 |
| SUBPDM | dummy | 1 if respondents live in suburbs and population density is low or medium; =0 otherwise | -0.329 | -2.64 |
| ODLOWHIGH | dummy | =1 if population density in the origin of the trip is low or medium and in the destination is high or very high; =0 otherwise | 0.489 | 3.39 |
| DISTANCE | continuous | Travel Distance | -0.009 | -3.84 |
| Number of observations | | | | 8348 |
| Log-likelihood with constant only LL(c) | | | | -3290.999 |
| Log-likelihood at convergence LL(β) | | | | -2704.681 |

*non-significant variables are omitted

Based on Table 3 results, motorcycle ownership seems to be higher for men (variable GENDER). Further, being a single woman has a strong negative effect towards motorcycle ownership (variable SINGLE). Also, individuals aged between 20 and 34 are associated with higher motorcycle ownership, while ages between 35 and 64 seem to go on a different direction. Possible explanations for these findings include:

1. Risk and safety perception associated with motorcycles, is probably lower for men and younger ages.
2. Social attitude and image for women driving motorcycles may, in some cases, be a discouraging factor for owning one.
3. Younger ages are expected to prefer lower cost transportation modes such as motorcycles.
4. The task of motorcycle driving is probably more attractive to men and younger adults.

Car ownership (variable CAR_OWNERSHIP) does have a negative effect on motorcycle ownership, which is an expected finding since it indicates households with a stronger preference towards (or dependence on) cars and higher income. On the contrary, lower income households (variable LOWERINC) tend to prefer modes of lower acquisition value and operating costs, thus the positive relation with moto ownership probability. Finally, with respect to trip characteristics, the following are observed:

1. Travel distance has a negative effect on motorcycle ownership (variable DISTANCE). Indeed, factors such as comfort, reduced traffic safety and limited protection from weather conditions could discourage motorcycle ownership, particularly when the motorcycle is expected to be used for longer distances.
2. In suburbs and areas of lower population density (variable SUBPDM), traffic conditions are expected to be better, incomes are higher in general, and travel distances are longer. These factors, associated with living in Athens suburbs, could negatively affect motorcycle ownership.
3. However, when (a) a trip destination is of higher population density (variable ODLOWHIGH), which implies poor traffic conditions and limited parking space, (b) a trip origin is of lower density, which may suggest lower coverage by public transportation and largest dependence on private modes, motorcycle ownership and usage is attractive to travelers.

Overall, it seems that lower costs of owning a motorcycle, a lower perception of the risk and the difficulty associated with its driving task, as well as reasons related to its advantages as a transportation mode, do affect motorcycle ownership in the Athens metropolitan area.

**Motorcycle Usage**

Results for motorcycle usage, by exploiting duration models with Weibull and Loglogistic distributions are presented in Table 4. The dependent variable in this case is the travel time (variable TTIME), representing driving time with a motorcycle. It should be noted here that the hazard functions for the loglogistic and Weibull distributions are given by $l\,p(l\,t)^{p-1}/1+(l\,t)^{p}$, $l\,p(l\,t)^{p-1}$ respectively; for the same distributions the survival functions are $1/1+(l\,t)^{p}$, $e^{-(l\,t)^{p}}$. It is evident, from Table 3, that the different functional forms of the Weibull and Lognormal distribution lead to very different qualitative conclusions. For example, the Weibull distribution is monotonically increasing indicating a continuously increasing hazard rate over time,

while the Loglogistic suggests an initial increase and then a decrease in the probability (Washington *et al.*, 2010). The obvious question then becomes how can the "best" fitting distribution be selected. Besides theoretical arguments, the statistical answer to this question is not straightforward. In general, for a model to be appropriate for the data the graph for each of the functional forms needs to be a straight line through the origin (for the exponential model, for example, it is the graph of the log of the survival versus *t*). However, it is interesting to note that the Weibull and log-logistic functional forms are all *nested* within the generalized gamma model, making it a simple matter to evaluate them with the likelihood ratio test. In the present case, likelihood ratio test results are also presented in Table 4. As can be seen, the value or $\rho^2$ is higher for the Weibull distribution, and therefore, results for the Weibull distribution are considered more appropriate for representing motorcycle usage.

**Table 4. Duration model results for motorcycle ownership**

| Variable Name | Type | Description | Distribution | | | |
|---|---|---|---|---|---|---|
| | | | Weibull | | Loglogistic | |
| | | | Coeff. | t-stats | Coeff. | t-stats |
| Constant term | - | - | 2.776 | 20.72 | 2.659 | 38.27 |
| LON7 | dummy | =1 if number of persons in the house <number of cars in the household; =0 otherwise | -0.390 | -2.51 | -0.267 | -1.92 |
| EVENING | dummy | =1 if time of the trip is between 18.00-22.00; =0 otherwise | -0.196 | -2.20 | - | - |
| DENSITY | ordinal | Population density of the destination area 1 (lower) – 4 (higher) | 0.054 | 1.65 | - | - |
| LOWINC | dummy | =1 if income is low; =0 otherwise | 0.177 | 2.35 | - | - |
| FREQUENCY | ordinal | Frequency of trips (1-5) | -0.720 | -1.77 | -0.947 | -3.08 |
| WORKEDU | dummy | =1 if trip purpose is work or education; =0 otherwise | 0.437 | 2.93 | 0.565 | 4.83 |
| Lambda | | | 0.058 | | 0.076 | |
| P | | | 1.732 | | 2.828 | |
| Number of observations | | | 285 | | | |
| Log-likelihood with constant only LL(c) | | | -293.8321 | | -282.5189 | |
| Log-likelihood at convergence LL(β) | | | -235.4139 | | -237.2974 | |
| $\rho_c^2$=(LL(c)-LL(b))/LL(c) | | | 0.248151 | | 0.1905689 | |

Model results for the Weibull distribution indicate that motorcycle usage is increased for people having work or education as a trip objective (variable WORKEDU). The same applies for trip destinations of lower density (variable DENSITY) and low income respondents (variable LOWINC). Indeed, advantages of motorcycles such as their maneuverability in traffic, limited parking requirements (motorcycles in Athens virtually park everywhere) and low operating costs may explain the effects of these factors. On the contrary, excessive car ownership (more cars that people in a household- variable LON7), which is a sign of higher income level and a tendency towards preferring cars, does have a negative effect on motorcycle usage. Interestingly, frequency of trips affects usage in a negative manner, implying that when motorcycles are used often, related trips are of a lower duration. Lower usage in the evening (variable EVENING) on the other hand may be

interpreted by increased traffic safety concerns at that time of the day. Finally, it is possible that the high constant term value accounts for unobservable factors that may have a considerable effect on motorcycle usage.

## CONCLUSIONS

Motorcycles are a popular transportation mode in both developed and developing countries of the world, having advantages that make the attractive to large population groups. Furthermore, they do have an impact to traffic conditions and safety, especially in urban areas. In this paper, appropriate econometric models were used in an effort to investigate factors affecting ownership and usage of motorcycles. Results indicated that income, gender, age and spatial trip characteristic do have an effect in motorcycle ownership and usage, with possible interpretations related with safety associated with motorcycle usage, lower acquisition and operating costs and advantages related to the mode. Future work should focus incorporating more explanatory factors to the model (for example weather conditions and also collect traveler perception on motorcycles, which would further validate interpretations drawn in the current study.

## REFERENCES

Acharya, S.R., Morichi, S. (2007). Motorization and role of Mass Rapid Transit in East Asian Megacities. *IATSS Research*, vol. 31, no.2, pp. 6-16.

Burge, P., Fox, J.B., Kouwenhoven, M., Rohr, C., Wigan M-R. (2007). Modeling of Motorcycle Ownership and Commuter Usage: A UK study. *Transportation Research Record, Journal of the Transportation Research Board*, no. 2031, pp. 59-68.

Chang, H-L., Yeh, T-S. (2007). Exploratory analysis of motorcycle holding time heterogeneity using a split-population duration model. *Transportation Research Part A*, vol. 41, pp. 587-596.

Chiou, Y-C., Wen, C-H., Tsai, S-H., Wang, W-Y. (2009). Integrated modeling of car/motorcycle ownership, type and usage for estimating energy consumption and emissions. *Transportation Research Part A*, no. 43, pp. 665-684.

Choo, S., Mokhtarian, P. (2008). How do people respond to congestion mitigation policies? A multivariate probit model of the individual consideration of three travel-related strategy bundles. *Transportation*, vol. 35, no 2, pp. 145-163.

Clark, S.D. (2009). Characterising and predicting car ownership using rough sets. *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 4, pp. 381-393.

Clark, S.D. (2007). Estimating local car ownership models. *Journal of Transport Geography*, vol. 15, no. 3, pp. 184-197.

Clark, S.D. (2009). The determinants of car ownership in England and Wales from anonymous 2001 census data. *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 5, pp. 526-540.

Department for Transport (DfT) (2007). Compendium of Motorcycling Statistics - 2007 Edition, UK, *Transport Statistics Bulletin*.

Duffy, M., Robinson, T. (2004). An econometric analysis of motorcycle ownership in the UK. *International Journal of Transport Management*, vol. 2, pp. 111-121.

Ettema, D., Bayarma, A., van Hagen, M. (2010). Spending time on the move: A comparison between travel modes. *Proceedings of the 89th Transportation Research Board*, Washington DC, U.S.A.

Gkritza, K. (2009). Modeling motorcycle helmet use in Iowa: Evidence from six roadside observational surveys. *Accident Analysis & Prevention*, vol. 41, no. 3, pp. 479-484.

Goulias, K., W. Brog, and E. Erl. (1985) Perceptions on Mode Choice Using Situational Approach. *Transportation Research Record, Journal of the Transportation Research Board*, no 1645, pp. 82-93.

Lee, T-C., Polak, J.W., Bell, M.G.H., Wigan, M.R. (2010). The PCU Value of Motorcycles in Congested Flows. *Proceedings of the 89th Transportation Research Board*, Washington DC, U.S.A.

Li, M-D., Doong, J-L., Huang, W-S., Lai, C-H., Jeng, M-C. (2009). Survival hazards of road environment factors between motor-vehicles and motorcycles. *Accident Analysis & Prevention*, vol. 41, no. 5, 938-947

Limtanakool, N., Dijst, M., Schwanen, T. (2006). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. *Journal of Transport Geography*, vol. 14, no. 5, pp. 327-341.

Matas, A., Raymond, J-L. Changes in the structure of car ownership in Spain, *Transportation Research Part A: Policy and Practice*, 42(1),187-202, 2008

National Highway Traffic Safety Administration (2007). Recent Trends in Fatal Motorcycle Crashes: An Update. *Annals of Emergency Medicine*, vol. 50, no. 2, pp. 193-194.

Nollan, A. (20100. A dynamic analysis of household car ownership. *Transportation Research Part A: Policy and Practice*, vol. 44, no. 6, pp. 446-455.

Oluwadiya, K.S., Kolawole, I.K., Adegbehingbe, O.O., Olasinde, O.O., Agodirin,O., Uwaezuoke, S.C. (2009). Motorcycle crash characteristics in Nigeria: Implication for control. *Accident Analysis & Prevention*, vol. 41, no. 2, pp. 294-298.

Pai, C-W, Hwang, K.P., Saleh, W. (2009). A mixed logit analysis of motorists' right-of-way violation in motorcycle accidents at priority T-junctions. *Accident Analysis & Prevention*, vol. 41, no. 3, pp. 565-573.

Paulozzi, L. (2005). The role of sales of new motorcycles in a recent increase in motorcycle mortality rates. *Journal of Safety Research*, vol. 36, no. 4, pp. 361-364.

Savolainen, P., Mannering, F. Effectiveness of Motorcycle Training and Motorcyclists' Risk Taking Behavior (2007). *Transportation Research Record, Journal of the Transportation Research Board*, no. 2031, pp. 52-58.

Senbil, M., Zhang, J., Fujiwara, A. (2007). Motorcycle Ownership and Use in Metropolitan Area of Jabotabek, Indonesia. *Proceedings of the 86th Transportation Research Board*, Washington DC, U.S.A.

Spyropoulou, I., Sermpis, D. (2007). Parameters related to modeling motorcycle movement. *Proceedings of the 86th Transportation Research Board*, Washington DC, U.S.A.

Steg, L., Geurs, K., Ras, M. (2001). The effects of motivational factors on car use: a multidisciplinary modeling approach. *Transportation Research Part A: Policy and Practice*, vol. 35, no. 9, pp. 789-806.

Tuan, V.A., Shimizu, T. (2005). Modeling of household motorcycle ownership behaviour in Hanoi city. *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 6, pp. 1751-1765.

Van Acker, V. , Witlox, F. (2010). Car ownership as a mediating variable in car travel behaviour research using a structural equation modeling approach to identify its dual relationship, *Journal of Transport Geography*, vol. 18, no. 1, pp. 65-74.

Viswanathan, K., K. Goulias, Jovanis, P. (2000). Use of traveler information in the Puget Sound Region. *Transportation Research Record, Journal of the Transportation Research Board*, no 1719, pp. 94-102.

Washington, S., Karlaftis, M., Mannering, F. (2010). *Statistical and Econometric Methods for Transportation Data Analysis – 2nd Edition*, Chapman & Hall / CRC Press, Boca Raton, FL, USA.

Yang, Z., Chen, G., Yu, B. (2008). Car ownership level for a sustainability urban environment. *Transportation Research Part D: Transport and Environment*, vol. 13, no 1, pp. 10-18.

Yannis, G., Golias, J., Spyropoulou, I., Papadimitriou, E. (2007). Mobility Patterns of Motorcycle and Moped Riders in Greece. *Transportation Research Record, Journal of the Transportation Research Board*, no. 2031, pp 69-75.

# SLIP RAMP SPACING DESIGN FOR TRUCK ONLY LANES USING MICROSCOPIC SIMULATION

Ghulam H. Bham[1], Manoj Vallati[2]

[1] Civil, Architectural and Environmental Engineering, Missouri University of Science and Technology (Missouri S&T), Rolla, Missouri 65409; PH: (573) 341-6286; email: ghbham@mst.edu
[2] Mechanical and Aerospace Engineering, Missouri University of Science and Technology (Missouri S&T), Rolla, Missouri 65409; PH: (573) 341-6539; email: mkv359@mail.mst.edu

## ABSTRACT

For the Corridors of the Future project, slip ramps will provide access to trucks between the proposed truck-only lanes (TOLs) and general purpose lanes (GPLs) for trucks to exit using at-grade interchanges. This paper proposes recommendations for slip ramp spacing between the TOLs and the general GPLs along Missouri rural interstate highways using a microscopic simulation model, VISSIM. Simulation of peak period traffic conditions indicated that heavy vehicle speeds were directly proportional to the lengths of the merge, diverge, and link sections. The proposed design recommendations for slip ramp spacing are based on the results of these section lengths. Central composite design was used to generate the cases required for simulation. To ascertain the findings of the simulation study, Level-of-Service for the segments on GPL was determined. This paper provides practitioners and state Departments of Transportation with design recommendations for slip ramp spacing and lengths of merge, link, and diverge for the Corridors of the Future project.

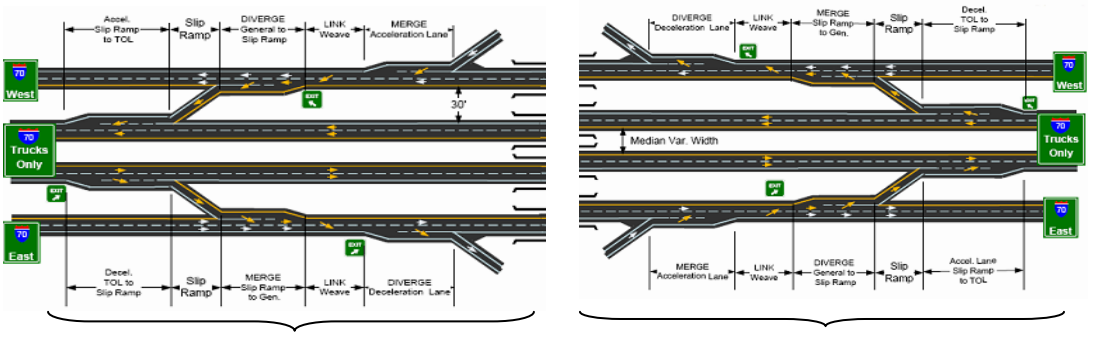
## INTRODUCTION

Truck-only lanes (TOLs), shown in Figure 1, are lanes designated exclusively for heavy vehicles to separate them from other vehicles to enhance safety and improve traffic flow. TOLs have proved to enhance safety by reducing conflicting movements and are cost effective when truck volumes are higher than 30 percent (Burke,2008). The Missouri Department of Transportation (MoDOT) has proposed TOLs along I-70 and I-44 to accommodate the high percentage of truck traffic and to minimize congestion along this freight-intensive corridor. Similar studies are also being conducted in other states. California performed a similar study and truck-only lanes were constructed on certain sections of I-5 in Los Angeles and Kern Counties.

In Missouri, in urban areas with intense truck activity grade-separated truck interchanges will be constructed, however, for most of the 250 mile rural corridor slip ramps will connect TOLs with GPLs for heavy vehicles to exit and enter the highway using general traffic interchange. Figure 1 shows a rural location served by slip ramps.

For efficient movement of heavy vehicles during peak hours, this paper proposes specific design lengths for the sections from the slip ramp to the general traffic interchange to exit the highway on level grades.

a) Proposed truck-only lane slip ramp configuration




Part 1 Part 2

**FIGURE 1. Schematic: truck-only lanes and general purpose lanes**

## LITERATURE REVIEW

Current publications do not address the spacing of slip ramps for TOLs. The AASHTO Design Guide (AASHTO, 2004) does not provide specifications for the design of slip ramps used exclusively by heavy vehicles. Previous studies have focused on the design of high occupancy vehicle (HOV) lanes and on truck lane restrictions. The studies reviewed below provide the details of ramp design and explains the selection of simulation parameters for use with VISSIM.

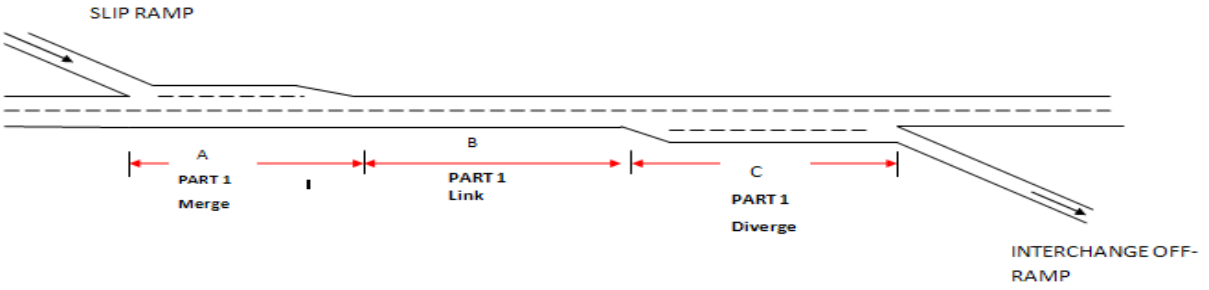
In Pasadena, California, VISSIM model was calibrated using field data for a 15-mile stretch of I-210 WB. This study simulated a freeway site consisting of three interacting bottlenecks and 20 metered on-ramps with and without HOV bypass lanes. The parameters in VISSIM were calibrated based on iterative runs and qualitative aspects of freeway operations such as initial and final times for mainline queues, the extent of queues, and on-ramp performance.

NCHRP report 505 classified acceleration into low-speed and high-speed acceleration. The report also provided the acceleration rates for heavy vehicles with varying weight to power ratios. The acceleration characteristics of buses were also reviewed as they will use the TOLs. Gattis et al. (Gattis, 1998) studied the acceleration behavior of buses. The study collected speed, time, and distance data for full size buses accelerating from a stopped position.

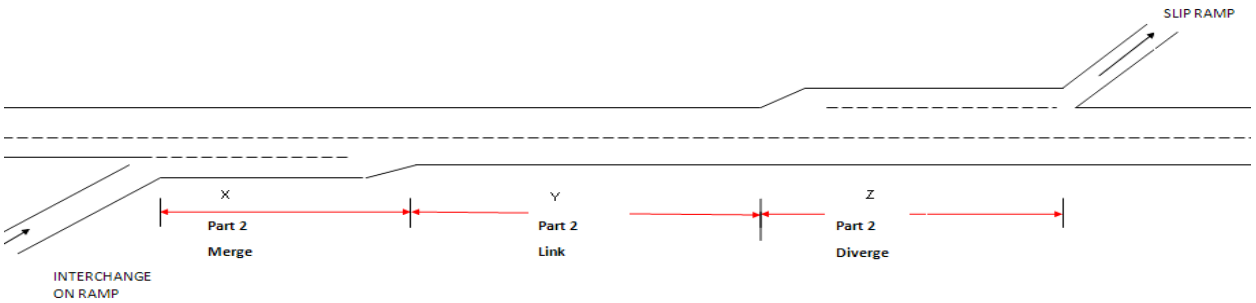
## METHODOLOGY

From this study the three key segments merge, link and diverge, also presented in Figure 2, were analyzed to determine the slip ramp spacing. These segments were used by heavy vehicles to merge with passenger vehicles on the GPLs and then diverge to exit the GPLs. The TOL model was constructed in VISSIM as per the

AASHTO Design Guide specifications (AASHTO, 2004) for lengths of on- and off-ramps. These segments were evaluated for adequate length using multiple simulation runs using VISSIM. Based on the analysis of simulation results, design length recommendations for these segments were developed.




a) Schematics for the Different Segments of Part 1 of General Purpose Lanes




b) Schematics for the Different Segments of Part 2 of General Purpose Lanes

**FIGURE 2. Truck Only Lanes**

## The VISSIM Simulation Model

For this study, VISSIM 5.1.0 was used for determining the optimum slip ramp spacing. VISSIM provides significant control over individual driver behaviors based on different vehicles categories for specific sections in terms of both car-following and lane changing phenomenon compared to other simulation models.

The default behavioral parameters in VISSIM were calibrated for conditions appropriate for TOL and GPL and to simulate realistic traffic conditions. Important car-following parameters that have a significant impact on driving behavior were identified from related studies. In particular, Gomes et al. (Gomes, 2004) study was used to select driver behavior parameters such as CC0 … CC4/CC5. NCHRP Report 505 was used to study the acceleration characteristics for heavy vehicles to be used in VISSIM (Douglas, 2003). Speed limits were established for vehicles to follow on the freeway, slip and the on- and off-ramps. The following sections provide details on the set up and the use of VISSIM to simulate the TOLs.

## The TOL Model

The TOLs were modeled in VISSIM using the highway geometry prepared for MoDOT based on the AASHTO Design Guide (AASHTO, 2004). The lengths of the merge, link, and diverge sections were specified as 800 ft, 900 ft, and 650 ft, respectively as shown in Figure 2. The lane widths specified were 12 feet, and the median spacing was 30 feet. The taper length for single-lane entrance and exit ramps, per Exhibit 10-69, was set at 300 feet and per Exhibit 10-72 at 250 feet. The slip

ramps were 800 ft long. Figure 2 presents the geometrics of TOLs and GPLs as coded in VISSIM.

**Traffic Composition and Traffic Inputs**

To simulate forecasted traffic conditions, traffic composition from a Missouri rural highway section was used as input. The data were provided by the MoDOT. Table 1 shows the traffic composition determined from the Annual Average Daily Traffic data for Missouri.

From the Missouri traffic data, heavy vehicle peak flow rates were determined to be around 1,220 trucks per hour (tph) for TOLs and the peak flow rate of passenger vehicles was 2,500 vehicles per hour (vph) for GPLs. During the peak hours, a maximum of 320 tph were expected to exit the TOLs using the slip ramps and the GPLs. Truck volumes higher than 320 tph would require grade-separated interchanges.

**TABLE 1. Vehicle Composition**

| Description | Percentage |
|---|---|
| **Truck Only Lanes** | |
| Buses | 1.50 |
| 2 axle 6 tire single units | 2.70 |
| 3 axle single units | 2.20 |
| 4 > axle single units | 0.03 |
| 4 < axle single units | 6.70 |
| 5 axle single trailers | 79.00 |
| 6 > axle single trailers | 0.48 |
| 5 < axle multi trailers | 4.60 |
| 6 axle multi trailers | 2.60 |
| 7 > axle multi trailers | 0.19 |
| **General Purpose Lanes** | |
| Motor Cycles | 0.64 |
| Passenger cars | 75.00 |
| 2 Axle 4 tire single units | 24.36 |

**Drivers' Behavioral Parameters**

Acceleration/Deceleration Behavior

The acceleration characteristics were plotted for heavy vehicles, cars, bikes and, buses.

Maximum acceleration rates for heavy vehicles on level terrain were obtained from University of Michigan Transportation institute UMTRI (Woodrooffe, 2009). Eq. 1 and 2 below plot the maximum acceleration versus speed for heavy vehicles with a gross vehicle weight (GVW) of 70,000 lbs and engine power (P) of 240 hp (250 lb/hp). The mass power acceleration relationship can be approximated to provide an estimate of non-dimensional acceleration (g), which can be expressed as:

$$g = \frac{2 * P}{GVW} \quad (1)$$

For differential speed changes between 30 and 70 mph, the following equation approximates the distance required to change speed.

$$D = \frac{V_f^2 - V_i^2}{2A} \cong \left(\frac{GVW}{2 * P}\right) * \left(\frac{V_f^2 - V_i^2}{2g}\right) \quad (2)$$

where:

D = distance traveled,
A = average acceleration which is related to the ratio of power to weight,
$V_f$ = final velocity,
$V_i$ = initial velocity,
$GVW$ = gross vehicle weight,
$P$ = HP, and
$G$ = acceleration due to gravity.

Equation 1 was used to plot the acceleration versus speed for heavy vehicles in VISSIM. Gattis et al. (Gattis, 1988) provided the acceleration rates used to plot acceleration versus speed for buses in VISSIM. The acceleration values for passenger cars were obtained from the Gomes study (Gomes, 2004). The acceleration values for motorcycles were obtained from FHWA study (Wicks, 1980).

Car-Following and Lane Change Parameters

VISSIM uses the Wiedmann-74 and 99 car-following models to simulate freeway behavior. The car-following model that represents freeway conditions, Wiedemann-99 car following model, has 11 user defined driving behavior parameters: CC0…CC10 (VISSIM, 2008) that can be adjusted to simulate realistic driving behavior. Gomes et al. (Gomes, 2004) and Lownes and Machemehl (Lownes, 2006) proposed driver-behavior parameters for the VISSIM Wiedmann-99 model that can be used for various sections of TOLs and GPLs, defining merging, diverging, link behavior and TOL freeway behavior for various segments. The parameters most significant for realistic simulation of TOL are CC0, CC1, and the CC4/CC5 values. These variables are defined briefly below.

Merge

Figures 2a and 2b indicate the merge sections for both Part 1 and Part 2. Heavy vehicles exit the TOL using the slip ramp and merge with vehicles on the GPLs. Similarly, heavy vehicles use the on-ramp to merge with vehicles on the GPLs. These behaviors affect passenger vehicle speeds on the main highway lines.

To model the realistic merging behavior of drivers in VISSIM, the CC parameters were specified. The CC1 value used was 1.4 seconds to simulate realistic headway times. The CC0 value for merge behavior was 5.58 ft, higher than the default value of 4.92 ft, to increase the safety distance between the vehicles on the merge section. The waiting time for diffusion was set at 60 seconds, which provided ample time for vehicles to change lanes and join the mainstream traffic during peak traffic flow conditions. The Safety Distance Reduction Factor (SDRF) value was set at 0.2 to generate realistic merge behavior on the freeways. This value was also used in a study by the Virginia Department of Transportation (Srividya, 2008).

Link

This section is defined as 'B' and 'Y' in Figure 2 between the merge and the diverge sections. The CC0 parameter selected was 5.58 ft, and values of CC4 and CC5 which specify vehicle following thresholds were modified to -2.0 and 2.0, respectively from the default values of -3.5 and 3.5. These values are similar to calibrated values for a freeway used by Gomes et al. (Gomes, 2004). The SDRF value on this section was 0.6, appropriate for a freeway section as no aggressive lane changing was required. The waiting time for diffusion on the freeway was set at 1 second

Diverge

Driver behavioral parameters similar to those used for merging were used for this diverging section as well. A SDRF value of 0.4 was used for off-ramps, in-between 0.2 (aggressive) and 0.6 (timid), that calibrated the SDRF factor for trucks and passenger vehicles using off-ramps. The waiting time for diffusion was kept at 60 seconds so that vehicles would have sufficient time to change lanes and join the main stream traffic during the peak flow conditions.

**Routing Decision Parameters**

Routes were defined for the TOLs and GPLs so that vehicles exit and enter the TOLs using the specified segments. Additionally, the routing decision parameters determined the percentage of truck traffic exiting from the slip ramp and the off-ramp. The percentages of trucks exiting the TOL via slip ramps were provided in such a way that passenger vehicles on GPL and heavy vehicles from the slip ramp reached peak flows at the same time, thereby critically testing the design lengths of the key segments for congestion

**Speed Decisions**

The speed limit for the TOLs, GPLs and slip ramps was set at 70 mph per Missouri interstate highway speed limits. To maintain homogeneity in truck speeds between the TOLs and GPLs when trucks exit or enter using the slip ramps, a 70 mph speed limit was set for the slip ramps. This limit ensured consistent speeds among heavy vehicles merging with the GPL traffic and with the TOL traffic. The speed limit for vehicles on the off- and on-ramps was set at 40 mph, per the AASHTO Design Guide (AASHTO, 2004).

**Design of Experiments**

Central Composite Design (CCD)

As part of designing the experiment, central composite design was used to generate design cases with different combinations of merge, link, and diverge lengths.

Figure 3 illustrates a RCC design. The stars in Figure 3 are located at a distance $\alpha$ from the center based on the experimental error and the number of factors in the design.

The star points establish extremes for the low and high lengths for all factors.

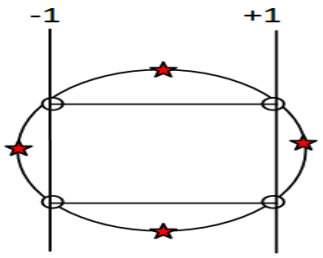



Figure 3. Rotatable Central Composite Design (Not To Scale)

JMP software [John, 2008] was used to calculate the design cases to be simulated. Pilot runs were conducted to determine the maximum and minimum lengths of each of the sections in which the speeds were acceptable, and the results were used as an input to JMP in generating the different cases. Table 2 provides the different segment lengths evaluated. The slip ramp spacing was determined by the sum of the lengths of the merge, link, and diverge segments.

Fifteen scenarios were each simulated 15 times using VISSIM with different random seeds to add variability to the results from the micro-simulation model. To ensure accurate results, the resolution of the simulation was set to two simulation seconds for each time step. Evaluation was performed based on the average results on the merge, diverge, and link sections. The traffic flow variables were averaged every five minutes for analysis.

**TABLE 2. Simulated Lengths for Part 1 and Part II of the GPLs Level Grade**

| Cases | Merge | Link | Diverge | Total | Results |
|---|---|---|---|---|---|
| *Case 1* | 800 | 900 | 650 | 2350 | Failed |
| *Case 2* | 800 | 900 | 1000 | 2700 | Failed |
| *Case 3* | 800 | 1500 | 650 | 2950 | Failed |
| *Case 4* | 800 | 1500 | 1000 | 3300 | Pass |
| *Case 5* | 900 | 1200 | 1100 | 3200 | Failed |
| *Case 6* | 900 | 1700 | 825 | 3425 | Pass |
| *Case 7* | 900 | 1200 | 550 | 2650 | Failed |
| *Case 8* | 1050 | 1200 | 825 | 3075 | Pass |
| *Case 9* | 900 | 1200 | 825 | 2925 | Pass |
| *Case 10* | 900 | 700 | 825 | 2425 | Failed |
| *Case 11* | 1000 | 1500 | 1000 | 3500 | Pass |
| *Case 12* | 1000 | 900 | 650 | 2550 | Failed |
| *Case 13* | 1000 | 1500 | 650 | 3150 | Failed |
| ***Case 14*** | **1000** | **900** | **1000** | **2900** | **Pass** |
| *Case 15* | 750 | 1200 | 825 | 2775 | Failed |

Bold values indicate recommended slip ramp spacing distance

**Level-of-Service (LOS)**

For this study, the Highway Capacity Software (HCS) (Highway Capacity Manual, 2000) was used to calculate the LOS and vehicle speeds obtained for the proposed

optimized slip ramp spacing distance from TOL slip ramp to off-ramp. HCS was used to compare the speeds of vehicles on the merge and diverge sections with the speeds of vehicles obtained in VISSIM for the same sections during peak flows for level, up- and down-grades. For the merge and diverge sections, merging and diverging procedures of the Highway Capacity Manual were used.

## ANALYSIS OF RESULTS

The slip ramp spacing was evaluated by studying the speed-flow behavior of vehicles on the key segments of the GPLs. Figure 2 shows the GPL segments that were studied for speed-flow plots before and after the interchange. These segments were chosen because heavy vehicles and passenger vehicles interact on these segments. The speeds of heavy and passenger vehicles were analyzed separately for more careful consideration for level grades.

For the level sections simulated, when vehicle speeds remained above 50 mph and there were no signs of congestion during peak flow on key segments, those cases were deemed acceptable. Design cases in which vehicles slowed down to below 50 mph, those cases were unacceptable and are indicated as failed in Table 2. It is to be noted that the speeds of heavy vehicles should be above 50 mph in the merge, link, and diverge sections for a case to be acceptable. Additional lengths per CCD were selectively provided to merge, link, and diverge segments and speed flow relationships were analyzed until an acceptable case for peak volume was found.

### Level Grade

Case 1 with a merge length of 800 feet, a link length of 900 feet, and a diverge length of 650 feet as recommended by the AASHTO Design Guide, consistently performed poorly with very low speeds for all vehicles during peak flows. Figure 4 presents the speed-flow relationship for this case, which indicates that heavy vehicles slowed down to 30 mph before exiting using the off-ramp. The speeds of passenger vehicles were also affected as they were reduced to below 50 mph. This case failed because the vehicle speeds were unacceptable for this type of facility.

For Case 4, the design lengths were found to be optimum to maintain vehicle speeds while entering and exiting. The plots, demonstrate that vehicle speeds for this particular case were consistently between 50 and 70 mph. This case was acceptable for heavy vehicles using the slip ramp with a peak flow rate of 320 tph. Greater design lengths and different combinations from the CCD were simulated and it was found that speeds of heavy vehicles and passenger vehicles in cases 6, 8, 9, 11 and, 14 were also acceptable for a truck volume of 320 tph. From Table 2, it can be concluded that the minimum length for diverge section must be 825 feet for the speeds of heavy vehicles to be above 50 mph.

The speed of passenger vehicles was consistent in the merge and link sections of Parts 1 and II, varying from 60 to 70 mph for all cases except for Case 1. In the diverge section for Case 1, the speeds of both passenger vehicles and heavy vehicles during the peak flows were very low compared to other cases with longer diverge lengths. Case 1, therefore, is not appropriate because vehicles slowed down considerably on the GPLs. Case 4, with a diverge length of 1000 feet, was acceptable, with passenger vehicles speeds above 60 mph. The simulation was run with both heavy vehicles and passenger vehicles reaching peak flows at the same time

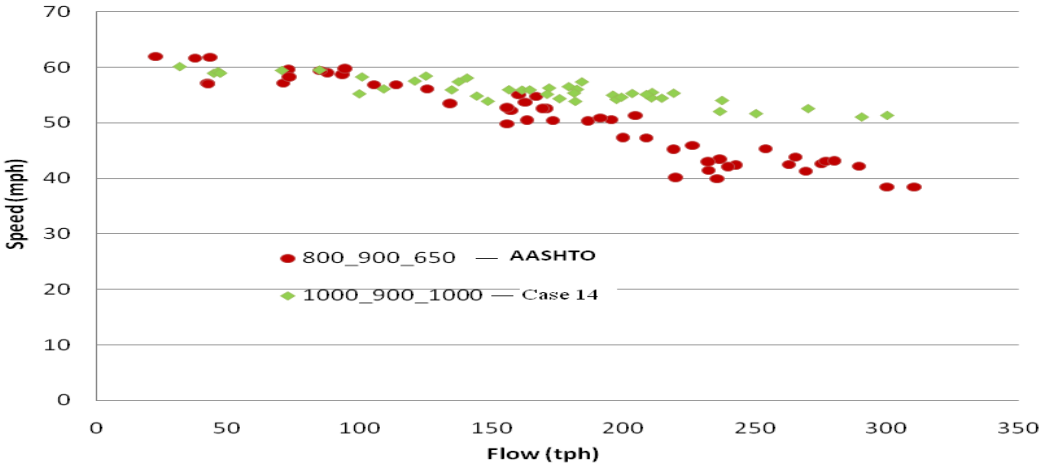



**FIGURE 4. Before interchange (level grade), heavy vehicles on diverge section of GPL.**

This study also analyzed the speed-flow characteristics for vehicles traveling on merge, link, and diverge sections of Part 2, as shown in Figure 2b. The trends observed were similar to those in Part I for the speed-flow curves, with the additional lengths permitting higher vehicular speeds. The cases that successfully passed for Part 1 also passed for Part 2.

Different combinations of merge, link, and diverge length were carried out for both Part 1 and Part 2. This analysis showed that the governing lengths of merge, link, and diverge sections were the distance needed for heavy vehicles to complete the lane change from the slip ramp to the off-ramp for Part I and from the on-ramp to the slip ramp for Part II. The recommended merge length was 1000 feet, the link length was 900 feet and diverge length was 1,000 feet for peak volume of heavy vehicles. Greater sections lengths were simulated, but significant improvements in vehicle speeds were not observed for the peak flow.

**Level-of-Service**

The LOS for merge and diverge sections for the optimized slip ramp spacing was calculated. The peak flow on the slip ramp was 320 tph and the peak flow of passenger vehicles on the GPL was 1000 vph and 1500 vph for the left and right lanes, respectively. The LOS was calculated to be 'B' for merge and diverge sections on GPL. The space mean speeds provided by HCS in the ramp influence area of diverge section was 57.5 mph compared to 55.63 mph for speeds of vehicles in VISSIM.

## CONCLUSIONS AND RECOMMENDATIONS

This paper studied the slip ramp spacing for safe and efficient operation of the GPLs and TOLs. Research indicated that the distance needed for trucks to weave in traffic was greater than to accelerate or decelerate; therefore, the weaving distance controls the length of the acceleration/deceleration lanes. Higher traffic flow rates can be achieved with no significant drop in the speed if additional merge, link, and diverge lengths are provided. Part 1 of the TOL merge section showed no signs of congestion for any of the design lengths tested. The average length of a merge section was found to be 1000 feet, the diverge lane length to be 1000 feet, and the weaving section requires 900 feet for heavy vehicle traffic of 320 tph. The total distance from slip

ramp to off-ramp is recommended to be between 2,900 and 3,500 feet. The results for Part 2, evaluating slip ramps after the interchange, followed a similar trend.. Operations on GPL degrade for traffics flow between 280 and 320 trucks per hour and between 2,500 and 3,000 vehicles per hour. LOS was found to be 'B' for the recommended merge, link and diverge sections for level grades on GPL. The speeds for vehicles on merge and diverge sections on GPL was found to be similar in both HCS and VISSIM softwares. To maintain homogeneity in speeds of heavy vehicles and to reduce the congestion on the merge section of a GPL, the speed limits on slip ramps should be the same as those on GPLs.

**ACKNOWLEDGEMENTS**

Authors acknowledge the research grant from the MoDOT

**REFERENCES**

AASHTO (2004), "A Policy on Geometric Design of Highways and Streets", Washington, DC.

Bham, G. H. and Benekohal, R. F. (2004), "A High Fidelity Traffic Simulation Model based on Cellular Automata and Car-Following Concepts", *Transportation Research Part C: Emerging Technologies*, 12 (1), 1-32.

Burke, N.A., H. Thomas, R. Crum, J. Plazak, J. David, and S. Omar (2007), "Dedicated Truck Lanes As A Solution To Capacity And Safety Issues On Interstate Highway Corridors", In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2008, Washington, D.C.

Douglas, W.H., and K. R. Richard (2003), "Review of Truck Characteristics as Factors in Roadway Design", *NCHRP Report 505,* Transportation Research Board of the National Academies, Washington, D.C.

Gattis, J.L., M. Bryant, and K.D. Lynette (2008), "Acceleration Lane Design for Higher Truck Volumes", *Mack-Blackwell Transportation Center*, Research and Special Programs Administration.

Gattis, J.L., S. Nelson, and J. Tubbs (1988), "School Bus Acceleration Characteristics", *Report No. MBTC FR 1054-2*, Arkansas State Highway and Transportation Department.

Gomes, G., A. May, and R. Horowitz (2004), "Congested Freeway Microsimulation Model Using VISSIM", In *Transportation Research Record*: *Journal of the Transportation Research Board*, *No. 1876*, Transportation Research Board of the National Academies, Washington, D.C.

Highway Capacity Manual 2000 (2000), *TRB*, National Research Council, Washington, D.C.

Lownes, N. E. and R. B. Machemehl (2006), "Sensitivity of Simulated Capacity to Modification of VISSIM Driver Behavior Parameters", *Transportation*

*Research Record: Journal of the Transportation Research Board*, No. 1988, Transportation Research Board of National Academics, Washington, D.C.

Montgomery (2001), *D.C. Design and Analysis of Experiments*, John Wiley & Sons, John Sall, Lee Creighton, and Ann Lehman (2008), JMP Start Statistics: A Guide to

Statistics and Data Analysis Using JMP, Fourth Edition, , SAS Press.

Sarvi, M (2009), "Freeway Operation Characteristics of Heavy Commercial Vehicles Traffic", *88th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Srividya, S., and B. Park (2008), "Development of VISSIM Base Model for Northern Virginia (NOVA) Freeway System", *Report No. UVACTS-13-0-124*,

John Woodrooffe (2009), UMTRI Truck Acceleration Characteristics.

VISSIM User Manual (2008), *PTV Planning Transport Verkehr AG*, Karlsruhe, Germany.

Wicks, D.A., and E.B. Lieberman (1980), "Development and testing of INTRAS: a microscopic freeway simulation model", *Vol. 3, Publication FHWA-RD-80-106*, FHWA.

# How Effective are "Flashing Yellow Arrow" Signals in Enhancing Safety?

**Srinivas S. Pulugurtha, Ph.D., P.E., M. ASCE[1], Mahesh Agurla, B.S.[2] and Khamar Salma Chittoor Khader, B.S.[3]**

## ABSTRACT

Crashes (in particular, those involving left-turn vehicles) at signalized intersections result in substantial economic losses to the society in terms of death, injury, and property damage. A wide variety of traffic signals are used by the traffic engineers, in the United States, to minimize left-turn crashes and provide maximum safety at signalized intersections. One latest innovation and addition to these traffic signals is the "flashing yellow arrow (FYA)" signal. This paper focuses on safety evaluation of FYA signals "before" and "after" their installation at six selected signalized intersections in the city of Charlotte, North Carolina. Empirical Bayes (EB) technique was used for evaluation. The EB method compares the number of crashes "after" the installation to the estimated number of crashes that would have occurred during the "after" period if the FYA signal had not been installed. Results obtained indicate that the number of crashes would have generally increased if the FYA signals had not been installed at the selected signalized intersections.

## INTRODUCTION

Left-turn is one of the most critical movements a motorist makes at an intersection. In the United States, nearly 30 percent of intersection fatalities occur at signalized intersections (Rice, 2007). Accommodating and improving safety of left-turning vehicles at signalized intersections, therefore, has been a high concern for traffic

[1] Associate Professor of Civil & Environmental Engineering, Assistant Director of Center for Transportation Policy Studies, The University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001, USA; Phone: (704) 687-6660, Fax: (704) 687-6953, E-mail: sspulugurtha@uncc.edu

[2] Graduate Student of Civil & Environmental Engineering, The University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001, USA; Phone: (704) 687-2769, Fax: (704) 687-6953, E-mail: magurla@uncc.edu

[3] Graduate Student of Civil & Environmental Engineering, The University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001, USA; Phone: (704) 687-2769, Fax: (704) 687-6953, E-mail: kchittoo@uncc.edu

engineers as they seek a balance between intersection capacity and safety through traffic signal phasing/timing techniques.

Exclusive left-turn lanes and protected left-turn phases are commonly used to minimize the effect of left-turning vehicles on operational performance and safety at intersections with high left-turn traffic volume. Based on traffic volume along both the streets and other needs, permissive left-turn phase or protected/permissive left-turn (PPLT) phase have also been in use to increase left-turn capacity and reduce delay at intersections. The PPLT phase can either be a lead (precede) or lag (follow) phase.

A key concern with PPLT control is the "yellow trap", which occurs during the change from permitted left-turn phase for both directions to a lagging protected left-turn phase for one direction (Kittleson Associates, Unknown Date). To avoid the yellow trap, most agencies use "Dallas Display" with permissive left-turn phases. This operation improves safety but may not be effective at all signalized intersections (Rice, 2007).

A latest edition to the list of PPLT or permissive left-turn signals is the "flashing yellow arrow (FYA)" signal. It is used to better accommodate and improve safety of left-turning traffic (due to high level driver comprehension) (Knodler et al. (2007). It also helps reduce vehicular delay at intersections. When looked at a broader perspective, FYA signals also reduce vehicle emissions and improve air quality at intersections.

Several State and local agencies in Arizona, Arkansas, California, Colorado, Florida, Georgia, Idaho, Iowa, Kansas, Kentucky, Louisiana, Michigan, Missouri, Nevada, New Hampshire, New Mexico, North Carolina, Oregon, Texas, Utah, Virginia, Washington and Wyoming have interim approval from the Federal Highway Administration (FHWA) to install and evaluate FYA signals (MUTCD, 2008).

This paper focuses on safety evaluation of FYA signals installed at six signalized intersections in the City of Charlotte, North Carolina.

## LITERATURE REVIEW

A survey conducted as a part of National Highway Cooperative Highway Research Program (NCHRP) 493 study (Brehmer et al. 2003) estimated that 29 percent of the signalized intersections have at least one approach with PPLT signal operation in the United States. In an attempt to better communicate to the left-turn driver that they must yield to opposing traffic before turning, traffic engineers in the United States implemented a variety of unique permissive left-turn signal indications such as flashing circular red (FCR), flashing red arrow (FRA), flashing circular yellow (FCY), and FYA signal. These signals vary in arrangement, indications, placement, and phasing sequence which could have a bearing on safety and operational effectiveness. The focus of the NCHRP 493 study was to evaluate each of the above signal indications and determine which was the most effective and could be uniformly applied throughout the United States. To accomplish this, a comprehensive array of research studies and experiments were completed including driver behavior laboratory analyses, operational and behavioral field evaluations, and full-scale driving simulator analysis. However, not much was done in terms of safety analysis.

Noyce et al. (2000) evaluated the operational characteristics associated with various protected/permitted left-turn signal displays such as FCR, FCY, FRA, and FYA. Their study found no differences in saturation flow rate and start-up lost time due to the type of PPLT signal display.

Noyce and Smith (2003) evaluated drivers' comprehension and response to various combinations of five-section PPLT signal displays (horizontal, vertical, and cluster) and permissive left-turn indications (circular green, FCR, FCY, FRA, and FYA). Results showed that five-section PPLT signal display arrangement has very little effect on driver comprehension of the permissive left-turn maneuver. The circular green, FCY and FYA were the best understood by drivers'.

Brehmer et al. (2003) conducted a study on installation and effectiveness of FYA signals. The study in Seattle, Washington revealed that FYA signal helps reduce left-turn crash occurrences at night more than a circular green signal. The results indicated that public reaction to the FYA signals was positive. It was, therefore, regarded as the most intuitively understood PPLT signal and recommended for approval by Manual on Uniform Traffic Control Devices (MUTCD).

To evaluate the safety effectiveness of the FYA permissive left-turn indication field installations, Knodler et al. (2005) used linear trend to estimate the crash frequency rate and Empirical Bayes (EB) method to project crashes at study intersections. The study concluded that there was no change in safety at intersections that operated only protected left-turn phasing prior to the implementation of the FYA permissive signal indication with PPLT phasing. No conclusions could be made at intersections that operated with permissive only left-turn phasing prior to the implementation of the FYA permissive signal indication due to a minimal number of implementation sites and data.

Mullinax (2006) was strongly in favor of FYA signals for a variety of reasons such as enhanced operational control, versatility, consistency and intuitive understanding to motorists.

Knodler et al. (2007) conducted a sequential experiment on driver's evaluation of the solid yellow arrow (SYA) to check if the FYA results in a change in drivers' perceived understanding of FYA indication in Madison, WI. The study concluded that there was no evidence to suggest that the FYA permissive indication might negatively affect drivers' understanding of SYA indication.

Noyce et al. (2007) evaluated the ability of the FYA signals to improve safety and reduce the number of crashes. Factors such as signal phasing, vehicle flow rates, posted speed limits and intersection geometry were used to interpret findings from the crash analysis. It was concluded that safety at intersections improved when FYA signals were added to existing PPLT signal phasing operations.

## STUDY INTERSECTIONS

Six signalized intersections with FYA signals in the city of Charlotte, North Carolina were considered for safety analysis and evaluation. The study intersections, the number of approaches with FYA signals and the posted speed limits, annual average daily traffic (AADT), and the average number of crashes "before" and "after" the installation of FYA signals are shown in Table 1. The crash analysis period varied

depending upon the date of installation and availability of data for each intersection. Three-year "before" period and one-year "after" period was considered for analysis.

## METHODOLOGY

EB method was used to evaluate the effectiveness of FYA signals. The advantage of EB method is that it minimizes regression-to-mean bias and gives valid results for even small samples. In addition, the "before" and "after" analysis time period need not be the same (Hauer, 2002).

The EB method compares the actual number of crashes "after" the installation of FYA signals and the estimated number of crashes during the "after" period had FYA signal not been installed. The process of evaluation involves the following four steps.

1. Safety performance function
2. Over-dispersion parameter
3. Relative weight
4. Estimate expected number of crashes

The methodology adopted in this research was discussed in detail by past researchers (example, Powers and Carson, 2004) and outlined in the following subsections.

### Safety Performance Function

The safety performance function is a unique mathematical model that estimates the number of crashes (say, at an intersection) connecting the expected crash rates to measurable features such as AADT, speed limit, number of lanes and/or other explanatory factors. The safety performance function is determined from the data collected for intersections without FYA signals to help estimate the number of crashes per year at an intersection had FYA signal not been installed.

In this paper, using crash and intersection data available for the "before" period (multiple years) for signalized intersections in the study area, a safety performance function was developed using a generalized linear model based on negative binomial distribution (as data exhibited over-dispersion). The independent variables initially considered for developing the safety performance function were posted speed limit, presence of median, skewness of intersection, the number of lanes along the FYA approach, lane width and AADT. All intersections considered for analysis were 4-legged intersections.

The correlation between the independent variables and their relation to the dependent variable (number of crashes) was examined before the development of safety performance function. The variables that had a strong correlation with the dependant variable (the number of crashes) based on Pearson chi-square estimate (significance level of 0.05 or 95 percent confidence level) were AADT and skewness.

Hence, only these independent variables were used to develop the safety performance function.

The generalized linear regression model representing the safety performance function developed using the before data is as follows.

$$SPF_i = \exp [1.844 + (0.00003757*AADT_i) + (0.887*Skewness_i)] \quad \ldots \text{Equation (1)}$$

where,

$SPF_i$ is safety performance function factor for intersection "i",

$AADT_i$ is traffic volume for intersection "i", and,

$Skewness_i$ is if intersection "i" is skewed or not skewed.

The numeric values preceding AADT and skewness are the computed coefficients for the independent variables. Note that the computed values from the safety performance function vary for each intersection based on AADT and skewness.

**Goodness of fit**

Goodness of fit for the generalized linear model is assessed using either Quasi likelihood criterion (QIC), Corrected Quasi Likelihood under Independence Model Criterion (QICC) statistic or Wald chi- square. In this paper, QIC, QICC and significance level were used to assess the strength of variable(s) and the goodness of fit.

**QIC and QICC value**

QIC statistic is used to evaluate the correlation, whereas QICC statistic helps the user to decide on the best subset of model predictors for a particular correlation structure. QICC statistic is designed to evaluate which predictors best explain the response within the model framework. If the difference between QIC and QICC values is low, the selected model has best correlation structure and predictors obtained best fit the model.

The QIC and QICC value obtained for the developed safety performance function were 4.996 and 10.064, respectively.

**Significance level**

Significance level is defined as the maximum allowable probability of committing a statistical error. In this paper, a significance level less than 0.05 are considered as allowable to explain the relationship between variables. The computed significance level for the model was 0.001. This indicates that one can be more than 99 percent confident about the role of AADT and skewness on crashes at the selected intersections.

**Over-dispersion Parameter**

Crash frequencies are often assumed to follow a Poisson distribution or Negative Binomial distribution. A Poisson distribution is appropriate if the mean and variance

of the dependent variable are equal. However, the computed variance was observed greater than the computed mean for the data considered in this paper. Hence, to account for over-dispersion of data, the distribution of crash frequencies was represented using a negative binomial distribution. The over-dispersion parameter, φ, was calculated using features available in SPSS© statistical analysis software (SPSS Inc., 2005).

### Relative Weight

A relative weight, $\alpha_i$, was applied to each intersection "i" to adjust for varying degrees of over-dispersion. The specific relative weight for each intersection ($\alpha_i$) was calculated using the following equation.

$$\alpha_i = 1 / (1 + (SPF_i/\varphi) \quad \text{… Equation (2)}$$

where,

$\alpha_i$ denotes the relative weight for intersection 'i'

### Estimate Expected Number of Crashes

The expected number of crashes had FYA not been installed, $\mathbf{C_{E,i}}$, for each intersection "i" is then estimated using the following equation.

$$C_{E,i} = (\alpha_i \times SPF_i) + ((1 - \alpha_i) \times C_{B,i}) \quad \text{… Equation (3)}$$

where,

$C_{E,i}$ is the expected number of crashes for intersection "i", and,

$C_{B,i}$ is the actual number of crashes for intersection, "i" during the before period.

## ANALYSIS & RESULTS

Six signalized intersections with FYA signals installed in 2007 and 2008 in the city of Charlotte, North Carolina were considered for analysis and evaluation. Crash data, network characteristics and traffic data were obtained from the City of Charlotte Department of Transportation. They are summarized in Table 1.

The minimum number of crashes (average) per year during the "before" period was 9 whereas the maximum number of crashes (average) per year during the "before" period was 33. On the other hand, the minimum number of crashes per year during the "after" period was 8 whereas the maximum number of crashes per year during the "after" period was 37.

The AADT (average for available years during the "before" period) varied from 22,000 vehicles per day to 45,000 vehicles per day at the selected signalized intersections. The speed limits generally varied from 30 mph to 45 mph along corridors with these signalized intersections.

Table 2 summarizes estimates using safety performance function, over-dispersion parameter, relative weight and the expected number of crashes per year. Projected

AADT during the "after" period was used to estimate the safety performance function and the expected number of crashes. A 3 percent background traffic growth rate recommended by regional agencies for projecting future growth in traffic was used.

Table 3 shows the estimated expected number of crashes, the actual number of crashes and the ratio of actual and estimated expected number of crashes per year had FYA not been installed at the intersections. A ratio less than 1.0 an improvement in safety after the installation of FYA signals.

It can be observed that the actual number of crashes is lower than the estimated expected number of crashes for five out of the selected six signalized intersections. The ratio was found to be less than 1.0 for these five intersections. This indicates an improvement in safety at five out of the selected six intersections "after" the installation of FYA.

## CONCLUSIONS

This paper focuses on safety evaluation of "flashing yellow arrow (FYA)" signals installed at six signalized intersections in the city of Charlotte, North Carolina. Empirical Bayes (EB) method was used to compare the actual number of crashes and the estimated number of crashes had FYA not been installed at the intersections. Results obtained showed improvements in safety at five out of the six signalized intersections used for evaluation in this paper.

While results are generally encouraging, it cannot be generally concluded that FYA have been very effective in reducing the crashes at all installed signal locations. Results considering larger sample size and considering only left-turn crashes needs an investigation. In addition, analysis is warranted based on intersection characteristics and signal phasing/timing patterns to identify where FYA signals are most effective.

## REFERENCES

1. Brehmer, Chris L., Kent C. Kacir, David A. Noyce and Michael P. Manser (2003) "Evaluation of Traffic Signal Displays for Protected/Permissive Left-Turn Control." NCHRP Report 493, National Cooperative Highway Research Program, Transportation Research Board.
2. Hauer, Ezra (2002). "Observational Before – After Studies in Road Safety." Pergamon/Elsevier Science, Inc.
3. Kittleson Associates (Unknown Date). "Dallas PPLT Displays (Horizontal Head) Lead left turn." Available Online: <http://projects.kittelson.com/pplt/displays/dallas_horiz_lead.htm> (Accessed October 29, 2010)
4. Knodler Jr., Michael A., David A. Noyce, Kent C. Kacir and Christopher L. Brehmer (2005) "Evaluation of Flashing Yellow Arrow in Traffic Signal Displays with Simultaneous Permissive Indications." *Transportation Research Record No. 1918*, p (46-55).
5. Knodler Jr., Michael A., David A. Noyce, Kent C. Kacir and Christopher L. Brehmer (2007) "An Evaluation of Driver Comprehension of Solid Yellow Indications Resulting from Evaluation of Flashing Yellow Arrow."

Compendium of Papers, *Transportation Research Board 86th Annual Meeting*, CD-ROM, Washington, DC.
6. Mullinax, Richard E. (2006) "Four-Section Protected-Permissive Flashing Yellow Arrow Traffic Signal Head." *Traffic Engineering Conference for Operations & Safety*, Wilmington, NC.
7. MUTCD (2008) Manual on Uniform Traffic Control Devices (MUTCD) - List of Authorized Requests for Interim Approval. Available Online: <http://mutcd.fhwa.dot.gov/resources/interim_approval/ialistreq.html> (Accessed November 19, 2009).
8. Noyce, D. A., Casey R. Bergh, and Jeremy R. Chapman. (2007) "Evaluation of the Flashing Yellow Arrow Permissive-Only Left-turn Indication Field Implementation." NCHRP Web Document, Available Online: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w123.pdf (Accessed June 30, 2010).
9. Noyce, David A. and C. R. Smith (2003). "Driving simulators for Evaluation of Novel Traffic Control Devices: Protected-permitted Left-Turn Signal Display Analysis." *Transportation Research Record 1844*, p (25-34).
10. Noyce, David A., Daniel B. Fambro and Kent C. Kacir (2000). "Traffic Characteristics of Protected/Permitted Left-Turn Signal Displays." *Transportation Research Record No. 1708*, p (28-39).
11. Powers, Meghan and Jodi Caron (2004). A Primer for Using the Empirical Bayes Method Tutorial. Available Online: http://www.mdt.mt.gov/research/docs/research_proj/crash_analysis/tutorial.pdf (Accessed June 12, 2010).
12. Rice, Ed (2007) "Taking Action to Reduce Intersection Fatalities." *Safety* Compass 1 (2), p (1–3).
13. SPSS Inc. (2005). SPSS® 14.0 *Brief Guide*, Copyright © 2005 by SPSS Inc., 233 South Wacker Drive, 11th Floor, Chicago, IL.

## ACKNOWLEDGMENTS

The authors acknowledge the staff of the City of Charlotte Department of Transportation (CDOT) for their help with data for this research.

## DISCLAIMER

The contents of this paper reflect the views of the author(s) and not necessarily the views of the University or CDOT. The author(s) are responsible for the facts and the accuracy of the data presented herein.

## TABLE 1 Study Intersections – Summary

| Study Intersection (i) | Before - Average # Crashes per Year ($C_{B,i}$) | After - Actual # Crashes per Year ($C_{A,i}$) | Before - AADT (veh/day) | Speed Limit (mph) | # Lanes (FYA Approach) |
|---|---|---|---|---|---|
| East 5th St / 7th St | 8.7 | 8 | 22,653 | 35 | 2 |
| I-77 SB Ramp / Sunset Rd | 28.3 | 22 | 44,658 | 45 | 2 |
| I-85 SB Ramp / Mallard Creek Ch Rd | 33.3 | 37 | 35,300 | 45 | 1 |
| Erwin Rd / Tryon St | 15.3 | 9 | 36,800 | 35 | 1 |
| Ballantyne Commons / Durant Blvd | 19.0 | 10 | 24,200 | 40 | 2 |
| Charlotte Town Ave / S Kings Dr | 18.0 | 11 | 35,605 | 35 | 3 |

## TABLE 2 EB Analysis – Summary

| Study Intersection (i) | Before - Average # Crashes per Year ($C_{B,i}$) | After - AADT (veh/day) | Skewness | SPF | Over-dispersion Parameter ($\varphi$) | Relative Weight ($\alpha_i$) | Expected # Crashes($C_E$) |
|---|---|---|---|---|---|---|---|
| East 5th St / 7th St | 8.7 | 23,333 | 0 | 15.19 | 0.29 | 0.0187 | 8.8 |
| I-77 SB Ramp / Sunset Rd | 28.3 | 45,998 | 0 | 35.59 | 0.29 | 0.0081 | 28.4 |
| I-85 SB Ramp / Mallard Creek Ch Rd | 33.3 | 36,359 | 1 | 60.16 | 0.29 | 0.0048 | 33.5 |
| Erwin Rd / Tryon St | 15.3 | 37,904 | 1 | 63.76 | 0.29 | 0.0045 | 15.6 |
| Ballantyne Commons / Durant Blvd | 19.0 | 24,926 | 0 | 16.13 | 0.29 | 0.0177 | 18.9 |
| Charlotte Town Ave / S Kings Dr | 18.0 | 36,673 | 0 | 25.07 | 0.29 | 0.0114 | 18.1 |

## TABLE 3 Estimated, Actual Numbers of Crashes and Ratio

| Study Intersection (i) | After - Estimated # Crashes per year ($C_{E,i}$) | After - Actual # Crashes per Year ($C_{A,i}$) | Ratio = ($C_{A,i}/C_{E,i}$) |
|---|---|---|---|
| East 5th St / 7th St | 8.8 | 8 | 0.9 |
| I-77 SB Ramp / Sunset Rd | 28.4 | 22 | 0.8 |
| I-85 SB Ramp / Mallard Creek Ch Rd | 33.5 | 37 | 1.1 |
| Erwin Rd / Tryon St | 15.6 | 9 | 0.6 |
| Ballantyne Commons / Durant Blvd | 18.9 | 10 | 0.5 |
| Charlotte Town Ave / S Kings Dr | 18.1 | 11 | 0.6 |

# A Methodology for Assessing Safety Impacts of Highway Shoulder Paving

Zongzhi Li[1*], Sang Huyk Lee[2], Yongdoo Lee[3], Bei Zhou[4], Radhika Bamzai[5]

* Corresponding author. Tel.: (312) 567-3556; fax: (312) 567-3519. E-mail address: lizz@iit.edu.

[1]Associate Professor, Department of Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, Chicago, IL 60616

[2345] Graduate Research Assistants, Department of Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, Chicago, IL 60616

## ABSTRACT

This paper introduces and applies an Empirical Bayesian (EB) analysis methodology for assessing highway shoulder paving impacts using Illinois data. Adding new outside paved shoulders is found to be more effective than widening existing paved shoulders and repaving the same width of paved shoulders. Shoulder paving is most effective for multilane highways, followed by two-lane and Interstate highways. Shoulder paving is effective for highways with per lane daily traffic not exceeding 10,000. It is found to be more effective in reducing injury and PDO crashes when the paved shoulder width does not exceed 8ft for Interstate highways and is between 4ft and 8ft for multilane and two-lane highways; and when the combined lane and outside paved shoulder width does not go beyond 20ft. No clear crash pattern is identified among lane widths of 11ft, 12ft, and 13ft for equal total combined lane and shoulder widths ranging from 12ft to 24ft.

## INTRODUCTION

The single greatest category of vehicle crashes on Interstate, multilane, and two-lane highways is run-off-the-road incidents. These crashes often occur at high speed and at night. A potential safety hazard can occur when a vehicle leaves the travel way. This hazard is due to significant material and elevation differences between highway pavement and shoulder surfaces, which can affect vehicle stability, reduce a driver's ability to handle the vehicle, and often cause head-on, sideswipe, rollover, and fixed object crashes. Shoulder paving is recognized as a positive countermeasure to reduce a shoulder dropoff hazard that will accommodate stopped vehicles to avoid encroachment from the travel way, facilitate conduction of maintenance work, provide access for emergency vehicles, and protect pavement structural integrity. A paved shoulder could assist in protecting the road structure damage caused by water infiltration and stray vehicles and help stray vehicles regain control, recover from error, and resume normal travel.

Over the last decade, a number of studies have been conducted to assess safety impacts of outside shoulder attributes. For Interstate highways as part of the National Highway System (NHS) and non-NHS highways, Zegeer et al. (1998) conducted a study on vehicle crash rates and the highway characteristic in seven states- California, Illinois, Maine, Michigan, Minnesota, North Carolina, and Washington. The majority of NHS highways had lane widths of 11ft or more, and many had shoulder width of 5ft or more. The overall crash rates on NHS highways were approximately 10 percent lower than those on non-NHS highways. The distribution of crashes by severity was quite similar for NHS and non-NHS highways. Ksaibati and Crowe (1999) collected Interstate data on 8,785 crashes in Wyoming for period 1991-1995 to evaluate safety impacts of shoulder attributes. The study results revealed that adding a 6-ft new paved shoulder would reduce crashes by 47.5 percent.

For multilane highways, In the Florida study led by Hadi et al. (1995), it was found that increasing unpaved shoulder width was estimated to decrease crash rates on four-lane rural highways. Furthermore, the use of an inside paved shoulder of 4-6ft wide was found to be very effective in decreasing crashes on rural freeways. In particular, using a 6-ft shoulder width could decrease crash rate by 15.7 percent. Souleyrette et al. (2001) collected data on 600 miles of rural four-lane highways in Iowa to assess the safety impacts of shoulder attributes. Based on the study results, effective value from paved shoulders could be obtained with a minimum width of 2-3ft. Wider paved shoulders might not be cost effective except with very high traffic volumes. Fitzpatrick et al. (2005) compared crashes on 882 miles of rural four-lane highways in Texas where 4,662 crashes from 1999-2001 were experienced. It was found that shoulder had a significant impact on safety of rural four-lane highways. Based on predictions of total crashes, the ratios of total crashes of no shoulder, and 1-10ft shoulders compared with 8ft wide shoulders were given as 1.64, 1.54, 1.45, 1.36, 1.28, 1.20, 1.13, 1.06, 1.00, 0.94, and 0.88, respectively.

For two-lane highways, Cottrell (1993) conducted cost analysis of paved shoulders. This study found that when an unpaved shoulder was changed to a 2ft paved shoulder with the remainder unpaved, a 2.6 percent reduction in crash frequency could be

realized. When the lane width was increased by 1ft and 1ft of the shoulder width was paved, a 6.9 percent reduction in crash frequency could be realized as compared with only paving a 2-ft shoulder. Abboud et al. (2001) evaluated the crash experience of two-lane rural highways in Alabama before and after installing 2-ft and 4-ft shoulders for 263 miles and 404 miles of highways in 11 rural Alabama counties. The study could not discern any statistically significant differences in either crash rate or severity rate between 2-ft and 4-ft shoulder installations. Fitzpatrick et al. (2005) compared 4,117 crashes occurred on 3,944 miles of rural two-lane highways in Texas from 1999-2001. It was found that shoulder had a significant impact on safety of rural two-lane highways. Based on predictions of total crashes, the ratios of total crashes of no shoulder, and 1-10ft shoulders compared with the 8ft wide shoulders were given as 1.62, 1.52, 1.43, 1.35, 1.27, 1.20, 1.13, 1.06, 1.00, 0.94, and 0.89, respectively. Örnek and Drakopoulos (2007) analyzed run-off-road crashes in relation to highway features and driver behavior using data on 5,792 miles of rural two-lane highways in the Wisconsin state highways that had 3-ft paved shoulders. It was found that the crash rate would reduce for two-lane rural highways with 3-ft shoulders when additional unpaved shoulder width was provided. However, it would taper off for additional unpaved shoulder widths in excess of 7ft.

This paper introduces an Empirical Bayesian (EB)-based methodology for analyzing safety impacts of shoulder paving in reducing vehicle crashes on rural and urban Interstates, multilane non-Interstates, and two-lane highways maintained by the Illinois Department of Transportation (DOT). This paper is organized as follows: The next section introduces the six basic analytical steps of the proposed EB methodology. The subsequent section discusses methodology application. The last section provides summary and conclusion.

## PROPOSED EB METHODOLOGY

### Main Analytical Steps

For the highway segments with pavement resurfacing and/or shoulder paving treatments, they are classified as three categories: Type I treatment segments with pavement resurfacing and shoulder paving, Type II treated segments with pavement resurfacing only, and other segments with shoulder paving only.

For either Type I treatment (pavement resurfacing and shoulder paving) or Type II treatment (pavement resurfacing only), the difference between the observed and expected EB-adjusted crash frequencies during the after treatment period is an estimate of the safety impacts of the treatment. Figure 1 illustrates the estimation of safety impacts of Type I Treatment and Type II Treatment, respectively. The EB approach for assessing safety impacts of shoulder paving are centered on i) estimating EB-adjusted crash frequencies for Type I and Type II treatments for the before and after treatment periods, ii) determining the types of crashes potentially affected by shoulder paving (regarded as target crashes classified by fatal, injury, and PDO crash severity categories) in Type I and Type II treatments for the after treatment period, and iii) quantifying the safety impacts of shoulder paving. These steps are discussed in details in subsequent sections.

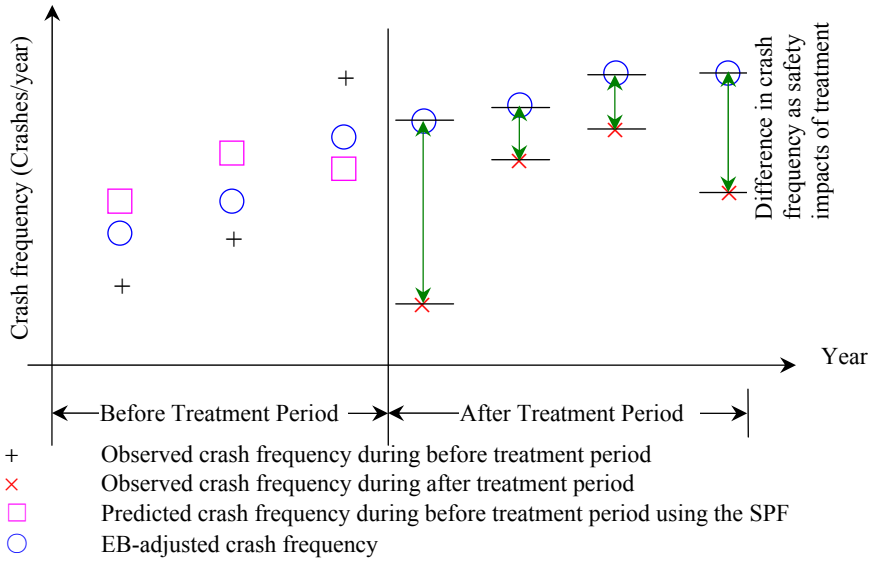



**Figure 1. Illustration of Safety Impacts of Type I Treatment or Type II Treatment**

*Calibration of Safety Performance Functions for Highway Segments*

Data on untreated highway segments (i.e., highway segments without receiving Type I treatment or Type II treatment) in comparison groups need to be used to calibrate safety performance functions that will be used for EB analysis. The three important considerations for a good model are an appropriate set of response variable, explanatory variables, and a model type that adequately explains the relationship between the selected response variable and explanatory variables.

Selection of Response Variable. Two forms of the dependent variables, crash counts and crash rates, were used in the past for the development of safety performance functions. Crash counts refer to the number of observed crashes while the crash rates are the crash counts divided by the traffic exposure (usually in millions of vehicle miles traveled). Recent studies by Hauer (1997) and Tarko et al. (2000) have shown that this assumption may not be valid. The choice of crash counts as the response variable and having the traffic exposure as an independent variable however allows for estimation of the exact relationship between number of crashes and traffic exposure. In this study, the crash counts in terms of number of fatal, injury, PDO, and total crashes per year per segment are chosen as the response variable for developing safety performance functions for rural and urban Interstate, multilane, and two-lane highways, respectively.

Selection of Explanatory Variables. A number of variables have been found to provide explanation for highway segment related vehicle crashes. The primary importance of traffic as an explanatory factor for vehicle crashes relative to other highway variables has long been acknowledged, as found in existing safety performance functions. The explanatory variables used in this study had to be selected from the data elements available in the database provided by the Illinois DOT. Key explanatory variables considered in the current study are:

- Segment Annual Average Daily Traffic (AADT)
- Segment truck volume
- Segment length
- Lane width
- Outside shoulder type
- Outside paved shoulder width
- Median type
- Median width
- Speed limit
- One-way or two-way operation
- Access control
- Truck route designation

Consideration of Model Types. The choice of model form typically depends on the nature of the response variable and its relationships with the explanatory variables. Linear models have been found to be unsuitable for crash modeling because of the non-negative and discrete nature of crash data. Secondly, the error terms of crash data are typically not normally distributed as implicitly assumed in linear modeling (Jovanis and Chang, 1985). Model forms such as the Poisson and negative binomial models have become well accepted for modeling discrete rare events such as crash occurrences (Miaou et al., 1993).

A major limitation of the Poisson regression model is that the variance of the dependent variable is constrained to be equal to its mean. If the data is over-dispersed, i.e., if the variance is greater than the mean then the coefficient vector β will be biased. To overcome such limitation, the negative binomial distribution (a generalized form of the Poisson model) has been recommended (Jovanis and Chang, 1985). The negative binomial distribution adds to the variance a quadratic term that represents the overdispersion. This allows for extra Poisson variation due to variables not included in the model.

The crash dataset may exhibit a large number of observations with zero crash occurrences (i.e., many segment without experiencing any crash in a given year). This case can be handled by zero-inflated Poisson model or negative binomial model. In the model calibration process, both model forms (with or without zero crash inflation) were investigated.

Model Evaluation. Besides selecting an appropriate model form for the analysis, the statistical significance of the estimated regression coefficient for each independent variable was investigated, That is, the null hypothesis that the coefficient is zero should be rejected if the variable is statistically significant. In addition, the sign and magnitude of each estimated coefficient should be practical and justifiable from an engineering viewpoint. For the present study, the modeling process using the Poisson and negative binomial models was accomplished using the NLOGIT econometrics software package (Greene, 2002). The NLOGIT software yields estimates of the coefficients and standard error for each coefficient from which the *p*-values and *t*-statistics can be computed. The likelihood ratio test is a common test used to assess two competing models.

*Computation of EB-Adjusted Crash Frequencies for the before Treatment Period*
For a site receiving either Type I treatment or Type II treatment, the EB adjusted crash frequency that corrects the regression-to-the-mean bias for the before treatment period needs to be estimated to begin with EB analysis. Typically, the EB adjusted crash frequency for the before treatment period is defined as

$$EB_B = w \times P_B + (1 - w) \times O_B \quad (1)$$

where $EB_B$ is the EB-adjusted crash frequency in *n* years for the before treatment period, *w* is the determined weight factor, $P_B$ is the number of predicted number of crashes for the before treatment period, $O_B$ is the total number of crashes observed in *n* years for the before treatment period. Based on this analysis, the longer the observations are made, the smaller the weight factor, which makes the EB-adjusted crash frequency, weighted more towards the observed number of crashes.

The weighting factor *w* for a given highway segment can be established on the basis of the overdispersion parameter determined in the process of calibrating the safety performance functions as

$$w = \frac{1}{1 + \left[ \alpha \times (N \times \text{Segment Length}) \times \sum_{n=1}^{N} P_{n,B} \right]} \quad (2)$$

where $\alpha$ is the overdispersion parameter for crash frequency per mile per year, $P_{n,B}$ is the predicted crash frequency for the highway segment in year *n* for the before treatment period, and *n*=1, 2, …, N.

*Computation of EB-Adjusted Crash Frequencies for the after Treatment Period*
For a site receiving either Type I treatment or Type II treatment, the EB adjusted crash frequency for the before treatment period can be used to predict EB-adjusted crash frequency for the after treatment period that represent the number of crashes to be expected for the after treatment period had such treatment not been implemented. The computation mainly needs to consider changes in traffic volume and segment length between the before treatment period and the after treatment period as

$$EB_A = EB_B \times \left( \frac{\sum_{m=1}^{M} AADT_m}{\sum_{n=1}^{N} AADT_n} \right) \times \left( \frac{Length_A}{Length_B} \right) \quad (3)$$

where $EB_B$ is the EB-adjusted crash frequency for the before treatment period, $EB_A$ is the EB-adjusted crash frequency for the after treatment period had the treatment not been implemented, $AADT_n$ and $AADT_m$ are annual average daily traffic for the before and after treatment periods, and $Length_B$ and $Length_A$ are highway segment length before and after treatment periods.

*Types of Crashes Affected by Shoulder Paving*

The EB-adjusted crash frequencies for fatal, injury, and PDO crashes for the after treatment period represent the number of crashes to be expected if no treatment were to be implemented in the after treatment period. Regardless of Type I treatment or Type II treatment, for fatal, injury, and PDO crash severity categories only some types of crashes are potentially affected by shoulder paving. These types of crashes are called target crash types, which mainly include the following: i) head-on crash with a vehicle in the opposite direction; ii) sideswipe crash with a vehicle in the same direction (for multilane highways); iii) sideswipe crash with a vehicle in the opposite direction; iv) fixed object on the left or right side of the roadway; and v) overturn in road or roadside, on the left or right.

Because shoulder-related crashes must show evidence of a vehicle leaving the road, the above five types of crashes can be re-grouped into the following: i) run-off-the-road right, then occurred head-on and sideswipe crashes with a vehicle in the opposite direction; ii) run-off-the-road right, then occurred sideswipe crash with a vehicle in the same direction (for multilane highways); iii) run-off-the-road right, then collided with fixed object; iv) run-off-the-road right, then overturn in road or roadside involving single vehicle; v) run-off-the-road right, then overturn in road or roadside involving multiple vehicles. As such, the safety impacts of shoulder paving on a highway segment with Type I treatment or Type II treatment in terms of reduction in number of fatal, injury, and PDO crashes for the after treatment period can be separately estimated as

Shoulder paving related reduction in fatal crashes: $\Delta_{A,F} = [EB_{A,F} - O_{A,F}] \times p_{A,F}$ (4)

Shoulder paving related reduction in injury crashes: $\Delta_{A,I} = [EB_{A,I} - O_{A,I}] \times p_{A,I}$ (5)

Shoulder paving related reduction in PDO crashes: $\Delta_{A,P} = [EB_{A,P} - O_{A,P}] \times p_{A,P}$ (6)

where $\Delta_{A,F}$, $\Delta_{A,I}$, and $\Delta_{A,P}$ are shoulder paving related reduction in number of fatal, injury, and PDO crashes for the after treatment period for highway segments with Type I treatment or Type II treatment; $EB_{A,F}$, $EB_{A,I}$, and $EB_{A,P}$ are the EB adjusted crash frequencies for the after treatment period had the treatment not been implemented; $O_{A,F}$, $O_{A,I}$, and $O_{A,P}$ are the actual crash frequencies for the after treatment period; and $p_{A,F}$, $p_{A,I}$, and $p_{A,P}$ are the proportion of head-on, sideswipe-same direction, sideswipe-opposite direction, fixed object, and overturn crashes involving single vehicle, and multiple vehicles in observed fatal, injury, and PDO crashes for the after treatment period, respectively.

*Comparison Groups of Highway Segments for Assessing Safety Impacts of Shoulder Paving*

Having estimated the reduction in fatal, injury, and PDO crashes for the after treatment period for individual highway segments with Type I treatment or Type II treatment, the treated sites need to be properly grouped so that the overall difference of reductions in crashes between the segments in the Type I treatment (pavement resurfacing and shoulder paving) group and Type II treatment (pavement resurfacing

only) group can be used to assess the safety impacts of shoulder paving. In this study, the comparison groups of highway segments with Type I treatment and Type II treatment are separately determined by i) daily traffic per lane, ii) outside paved shoulder width, iii) lane width, and iv) combined lane width and outside paved shoulder width during the after treatment period.

*Safety Impacts of Shoulder Paving*

For rural and urban Interstate, multilane, and two-lane highways, the safety impacts of shoulder paving was assessed using information on shoulder paving related reduction in fatal, injury, and PDO crashes for comparable highway segments with Type I and Type II treatments grouped by traffic volume range, outside paved shoulder width, lane width, and combination of lane width and outside paved shoulder width, respectively. For each group of highway segments with Type I and Type II treatments, the safety impacts of shoulder paving is separately expressed by percentage reduction in fatal, injury, and PDO crashes as

$$Eff_F = \frac{\sum_{k=1}^{K}\left(\frac{\Delta_{k,A,F,Type\ I}}{EB_{k,A,F,Type\ I} \times p_{k,A,F,Type\ I}}\right)}{K} - \frac{\sum_{l=1}^{L}\left(\frac{\Delta_{l,A,F,Type\ II}}{EB_{l,A,F,Type\ II} \times p_{l,A,F,Type\ II}}\right)}{L} \quad (7)$$

$$Eff_I = \frac{\sum_{k=1}^{K}\left(\frac{\Delta_{k,A,I,Type\ I}}{EB_{k,A,I,Type\ I} \times p_{k,A,I,Type\ I}}\right)}{K} - \frac{\sum_{l=1}^{L}\left(\frac{\Delta_{l,A,I,Type\ II}}{EB_{l,A,I,Type\ II} \times p_{l,A,I,Type\ II}}\right)}{L} \quad (8)$$

$$Eff_P = \frac{\sum_{k=1}^{K}\left(\frac{\Delta_{k,A,P,Type\ I}}{EB_{k,A,P,Type\ I} \times p_{k,A,P,Type\ I}}\right)}{K} - \frac{\sum_{l=1}^{L}\left(\frac{\Delta_{l,A,P,Type\ II}}{EB_{l,A,P,Type\ II} \times p_{l,A,P,Type\ II}}\right)}{L} \quad (9)$$

where $Eff_F$, $Eff_I$, and $Eff_P$ are safety impacts of shoulder paving in terms of percentage reduction in fatal, injury, and PDO crashes; $\Delta_{k,A,F,Type\ I}$ and $\Delta_{l,A,F,Type\ II}$, $\Delta_{k,A,I,Type\ I}$ and $\Delta_{l,A,I,Type\ II}$, $\Delta_{k,A,P,Type\ I}$ and $\Delta_{l,A,P,Type\ II}$ are net reductions in fatal, injury, and PDO crashes resulted from shoulder paving for the after treatment period for highway segment *k* with Type I treatment and highway segment *l* with Type II treatment; $EB_{k,A,F,Type\ I}$ and $EB_{l,A,F,Type\ II}$, $EB_{k,A,I,Type\ I}$ and $EB_{l,A,I,Type\ II}$, $EB_{k,A,P,Type\ I}$ and $EB_{l,A,P,Type\ II}$ are EB-adjusted fatal, injury, and PDO crashes for the after treatment period for highway segment *k* with Type I treatment and highway segment *l* with Type II treatment; $p_{k,A,F,Type\ I}$ and $p_{l,A,F,Type\ II}$, $p_{k,A,I,Type\ I}$ and $p_{l,A,I,Type\ II}$, $p_{k,A,P,Type\ I}$ and $p_{l,A,P,Type\ II}$ are the proportion of shoulder-related fatal, injury, and PDO crashes for the after treatment period for highway segment *k* with Type I treatment and highway segment *l* with Type II treatment; *K* and *L* are number of highway segments in Type I treatment group and Type II treatment group, respectively.

## METHODOLOGY APPLICATION

The EB analysis methodology proposed in the previous section was applied using the processed data on untreated and treated segments being involved with Type I or Type

II treatments for period 2000-2006. The following sections discuss calibration of safety performance functions for fatal, injury, and PDO crashes associated with rural and urban Interstate, multilane, and two-lane highways; and EB analysis results.

**Calibration of Safety Performance Functions**

Table 1 summarizes total number of untreated highway segments in comparison groups, mileage, and number of crashes by crash severity from 2000-2006 that was utilized for calibrating safety performance functions for rural and urban Interstate, multilane, and two-lane highway segments, respectively.

**Table 1. Total Number and Mileage of Untreated Highway Segments for Period 2000-2006**

| Highway Class | | Seg | Miles | Total Number of Crashes | | |
|---|---|---|---|---|---|---|
| | | | | Fatal | Injury | PDO |
| Rural | Interstate | 897 | 432 | 113 | 4,541 | 20,583 |
| | Multilane | 765 | 158 | 72 | 6,366 | 23,433 |
| | Two-lane | 1,248 | 583 | 195 | 6,835 | 23,805 |
| Urban | Interstate | 816 | 180 | 137 | 9,257 | 49,217 |
| | Multilane | 2,514 | 375 | 293 | 8,647 | 106,836 |
| | Two-lane | 1,728 | 255 | 100 | 8,041 | 29,575 |

The steps followed for calibrating SPFs are briefly explained as below. First, a preliminary trend analysis was conducted to obtain the distribution of explanatory variables and identify factors that affect occurrence of crashes on roadway segment. Then, the model of choice was formulated by adding each explanatory variable in a stepwise manner to test the impact of its inclusion. The variables were added starting from that deemed most significant on the basis of the preliminary trend analysis results. Finally, SPFs were calibrated by land area/highway functional class (rural/urban Interstate, multilane, and two-lane highway segments) and by crash severity category (fatal, injury, and PDO crashes). Emphases of model validation were given on multicolinearity of explanatory variables, autocorrelation of error terms, and heteroskedasticity of error terms.

**Estimation of Safety Impacts of Shoulder Paving**

*Data Preparation*

Table 2 presents summary information on total number of highway segments for rural and urban Interstate, multilane, and two-lane highways with Type I and Type II treatments implemented in 2000-2006, along with mileage, total number of crashes by crash severity, and shoulder-related target crash percentages out of the total crashes, respectively. These segments are only associated with Illinois DOT maintained highways and do not include tollways.

**Table 2. Data Summary of Type I and Type II Treated Highway Segments for EB Analysis**

| Treatment | Highway Class | | Seg | Miles | Total Number of Crashes | | |
|---|---|---|---|---|---|---|---|
| | | | | | Fatal | Injury | PDO |
| Type I | Rural | Interstate | 68 | 57.52 | 11 | 268 | 1,322 |
| | | Multilane | 4 | 1.29 | 3 | 15 | 152 |
| | | Two-lane | 21 | 22.98 | 2 | 296 | 1,007 |
| | Urban | Interstate | 59 | 15.03 | 16 | 555 | 2,527 |
| | | Multilane | 17 | 5.29 | 0 | 71 | 547 |
| | | Two-lane | 17 | 2.01 | 0 | 14 | 84 |
| Type II | Rural | Interstate | 21 | 12.45 | 1 | 68 | 382 |
| | | Multilane | 55 | 17.64 | 4 | 220 | 944 |
| | | Two-lane | 123 | 68.75 | 18 | 605 | 1,863 |
| | Urban | Interstate | 33 | 7.33 | 8 | 399 | 1,571 |
| | | Multilane | 166 | 26.99 | 17 | 1,839 | 6,312 |
| | | Two-lane | 149 | 22.75 | 10 | 494 | 1,924 |

*EB Analysis Results*

The safety impacts of shoulder paving for rural and urban Interstate, multilane, and two-lane highways were separately assessed by grouping the highway segments with Type I treatment and Type II treatment by per lane daily traffic range, outside paved shoulder width, lane width, and combined lane and outside paved shoulder width. For each of these assessments, safety impacts in terms of percentages of fatal, injury, and PDO crash reductions were separately established for segments involved with repaving the same width of a paved shoulder, widening the paved shoulder from a certain paved width, and adding a new paved shoulder from a "0" paved shoulder width, respectively.

Safety Impacts of Shoulder Paving for Highway Segments Classified by Per Lane Daily Traffic. In general, adding new paved shoulders and widening existing paved shoulders are more effective than repaving the same width of paved shoulders in crash reductions. Shoulder paving is most effective for multilane highways, followed by two-lane highways and then Interstate highways. Shoulder paving is effective for repaving and widening shoulders of rural Interstate highways with per lane daily traffic not exceeding 10,000, for repaving and widening shoulders of urban Interstate highways with per lane daily traffic within 5,000-10,000, for repaving shoulders of multilane highways with per lane daily traffic not exceeding 10,000, for widening and adding shoulders of multilane highways with per lane daily traffic ranging 5,000-10,000, and for repaving, widening, and adding paved shoulders of two-lane highways with per lane daily traffic within 5,000-10,000, respectively.

Safety Impacts of Shoulder Paving for Highway Segments Classified by Outside Paved Shoulder Width. For all classes of highways with sufficient data for EB analysis, repaving existing shoulder and adding new paved shoulders are found to be effective in reducing crashes when the paved shoulder width is within 8ft. In many case, shoulder related crashes generally increase for fatal, injury, and PDO crashes when the paved shoulder width is wider than 8ft, with the exception of widening

shoulders of urban Interstate highways that still experience crash reductions when the outside paved shoulder width exceeding 8ft.

Safety Impacts of Shoulder Paving for Highway Segments Classified by Lane Width. For all classes of highways with sufficient data for EB analysis, shoulder paving is found to be effective in reducing crashes when the lane width is within 12ft. Shoulder related crashes generally increase for fatal, injury, and PDO crashes when the lane width exceeds 12ft.

Safety Impacts of Shoulder Paving for Highway Segments Classified by Combined Lane/Outside Paved Shoulder Width. For all classes of highways with sufficient data for EB analysis, shoulder paving is found to be effective in reducing crashes when the combined width is within 20ft. Shoulder related crashes generally increase for fatal, injury, and PDO crashes when the paved shoulder width is wider than 20ft. In general, widening existing paved shoulders and adding new paved shoulders is more effective for multilane highways than two-lane highways.

## SUMMARY AND CONCLUSION

This study has conducted an in-depth investigation into safety impacts of highway shoulder attributes in Illinois. It began with a review of existing studies on safety impacts of shoulder paving and then discussed study objectives. An EB-based methodology consisting of six major analytical steps was proposed for assessing safety impacts of shoulder paving. Data on Illinois State-maintained highways for period 2000-2006 were used to apply the individual analytical components. The folllowing findings are drawn:

- The existing studies have revealed that adding outside paved shoulders up to no more than 8ft could reduce shoulder-related crashes with considerably large variations. The outside paved shoulder width goes beyond 8ft may increase shoulder-related crashes, especially severe crashes. Several studies also suggested that shoulder paving is more beneficial for high volume roads.

- Adding new outside paved shoulders appears to be more effective than widening existing outside paved shoulders and repaving the same width of outside paved shoulders. Among all highway classes, shoulder paving is most effective for multilane highways, followed by two-lane and Interstate highways. Shoulder paving is effective for repaving and widening shoulders of rural Interstate highways and for repaving shoulders of multilane highways with per lane daily traffic not exceeding 10,000. It is effective for repaving and widening shoulders of urban Interstate highways, for widening and adding shoulders of multilane highways, and for repaving, widening, and adding new paved shoulders of two-lane highways with per lane daily traffic within 5,000-10,000.

- Shoulder paving was found to be more effective in reducing injury and PDO crashes when the paved shoulder width does not exceed 8ft for Interstate highways and is 4-8ft for multilane and two-lane highways; and when the combined lane and outside paved shoulder width does not go beyond 20ft. No clear pattern in terms of increase in shoulder-related crashes was identified among

lane widths of 11ft, 12ft, and 13ft for equal total combined lane and outside paved shoulder widths ranging from 12ft to 24ft.

## REFERENCES

Abboud, N.K. (2001). Evaluation of Two- and Four-Foot Shoulders on Two-Land State Routes. ITE Journal. Institute of Transportation Engineers, Washington, D.C.

Fambro, D.B. (1981). Safety Benefits of Paved Shoulders on Rural Two-Lane Highways. *SAFE Journal* 11(2), 22-25.

Fitzpatrick, K., W.H. Schneider, and E.S. Park (2005). Comparisons of Crashes on Rural Two-Lane and Four-Lane Highways in Texas. Texas A&M University, College Station, TX.

Greene, W.H. (2002). NLOGIT Version 3.0 User's Manual. Econometric Software Inc., Bellport, NY.

Hadi, M.A., J. Aruldhas, L-F. Chow, J.A. Wattleworth (1995). Estimating Safety Effects of Cross-Section Design for Various Highway Types Using Negative Binomial Regression. *Transportation Research Record* **1500**, 169-177.

Hauer, E. (1997). *Observational Before-After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*. Pergamon Press, Elsevier Science, Ltd, Oxford, UK.

Jovanis, P.P., H. Chang (1985). Modeling the Relationship of Crashes to Miles Traveled. *Transportation Research Record* **1068**, 42-51.

Ksaibati, K., I. Crowe (1999). Evaluation of Pavement Shoulders. Department of Civil and Architectural Engineering, University of Wyoming, Laramie, WY.

Miaou, S. P., P.S. Hu, T. Wright, S.C. Davis, A.K. Rathi (1993). Development of Relationship between Truck Crashes and Geometric Design: Phase I. *Report Number FHWA-RD-91-124.* Federal Highway Administration. Washington, D.C.

Örnek, E. (University of Wisconsin, Madison), A. Drakopoulos (Marquette University) (2007). Analysis of Run-Off-Road Crashes in Relation to Roadway Features and Driver Behavior. Proceedings of the 2007 Mid-Continent Transportation Research Symposium, Ames, Iowa.

Souleyrette, R., T. McDonald, Z. Hans, A. Kamyab, T. Welch, B. Storm (2001). Paved Shoulders on Primary Highways in Iowa: An Analysis of Shoulder Surfacing Criteria, Costs, and Benefits. Iowa State University, Ames, IA.

Tarko, A.P., K.C. Sinha, S. Eranky, H. Brown, E. Roberts, R. Scinteie, S. Islam (2000). Crash Reduction Factors for Improvement Activities in Indiana. Joint Transportation Research Program, Purdue University, West Lafayette, IN.

Zegeer, C.V., H.F. Huang, J.R. Stewart, R. Pfefer, J. Wang (1998). Effects of a Towaway Reporting Threshold on Crash Analysis Results. *Transportation Research Record* **1635**, 49-57.

# Opinions of Drivers on Set Belt Use based on Voluntary Responses

Sunanda Dissanayake[1] and Abhishek Parikh[2]
[1]Department of Civil Engineering, 2118 Fiedler Hall, Kansas State University, Manhattan, KS, 66506, Tel: (785) 532 – 1540, Fax: (785) 532 – 7717, Email: Sunanda@ksu.edu
and
[2]Dowling Associates Inc.428 J Street, Sacramento, CA 95833, Tel: (916) 266-2190, Email: abhiparikh@yahoo.com

## ABSTRACT

Even though the effectiveness of seat belts is widely known and accepted, seat belt usage remains relatively low in the United States. In addition to the type of seat belt law and associated enforcement practices, there seem to be many driver characteristics and behavior related to non-use of seat belts. Therefore, this study conducted road-user surveys in Kansas with the intention of identifying human factor-related issues that are playing a role in relation to seat belt use. Participation in the survey was completely voluntary even though an effort was made to cover as many groups as possible. Perceptions, attitudes, understandings, stated compliance levels, potential motivators, etc. of road users were obtained through the survey in order to suggest more effective countermeasures to improve seat belt use. Based on identified critical areas, more focused education and training programs need to be developed. In addition, enforcement levels need to be increased, specifically in areas where the most benefit in terms of seat belt usage could be achieved. It was however, interesting to observe that even drivers themselves agree that stricter laws, higher fines, and other penalties are helpful in increasing the self-discipline needed to wear seat belts more frequently.

## INTRODUCTION

Even though the effectiveness of seat belts is widely known and accepted, seat belt use remains low in the United States compared to Australia, Canada, and some European countries, which have average usage rates above 90 percent. Seat belt use rates in the U.S. vary from state to state reflecting differences in seat belt laws, public attitude, enforcement practices, legal provisions, education programs, etc. (Balci and Vertiz, 2001).

In the United States, a mandatory seat belt law was first enacted in New York in 1984. Lund et al. (1987) found a nine percent decline in traffic fatalities in the first nine months when New York enacted mandatory seat belt law. By 1996, all states except New Hampshire had enacted mandatory seat belt laws, which consist of two categories, known as Primary Seat Belt Law and Secondary Seat Belt Law. Primary law has more power and allows police officers to stop and cite motorists solely for not wearing seat belts. However, under secondary law, police officers can penalize motorists for not wearing seat belts only if they are stopped for some other traffic infraction. As of early 2007, 26 states plus District of Columbia had enacted primary seat belt law and 24 states had secondary seat belt law (Insurance Institute for Highway Safety 2007). New Hampshire is the only state where adults

are legally not obliged to wear seat belts. General effectiveness of primary seat belt law in increasing seat belt usage is evident from the fact that average seat belt usage rate for states with primary law is about 11 percent more than that of secondary law the states (Glassbrenner and Ye, 2006).

In addition to the type of seat belt law and associated enforcement practices, there seem to be human factors related to non-use of seat belts. These non-quantifiable matters and their effects are rather difficult to be captured through methodologies such as statistical modeling. Therefore in this study, road-user surveys were conducted with the intention of identifying human factor-related issues that are playing a role in relation to seat belt use. Perceptions, attitudes, understandings, stated compliance levels, potential motivators, etc. of road users were obtained through the surveys in order to suggest more effective countermeasures to improve seat belt use.

## METHODOLOGY

Surveys were conducted in Kansas at various locations throughout the state. The survey was expected to directly capture general understandings, compliance, attitudes, and behaviors of road users in Kansas and tie those to characteristics of the respondents. A survey questionnaire of 37 questions was prepared. The questions can be divided into four main categories: general characteristics, awareness of seat belt issues, seat belt use patterns and related factors, and others. For example, the category of "general characteristics" included questions about gender, median income, age group, marital status, etc. The "awareness of seat belt issues" category included questions about type of seat belt law in Kansas, penalty for not wearing seat belts, etc. Similarly, the category "seat belt use patterns and related factors" included questions about frequency of seat belt usage, factors that would motivate higher seat belt use, factors influencing seat belt usage negatively, etc. Each category of question consisted of a reasonable amount of options to accommodate various types of responses. The survey form was developed considering factors affecting seat belt usage in United States, as discussed in the literature review, while referring to a compendium of surveys (Huang and Preston, 2004) which consists of questionnaires prepared in various countries for similar purposes.

The surveys were conducted by a group of four people going door to door in residential areas. Locations included both rural and urban areas and clusters of residential areas were selected without any particular regard, but with efficiency in mind. Surveys were administered during daytime hours on both weekdays and weekends. The purpose of the survey was explained and the general public was requested to complete the survey form, where a reasonably satisfactory response rate was observed. Residential communities and other areas were selected randomly, and a total of 794 surveys responses were obtained. After discarding incomplete survey forms, about 753 survey forms were used for analysis purposes.

## RESULTS OF THE SURVEY

This section presents details of the self-reported behavior of drivers in relation to seat belt usage. The first part briefly shows characteristics of the respondents and relationships among seat belt use, awareness of seat belt law, and other factors influencing use of seat belts are then presented in the form of charts.

**Characteristics of Respondents**

Characteristics and distribution of respondents by categories such as age group, gender, median household income, employment status, type of employment, educational background, frequency and amount of driving, type of vehicle driven, etc. are considered in the study. The main purpose of analyzing the characteristics of the respondents was to see whether they were a representative sample of the population being studied. However, it was not possible to strictly control the characteristics of the respondents as the participation was voluntary. Young drivers for example, were much more willing to complete the survey form compared to other age groups, making them over-represented among the respondents. In the absence of an approach to require road users to respond, distribution of respondents within each category was assumed to be reasonably acceptable.

**Seat Belt Law Awareness**

As shown in Figure 1, 50% of the respondents indicated that the seat belt law in Kansas is primary, which implies that half of the population is unaware that the seat belt law in Kansas is secondary. (Note: Type of law was explained in the survey form instead of using the words primary and secondary.) Only 43% of respondents answered the question correctly by saying Kansas has secondary seat belt law, and 6% said they did not know of the type of existing seat belt law. Hereafter, those who responded that the seat belt law in Kansas was primary, or that they were unaware of the seat belt law are referred as "Unaware" about the seat belt law and those who answered that seat belt law was secondary are referred as "Aware".

The survey also included a question on the penalty in Kansas for violating the seat belt law. Responses to this question are presented in Figure 2, which indicates that only 15% of respondents were aware of the correct amount of penalty/fine, which was $ 10 at the time of the study.

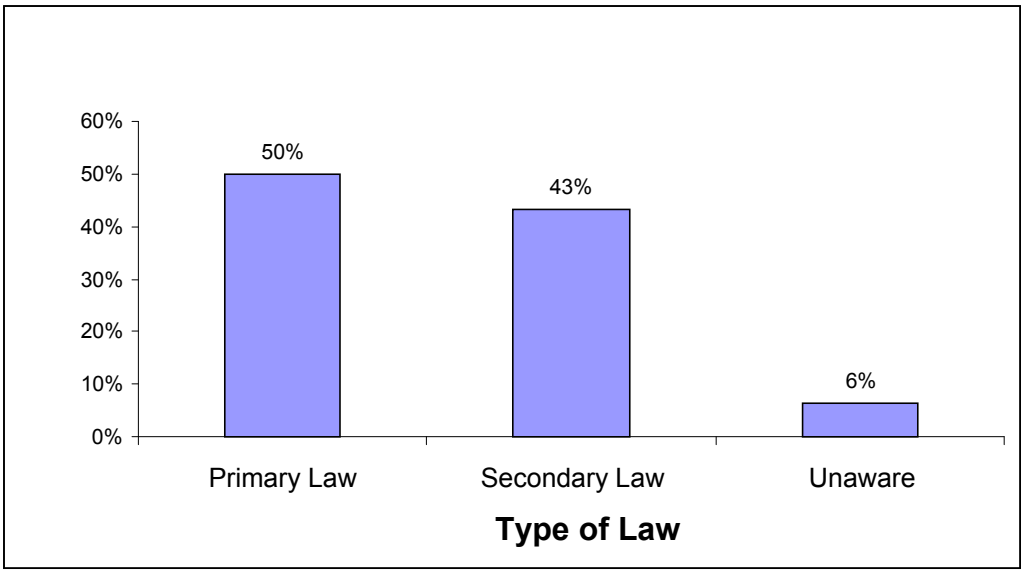



**FIGURE 1 Awareness of the existing seat belt law.**

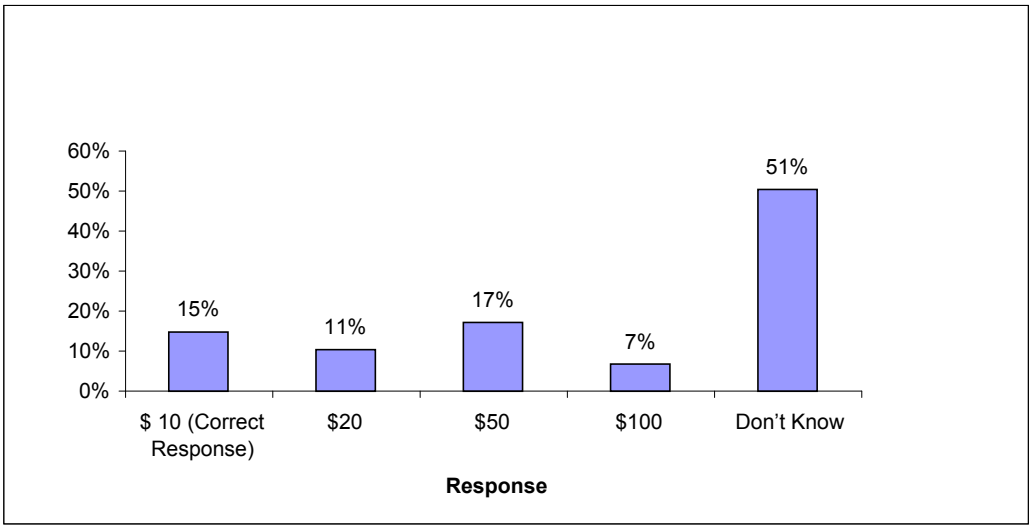



**FIGURE 2 Awareness of penalty for seat belt violation.**

Awareness of the seat belt law was then tied to the characteristics of the respondents. Relationships between gender, educational background, age group, median household income, etc. and awareness of the law were developed. Gender and educational background did not show any clear differences in terms of awareness of the seat belt law. However, it was found that seat belt law awareness was higher among higher income groups and higher age groups. In general, it could be said that the higher the income and age, the more knowledgeable people were of the seat belt law, which conforms to the perceived outcome. When considering awareness of the seat belt law by ethnicity, clear differences existed between groups. Hispanics were found to have the lowest awareness, where only 27 % were familiar with the type of law, followed by African-Americans (33 %) and Caucasian whites (47%). More than 59 % of Asian/Pacific islanders were aware of the seat belt law, but as the number of responses from this group was only 22 out of a total 683 completed responses on this question, the results may not be statistically reliable.

**Seat Belt Use Patterns**

Self-reported behavior related to seat belt use is presented in Figure 3, where 59% of respondents said that they 'always' use seat belts and about 2% said they 'never' use seat belts. Twenty percent of respondents indicated they almost always use seat belts as compared to 4% who said they use seat belts just occasionally.

Self-reported seat belt use has also been categorized by various other characteristics of the respondents. Figure 4 represents stated seat belt use by gender, which indicates that 70% of female respondents belong to the 'always use seat belt' category as compared to 48 % of males. Similarly, only 1 % of females belong to the 'never use seat belts' category as compared to 2 % males. This indicates that self reported seat belt usage was lower among males than females by a huge margin.

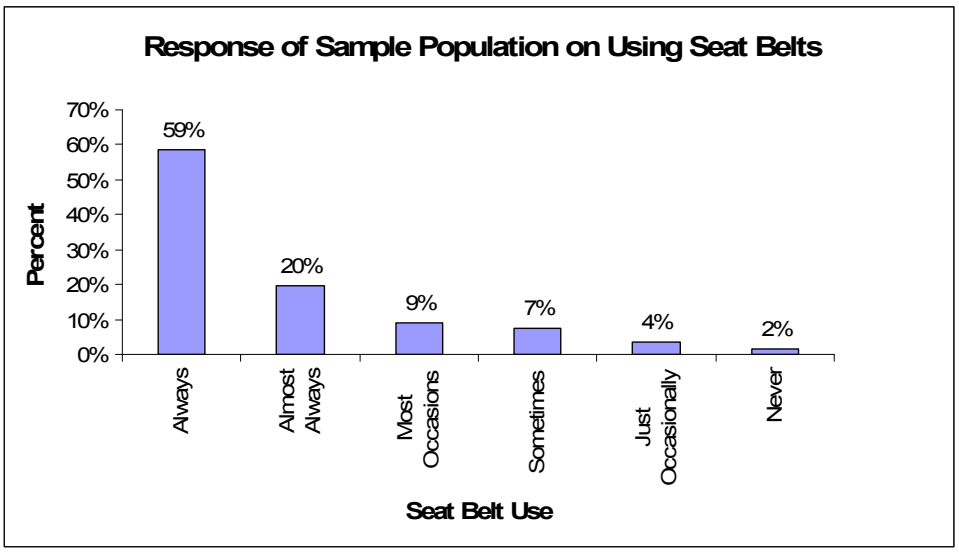



**FIGURE 3 Self-reported seat belt use among respondents.**

Another general perception is that older adults have higher seat belt usage compared to teens and younger age groups. From Figure 5, it can be seen that the percentage of people belonging to the 'always use seat belt' category increases with increased age. For example, 48 % of 16-24 year olds, 69 % of 45-54 year olds, and 78 % of older than 65 years belong to the 'always use seat belt' category. Similarly, respondents in the 'never use seat belts' category decrease as age increases. This confirms that seat belt use is lower among younger age groups, which is consistent with the general understanding and previous research. Therefore, young drivers could be identified as one of the critical groups that need focused attention to increase seat belt usage rates.

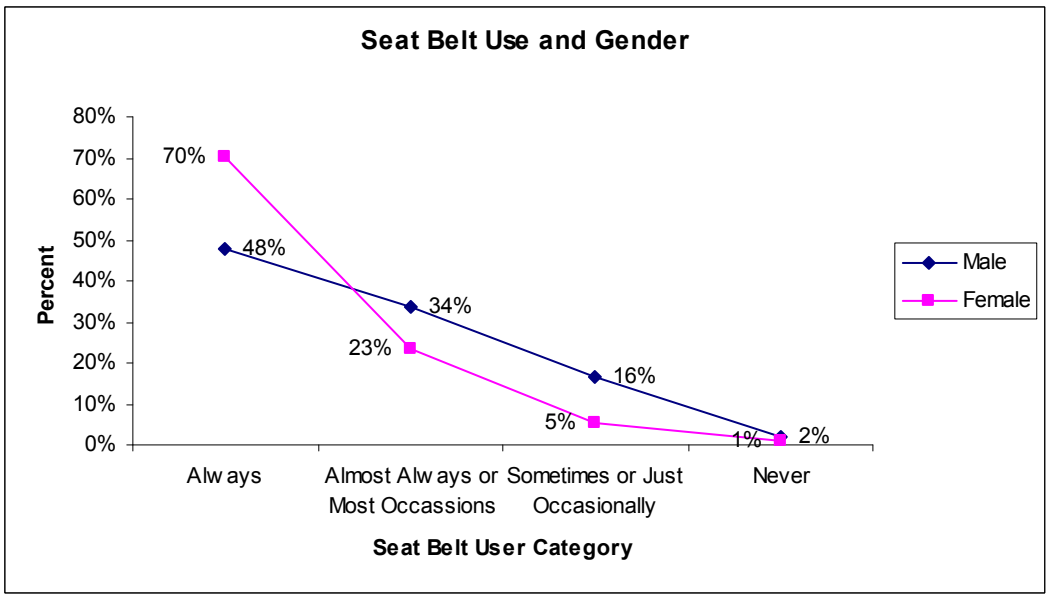



**FIGURE 4 Self-reported seat belt use by gender.**

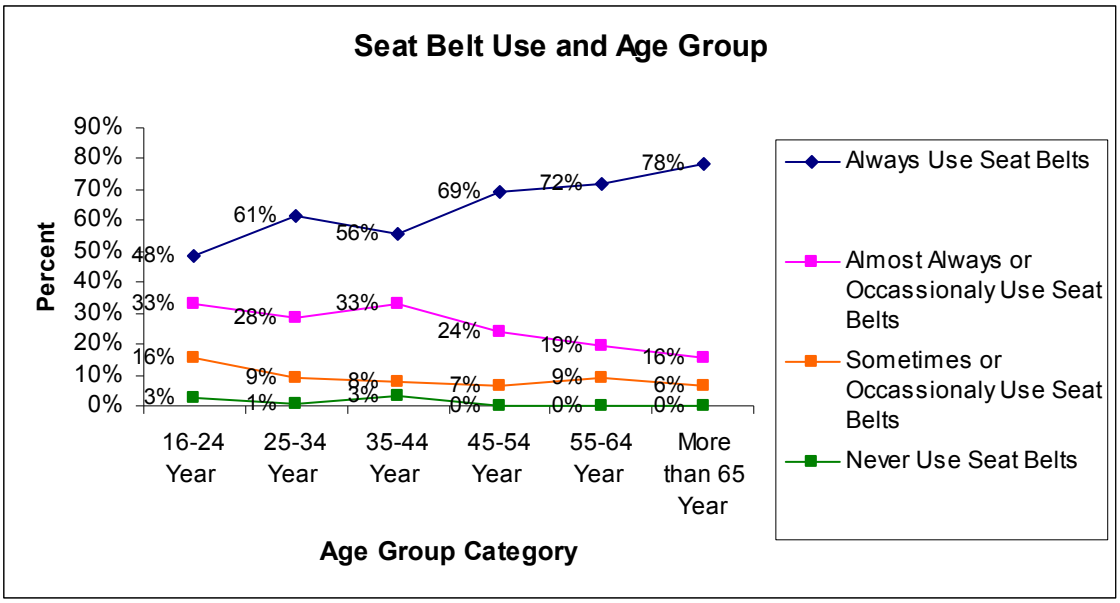



**FIGURE 5 Self-reported seat belt use by age group.**

It is generally believed that seat belt use is low among some ethnic groups compared to others, even though previous research has indicated mixed results. Some studies have found lower seat belt usage among African Americans and Hispanics as compared to Caucasian whites, whereas other studies have found no difference in seat belt usage among ethnic groups. Results of this analysis, as shown by Figure 6, suggest that respondents in the 'always use seat belt' category are highest among Asians and Pacific Islanders at 76 % and lowest among Hispanics at 52 %. However, the percentage of African Americans who stated that they always use seat belts is high at 64 % as compared to 58 % for Caucasian whites.

**Factors Influencing Seat Belt Usage**

Along with the characteristics of the respondents there could be many other factors that influence use of seat belts either positively or negatively. Figure 7 presents reasons given by the respondents for not wearing seat belts. The highest number of respondents said they were less likely to use seat belts while driving a short distance, about 25% said they forget to put on the seat belts, and about 19% said they use them less often when they are in a hurry.

Along with these reasons that discourage use of seat belts, there were also some factors that affected seat belt usage positively, as shown in Figure 8. For example, it was found that 22% of respondents were more likely to use seat belts while driving in bad weather conditions. Seventeen percent said they were more likely to use seat belts while driving with a child and another 17% responded that they were more likely to use seat belts while driving at night.

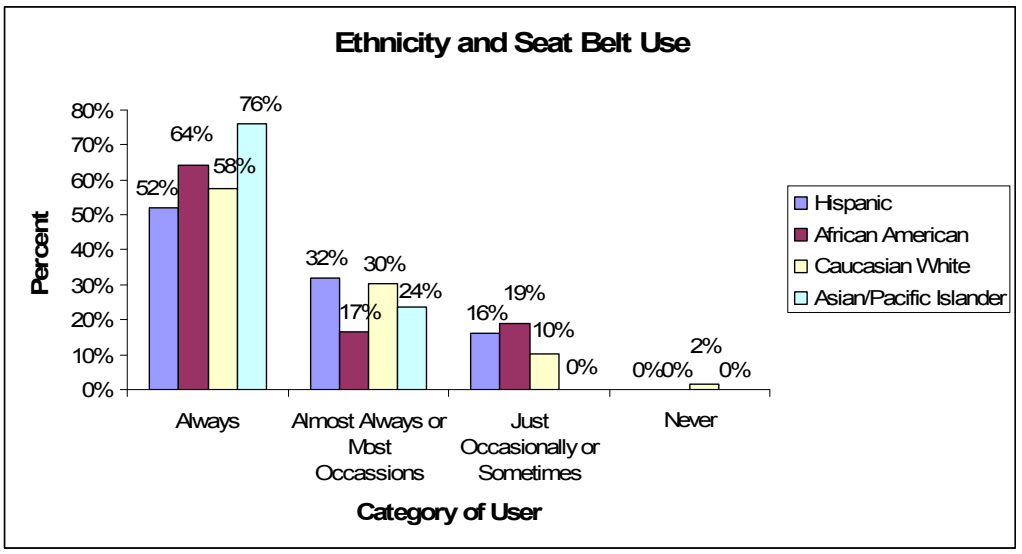
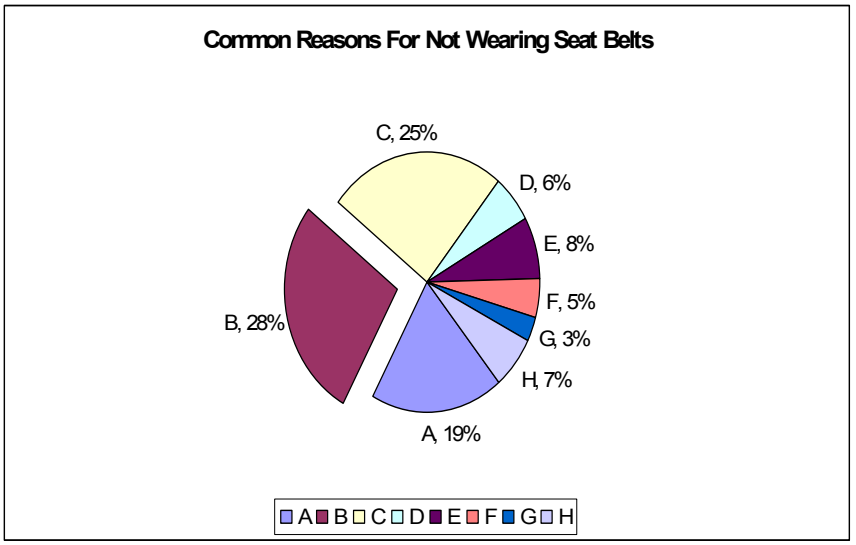


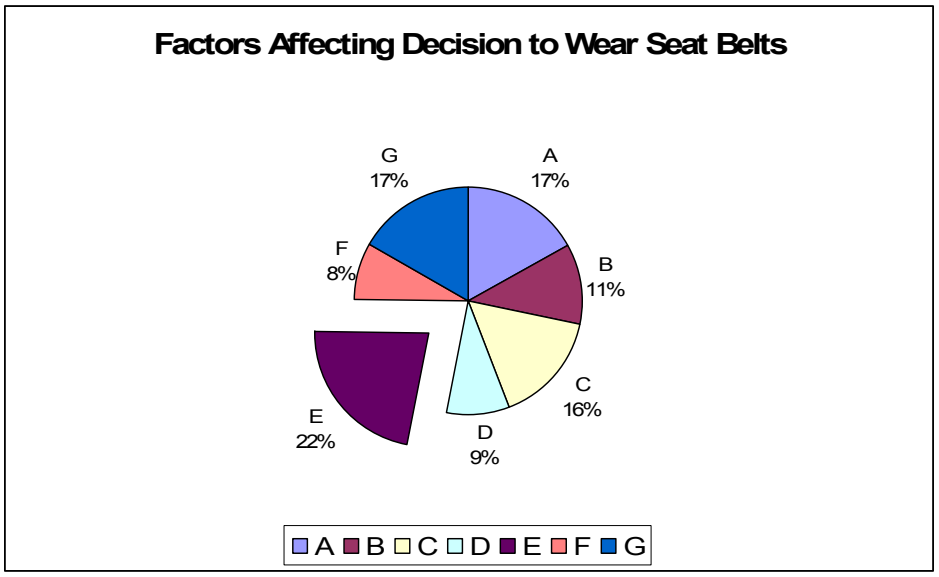

**FIGURE 6 Self-reported seat belt use by ethnicity.**




Do not wear seat belts: A = when in a hurry, B = when driving a short distance, C = as they forget to put it on, D = when driving in light traffic, E = since it is uncomfortable, F = when they assume that chances of crashes are low, G = when people traveling with them don't wear seat belts, and H = when they are driving in rural areas

**FIGURE 7 Reasons for not wearing seat belts.**

Other factors that affected seat belt use behavior were type and characteristics of trips and roads. As shown in Figure 9, based on the question of type of trips where respondents were more likely to wear seat belts, 38% said they were more likely to wear seat belts for work/school-related trips as compared to 27 % saying the same for personal trips/running errands. These results may be useful in identifying locations where seat belt usage is likely to be low, and increasing enforcement accordingly in those locations would be more effective. When type of road was considered, 51 % of respondents indicated they were more likely to wear seat belts on interstate roads and freeways as compared to 14% on two-lane roads without a median.




A = Driving at night, B = After being involved in a crash, C = Strict law enforcement, D = When driving under influence of alcohol, E = Bad weather, F = When passengers are under the influence of alcohol, G = While driving with a child

**FIGURE 8 Factors positively affecting decision to wear seat belts**

In order to identify measures for increasing seat belt usage rates in the target area, respondents were asked what would motivate them to use seat belts more often. These results are presented in Figure 10, which indicate that one-fifth of the surveyed sample identified stricter enforcement as a factor that would motivate them to wear seat belts. Similarly, 'reminder from someone', 'increase in insurance if ticketed for not wearing seat belts', 'higher fine' and 'stricter seat belt laws' were factors supported by 18, 17, 15 and 14 % respectively of the respondents, who said those factors would motivate them to use seat belts more often. Education/training program was the least supported factor, where only 4% respondents said that it would motivate higher seat belt use. These findings indicate that drivers themselves like stricter laws and other punishments to improve their own safety, perhaps because they are not able to maintain high levels of self-discipline.

## SUMMARY AND CONCLUSIONS

From the results obtained by this analysis, some important conclusions can be drawn, which are summarized as follows.

**Awareness of Seat Belt Law** - General awareness of seat belt law among Kansans is quite low. It was found that awareness is lower among lower income groups, younger age groups, and among Hispanics and African Americans compared to Caucasian whites.

**Seat Belt Use –** About 59 % of the surveyed sample said they always use seat belts. Stated usage was found to be lower among males, younger drivers, and lower income groups. Hispanics were found to have the lowest self-reported seat belt usage, and Asian/Pacific islanders were found to have the highest usage.

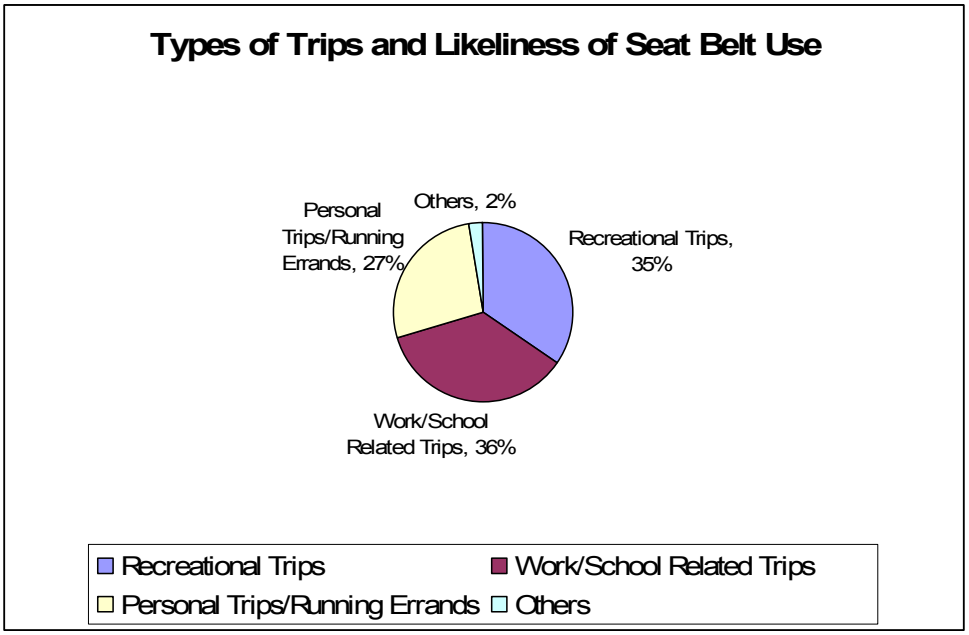
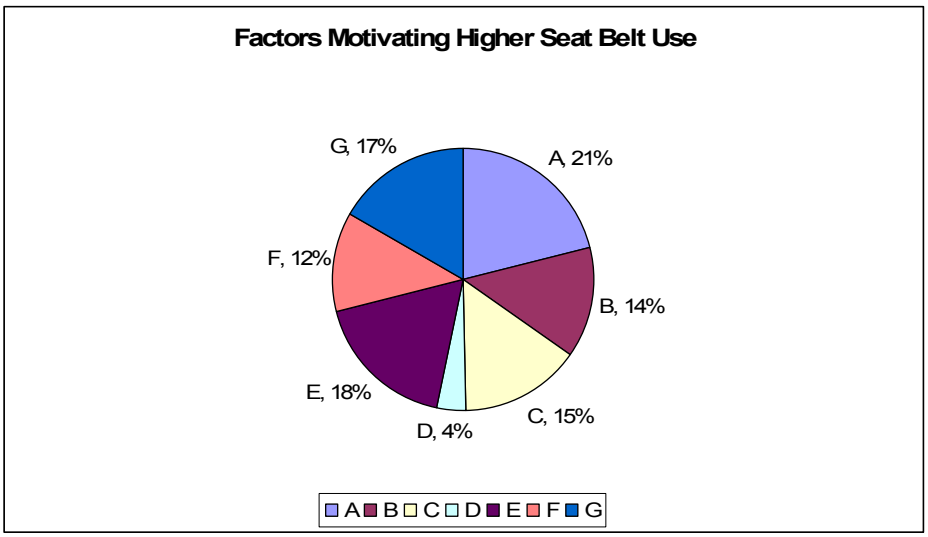



**FIGURE 9 Types of trips where seat belt use is more likely.**

**Reasons for not using seat belts more often** – Some of the more frequent reasons identified by respondents for not using seat belts were when driving a short distance, because of forgetfulness, and when in a hurry. About two-thirds of the sample gave one or more of these reasons for not wearing seat belts. Moreover, by category of trips, seat belt usage was more likely to be high for work/school-related trips and less likely for personal trips/running errands. By type of facility, stated usage was higher on interstate roads and lower on rural roads.

**Factors Motivating Seat Belt Usage –** Some of the factors that were found to positively affect the decision to wear seat belts included bad weather conditions, driving with a child, driving at night, strict law enforcement, reminder from someone for using seat belts, higher fines, and increase in insurance if ticketed for not wearing seat belts. Even though some of these factors are uncontrollable, other findings indicated that drivers like stricter laws and other punishments to improve their own safety, perhaps because they are not able to maintain high levels of self-discipline by themselves.

The study findings indicated that human factors play a major role in the decision to use or not use seat belts. Based on the identified critical areas more focused education and training programs need to be developed. In addition, enforcement levels need to be increased, specifically in areas where most benefits in terms of seat belt usage could be achieved. It is however, interesting to observe that even the drivers themselves feel that stricter laws and higher fines and other penalties are going to motivate them to increase the self-discipline to use seat belts more frequently. Based on the findings of this study, state of Kansas was considering making changes to the existing seat belt law to improve the usage rates.




A = Strict law enforcement, B = Stricter seat belt law, C = Higher fine, D = Education/Training program, E = Reminder from someone, F = Automatic reminder system in car, G = Increase in insurance, if ticketed for not wearing seat belts.

**FIGURE 10 Factors motivating increased use of seat belts.**

## REFERENCES

Balci, R. and A. Vertiz, "Comfort and Usability of Seat Belts, SAE Technical Paper Series." SAE 2001 World Congress, 2001.

Glassbrenner, D. and J. Ye. "Seat Belt Use in 2006-Overall Results." *Traffic Safety Facts Research Note,* DOT HS 810 677, National Highway Traffic safety Administration, National Center for Statistics and Analysis, U.S. Department of Transportation, Washington, D.C., 2006.

Huang, B. and J. Preston. *Road User's Attitude and Behavior: Compendium of Survey Contents*. University of Oxford, Transport Studies Unit. 2004.

Insurance Institute for Highway Safety. Laws and Regulations Homepage, Seat Belt Use Laws as of May 2007. http://www.iihs.org/laws/state_laws/restrain3.html, Accessed July 10, 2007.

Lund, A. K., D. F. Preusser, and A. F. Williams. (1987). "The effect of New York's Seat Belt Use Law on Teenage Drivers," *Accident Analysis and Prevention*, Vol. 19, (1987): 73-80.

National Highway Traffic Safety Administration. *Seat Belt Use in 2006-Use Rates in the States and Territories,* Traffic Safety Facts Crash Stats, DOT HS 810 690, National Center for Statistics and Analysis, U.S. Department of Transportation, Washington, D.C., 2007

National Highway Traffic Safety Administration. *Traffic Safety Facts 2008,* Early Edition, DOT HS 811 170, National Center for Statistics and Analysis, U.S. Department of Transportation, Washington, D.C., 2009.

# Driver Perceptions at Free Right-Turn Channelized Intersections

Gregory S. Macfarlane[1], Mitsuru Saito[2], and Grant G. Schultz[3]

[1] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332; PH (801) 616-9822; email: gregmacfarlane@gmail.com
[2] Department of Civil and Environmental Engineering, Brigham Young University, Provo, UT 84602; PH (801) 422-6326; email: msaito@byu.edu
[3] Department of Civil and Environmental Engineering, Brigham Young University, Provo, UT 84602; PH (801) 422-6332; email: gschultz@byu.edu

## ABSTRACT

Free right-turn channels are sometimes used to alleviate vehicle delay at signalized intersections, but motorists often yield unnecessarily when using these channels. To better understand why motorists do not proceed when they have an available path, the authors conducted a national survey. Approximately one thousand survey participants viewed flashcards of intersection treatments and were asked to identify the proper driving behavior from a list of multiple choices. The answers to the questions were compared in Chi-squared cross tabulation analysis. The results showed that statistically significant proportions of people incorrectly indicate that right-turn behavior is similar at standard right-turn pockets or lanes, yield right-turn channels, and free right-turn channels. Men and women were also noted to have statistically different responses. The survey results further indicate that motorists may not distinguish between intersection treatments, prompting better signage and motorist education or the revision of right-turn delay models for different intersection types.

## INTRODUCTION

Right-turns at signalized intersections are normally not a major design consideration, as almost all states permit turns to be made even on red lights (Roess et al. 2004). But sometimes traffic patterns at individual intersections necessitate special treatments. As the demand of right-turning vehicles increases, engineers typically move through the following hierarchy in designing for the right-turn demand:

- Right-turn lanes or pockets,
- Right-turn channelization,
- Right-turn channelization with deceleration and acceleration lanes, and
- Right-turn channelization with an added lane exiting the channel, sometimes called a free right-turn channelization.

In almost all cases, yield signs are used to ensure the safe exit of right-turning vehicles from the intersection. With the free right-turn channelization, however, no

signs are used because there is no actual conflict with merging vehicles at the exit of the channel. Schematic diagrams of the intersection types are provided in Figure 1.

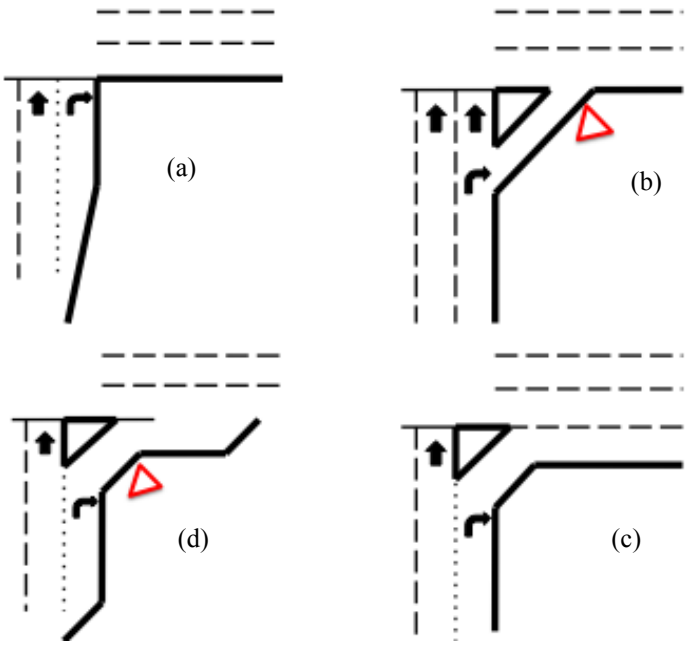



**Figure 1. Right-turn treatments at intersections (clockwise from top left): (a) standard right-turn lane or pocket; (b) yield right-turn channel; (c) free right-turn channel; (d) yield right-turn channel with acceleration lanes.**

It is standard engineering practice to assume that vehicles using a free right-turn channel will not contribute to vehicle delay at an intersection. Neither the Highway Capacity Manual (HCM) (TRB 2000) nor the Intersection Capacity Utilization (ICU) Method (Husch and Albeck 2003) provide any procedures for calculating the delay based on right-turning vehicles when there is a free right-turn channel. Engineers merely assume that automobiles will make the movement at a comfortable turning speed without stopping for conflicting vehicles. It is important to remember throughout this paper that although no actual conflict exists, motorists sometimes anticipate a conflict; thus "conflicting" vehicles in this report are only conflicting psychologically. Because no real conflict exists, currently used analytical techniques assume a "zero delay" condition at free right-turn channels.

## Purpose and Need

Rudimentary observation of free right-turn channels demonstrates that the zero delay assumption may be unfounded. The authors have demonstrated in a related study that many drivers will in fact remain inside the turn channel while waiting for conflicting vehicles to clear, even during peak hours when more familiar drivers use the intersections (Macfarlane 2009). In some severe scenarios, this could contribute to significant vehicle delay at an intersection that would escape current analytical methods. An explanation of this delay is necessary to improve the accuracy of level of service (LOS) predictions.

Because right-turn channels, and free right-turn channels in particular, are relatively rare, it may be possible that motorists are not aware of proper driving behavior. It is important that engineering practice and public perception align, otherwise, fundamental assumptions built into engineering models could prove false. The authors conducted a survey to ascertain if the driving public's understanding of varying right-turn treatments matches the expectations of engineering models. In the survey, participants were directed to look at flashcards with intersection diagrams, and respond with their understanding of the proper driving behavior. This paper presents a statistical analysis of the survey responses.

## LITERATURE REVIEW

The 2000 edition of the HCM (TRB 2000) contains specific, and sometimes lengthy, procedures for analyzing the delay at intersections estimated by calculating vehicle volumes, capacities and necessary parameters. At present there is no specific methodology for examining the delay of vehicles in a turn channel, much less a free right-turn channel.

When calculating the capacity of a right-turn lane, the saturation flow rate of the lane is simply multiplied by a factor of 0.85 to account for the slower speed of vehicles preparing to turn (TRB 2000). On a red light, the right turn on red (RTOR) volume could be estimated using the unsignalized intersection methodology. This involves calculating the probability of gaps in the opposing traffic stream, using standard average values provided in the HCM. The potential capacity of a Stop Sign (and therefore RTOR as well) is described with a Poisson-type probability equation (Roess et al. 2004).

While this model could produce results that fairly accurately predict delay at right-turn channels, the fact remains that this is not the procedure. When channels are added to an intersection, the procedure is to treat those vehicles as though they are not part of the intersection, and thus do not contribute to the delay at the intersection (Roess et al. 2004). In software models such as the Synchro HCM module and the Highway Capacity Software (HCS) channels are treated the same as standard turn lanes.

There seems to be a dearth of research into right-turn operations. Perhaps this is because left-turn or through movements are usually the critical phases for an intersection, or because varying RTOR laws exist in many jurisdictions nationwide.

The Texas Department of Transportation (TxDOT) performed a study on the merits and demerits of various types of right-turn lane designs (Fitzpatrick et al. 2006). The researchers in this study assert that free right-turn channels lower emissions and reduce delay. The main focus of the TxDOT study is to quantify turning speed patterns at numerous intersections throughout Texas. The turning speeds are subjected to analysis of variance, which shows that geometry plays a statistically significant role in turning speed, with channelized turns having a slightly higher turning speed. They acknowledge, "Vehicles are observed to frequently stop prior to entering the cross street even with an available dedicated lane, because drivers do not know they have a dedicated lane or its length." (Fitzpatrick et al. 2006, 56) While the unnecessary delay at the basis of this paper is acknowledged in the

TxDOT study, an attempt to quantify or explain the delay appears to be outside of the scope of the study. The study authors recommend controlling accesses coming out of the intersection to discourage weaving, under the assumption that downstream accesses are a major contributing factor.

The most voluminous topic area on right-turn traffic operations is the need for deceleration lanes of any kind. Hadi and Thakkar (2003) provide methodology for determining operationally if a deceleration lane is needed. Their methodology involves estimating the speed differential between through and right-turning vehicles to decide if removing slower turning traffic from the stream will significantly improve the safety of an approach. Similarly, Potts et al. (2007) show the benefits and costs of deceleration lanes can be reliably predicted using VISSIM micro simulations, thus providing a methodology for determining benefit/cost ratios for planning projects.

While some researchers have made pioneering efforts in the field of right-turn operations, none have focused on the actual delay of right-turning vehicles in channels. Nor have any studies focused on the underlying causes for this delay. As right turn channels in general are a common intersection treatment, and free right channels are used in particularly important situations, it is imperative to understand the apparent inconsistencies of motorist behavior related to turn lanes of this type.

## METHODOLOGY

The authors assembled a survey using the Qualtrics (Qualtrics 2008) statistical survey software, leased through Brigham Young University's site license. This survey was distributed via mass email to respondents across the United States on 24 August 2009.

The survey asked respondents to identify their age, sex, whether or not they have a driver's license, and in which state they earned their first driver's license. After these initial questions the respondents saw five flashcards, each with a picture of a different intersection treatment. The respondents selected how they are supposed to behave when making a right turn at a roundabout, a four-way stop, a typical right-turn lane, a right-turn channel with a yield sign, and a free right-turn channel. Respondents then select one of three or four choices, listed from "aggressive" to "passive" driving behavior as below:

A. Proceed: Proceed directly through without stopping or yielding
B. Yield: Yield to oncoming traffic
C. Stop: Stop completely, yield, and then proceed
D. Wait: Wait for the light to turn green before proceeding

The full list of questions and the figures used are available in Macfarlane (2009). A total of 957 partial responses were received, though only 817 surveys were fully completed. The responses to questions of particular significance are compared using a Chi-squared independence test (Weiss 2002).

## RESULTS

The survey participants' responses to different intersection types are compared in cross-tabulation analysis. Because this study primarily addresses drivers' perceived behavior at signalized intersections, the following comparisons are used:

- Free right-turn channel and yield right-turn channel
- Free right-turn channel and standard right-turn lane or pocket

Also, because a majority of survey participants were women, the answers of men and women to the free right-turn channel are analyzed to present potential bias.

### Free Right-turn and Yield Right-Turn Channels

Table 1 presents the cross-tabulation of responses for free right-turn channels and yield right-turn channels. The answers that are "correct" in terms of engineering expectation are depicted in bold. Only 154 participants out of the 821 total correctly described the proper behavior in both turn types. While 315 participants correctly identified "B. Yield" as the proper behavior at a yield right-turn channel, only 238 correctly identified "A. Proceed" as the proper behavior at a free right-turn channel. The most frequent responses for each question were both "C. Stop," which is an incorrect response. The marginal distribution of the answers allows the authors to predict the "expected" response rates, were the responses to the two questions entirely independent. The differences between the expected and actual responses are given in Table 2. If the sum of the squared differences is sufficiently high, then we can infer survey participants view the two intersection types as functionally similar. This hypothesis is partially supported by the positive numbers on the diagonal, but can be confirmed with a Chi-squared independence test.

**Table 1. Cross-Tabulation of Free Right-Turn and Yield Right-Turn Channel Responses**

| *Free Right-Turn Channel* | *Yield Right-Turn Channel* | | | | |
|---|---|---|---|---|---|
| | A. Proceed | **B. Yield** | C. Stop | D. Wait | Total |
| **A. Proceed** | 25 | 154 | 49 | 10 | 238 *29.1%* |
| B. Yield | 0 | 114 | 54 | 10 | 178 *21.7%* |
| C. Stop | 1 | 39 | 227 | 19 | 286 *34.9%* |
| D. Wait | 1 | 8 | 31 | 77 | 117 *14.3%* |
| Total | 27 *3.3%* | 315 *38.4%* | 361 *44.1%* | 116 *14.2%* | 819 |

**Table 2. Difference between Actual and Expected Responses for Free Right-Turn and Yield Right-Turn Channel Cross-Tabulation**

| *Free Right-Turn Channel* | *Yield Right-Turn Channel* | | | |
|---|---|---|---|---|
| | A. Proceed | **B. Yield** | C. Stop | D. Wait |
| **A. Proceed** | 17 | 63 | -56 | -24 |
| B. Yield | -6 | 45 | -24 | -15 |
| C. Stop | -8 | -71 | 101 | -21 |
| D. Wait | -3 | -37 | -21 | 60 |

The data meet all assumptions of the Chi-squared independence test. The parameters for the test are described as follows:

- $\alpha_{0.05} = 16.9$ for 9 degrees of freedom
- $h_0 : \chi^2 \leq \alpha$ Null hypothesis; the responses are not related
- $h_a : \chi^2 \geq \alpha$ Alternative hypothesis: there is a relationship between the responses

The Chi-squared value is 586.25, which is significant at a 5% significance level. We therefore reject the null hypothesis and infer there is a statistical relationship between the answers to the two questions.

While 63 more participants answered both questions correctly than would be expected from the marginal distributions, this only indicates that some motorists do distinguish between free right-turn channels and yield right-turn channels; the vast majority of motorists statistically do not.

## Standard Right-Turn Lane or Pocket vs. Free Right-Turn Channel

Table 3 presents the cross-tabulation of responses for free right-turn channels and standard turn lanes. The answers that are "correct" in terms of engineering expectation are depicted in bold. Only 188 participants out of the 820 total correctly described the proper behavior in both turn types. While 598 participants correctly identified "C. Stop" as the proper behavior at a standard turn lane, only 238 correctly identified "A. Proceed" as the proper behavior at a free right-turn channel.

**Table 3. Cross-Tabulation of Free Right-Turn Channel and Standard Right-Turn Lane or Pocket Responses**

| *Free Right-Turn Channel* | *Standard Right-Turn Lane or Pocket* | | | | |
|---|---|---|---|---|---|
| | A. Proceed | B. Yield | **C. Stop** | D. Wait | Total |
| **A. Proceed** | 4 | 4 | 188 | 43 | 239 *29.1%* |
| B. Yield | 1 | 13 | 140 | 24 | 178 *21.7%* |
| C. Stop | 1 | 4 | 235 | 46 | 286 *34.9%* |
| D. Wait | 0 | 3 | 35 | 79 | 117 *14.3%* |
| Total | 6 *0.7%* | 24 *2.9%* | 598 *72.9%* | 192 *23.4%* | 820 |

**Table 4. Difference between Actual and Expected Responses for Free Right-Turn Channel and Standard Turn Lane or Pocket Cross-Tabulation**

| *Free Right-turn Channel* | *Right-turn Lane* | | | |
|---|---|---|---|---|
| | A. Proceed | B. Yield | **C. Stop** | D. Wait |
| **A. Proceed** | 2 | -3 | 14 | -13 |
| B. Yield | 0 | 8 | 10 | -18 |
| C. Stop | -1 | -4 | 27 | -21 |
| D. Wait | -1 | 0 | -50 | 52 |

Some of the 192 participants who answered "D. Wait" for behavior at a standard turn lane may live in non-RTOR jurisdictions, though the authors did not control for this. In this case, the vast majority of survey participants correctly selected "C. Stop" as the behavior at a standard turn lane. Because standard turn lanes are the most common right-turn type, it is reasonable that such a high percentage of participants would respond correctly.

The marginal distribution of the answers allows the authors to predict the "expected" response rates, were the responses to the two questions entirely independent. The differences between the expected and actual responses are given in Table 4. If the sum of the squared differences is sufficiently high, then we can infer survey participants view the two intersection types as functionally similar. This hypothesis is partially supported by the positive numbers on the diagonal of Table 4, but can be confirmed with a Chi-squared independence test.

The data meet the assumptions of the Chi-squared independence test after combining the first two columns. This is reasonable because either "A. Proceed" or "B. Yield" is an irresponsible action at a standard turn lane or pocket. The parameters for the test are described as follows:

- $\alpha_{0.05} = 12.6$ for 6 degrees of freedom
- $h_0 : \chi^2 \leq \alpha$ Null hypothesis; the responses are not related
- $h_a : \chi^2 \geq \alpha$ Alternative hypothesis: there is a relationship between the responses

The Chi-squared value is 161.1, which is significant at a 5% significance level. We therefore reject the null hypothesis and conclude there is a statistical relationship between the answers to the two questions, and that a significant majority of motorists will likely exhibit similar behavior at different intersection right-turn types.

## Responses by Gender

Respondents were also asked to identify their gender. Because twice as many women participated in the survey as men, the authors were concerned that a participant's gender could be a lurking variable. A tabulation of answers by gender is given in Table 5.

As before, the marginal distribution of the responses allows the authors to compute expected values, were the responses not dependent on the gender of the participant. The differences between the expected and actual responses are given in

Table 6. If the sum of the squared differences is sufficiently large, there is a statistical relationship between the responses to the two questions.

The parameters for the test are described as follows:

- $\alpha_{0.05} = 7.8$ for 3 degrees of freedom
- $h_0 : \chi^2 \leq \alpha$ Null hypothesis; the responses are not related to gender
- $h_a : \chi^2 \geq \alpha$ Alternative hypothesis: there is a relationship between the responses and the participant's gender

The Chi-squared value is 12.9, which is significant at a 5% significance level. We therefore reject the null hypothesis and conclude there is a statistical relationship between stated behavior and the gender of the participants. The data show this as well; the expected values clearly indicate that there is a statistically significant difference between the answers of men and women.

**Table 5. Cross-Tabulation of Free Right-Turn Channel by Gender**

| *Free Right-turn Channel* | Men | Women | Total |
|---|---|---|---|
| **A. Proceed** | 95 | 143 | 238 *29.1%* |
| B. Yield | 51 | 127 | 180 *21.7%* |
| C. Stop | 77 | 207 | 286 *34.9%* |
| D. Wait | 30 | 87 | 117 *14.3%* |
| Total | 253<br>31.0% | 564<br>69.0% | 817 |

**Table 6. Difference between Actual and Expected Responses for Free Right-Turn Channel by Gender**

| *Free Right-turn Channel* | Men | Women |
|---|---|---|
| **A. Proceed** | 21 | -21 |
| B. Yield | -4 | 4 |
| C. Stop | -11 | 11 |
| D. Wait | -6 | 6 |

## DISCUSSION OF RESULTS

The statistical significance of the Chi-squared independence tests show that respondents' answers to questions were dependent on their answers to other questions; that is, people generally tended to give identical or similar answers to different questions. Practically, this means people view different right-turn treatments as being similar in operation.

The difference between expected and actual response rates shows some interesting trends as well. The positive trend along the diagonal of both tables reinforces the validity of the Chi-squared results. The only other positive values on either table are more "aggressive" answers to the free right-turn channel question.

This means that people who answered differently for the two questions tended to answer more correctly than incorrectly.

The 73% of "correct" answers for the right-turn lane question indicate that almost all people understand the proper behavior at this more prevalent intersection treatment, with drastically lower percentages for more atypical treatments.

### Limitations

The responses were collected as a voluntary response email, not as a random sample. Approximately 30,000 emails were sent out to addresses that the Qualtrics Corporation had on file, and 957 responses were collected; there may be lurking variables in the data related to the type of people willing to respond to this kind of survey. No demographics of the respondents are known, other than the answers to the survey questions. More resources would have permitted a simple random sample; future studies could attempt a more statistically proper sample.

The authors will make the survey data readily available to any future researchers who wish to continue this aspect of the study.

## CONCLUSION AND RECOMMENDATIONS

The statistical models indicate that in general, motorists tend to treat all intersections the same. That is, a driver is likely to behave at a free right-turn channel exactly as he would at a typical right-turn lane. However, it should also be emphasized that when motorists recognize operational distinctions in intersection treatments, they generally respond more correctly than incorrectly. This observations leads to the possibility that improved signage and education can overcome motorist error.

The survey results show that the general driving population does not distinguish between different intersection right-turn treatments. Some motorists do understand the subtle distinctions between channel types, but this proportion is not large enough to justify the "zero-delay" assumption for free right-turn channels. The public should be educated to improve the effectiveness of right-turn channels.

### Limitations and Opportunities

A simple random sample should be used to increase the statistical validity of the survey findings. Nevertheless, there is a potential wealth of information in the survey data that this study does not have the resources to address. Future researchers should consider the role that driver age, experience, or familiarity with a given intersection has on perceived turn behavior.

### Recommendations

The results indicate that there is a gap between engineering practice and public perceptions. Engineers should try to inform the public of proper driving behavior through education and clear signage.

A potential mitigation is to include the "Yield to Pedestrians" (R1-5 or R1-6) sign, at the crosswalk and a "Right Turns Do Not Stop" regulatory message

approaching the channel (FHWA 2003). These signs, if used consistently, could help reassure motorists that there is in fact an auxiliary lane at the exit of the channel, and they could help to teach proper driving behavior. Typically engineers are reticent to ask the public to drive more aggressively, but in this case unnecessary delay could be safely reduced with a change in driving practice.

## REFERENCES

Federal Highway Administration (FHWA) (2003). "Chapter 2- Regulatory Signs" *Manual on Uniform Traffic Control Devices,* U.S. Department of Transportation, Washington, D.C..

Fitzpatrick, K., Schneider, W.H., and Park, E.S. (2006). "Operation and safety of right-turn lane designs." *Transportation Research Record No. 1961*, National Research Council, Washington, D.C., 55-64.

Hadi, M.A. and Thakkar, J. (2003) "Speed differential as a measure to evaluate the need for right-turn deceleration lanes at unsignalized intersections." *Transportation Research Record No. 1847*. National Research Council, Washington, D.C., 58-65.

Husch, D. and Albeck, J. (2003). *Intersection capacity utilization*. Trafficware, Albany, Calif.

Macfarlane, G.S. (2009). "Delay patterns and perceptions at free right-turn channelized intersections." Undergraduate Honors Thesis. Brigham Young University.

Potts, I.B., Rignert, J.F., Harwood, D.W., and Bauer, K.M. (2007). "Operational and safety effects of right-turn deceleration lanes on urban and suburban arterials." *Transportation Research Record No. 2023*, 52-62.

Roess, R.P., Prassas, E.S., and McShane, W.R. (2004). "Analysis of signalized intersections." *Traffic engineering, 3rd ed.* Pearson Prentice Hall, Upper Saddle River, New Jersey, 599-606.

Qualtrics, Incorporated. (2008) "Qualtrics statistical survey software." <www.qualtrics.com> (July-September 2009).

Transportation Research Board (TRB) (2000). "Chapter 16 Signalized intersections" *Highway capacity manual.* National Research Council, Washington, D.C..

Weiss, N.A. (2002). "Chi-square procedures." *Introductory statistics*. Addison-Wesley, Boston, 653-675.

# Common or Variable Cycle Length Policy for a more Efficient Network Performance?

Ali Hajbabaie
PhD Candidate, Civil and Environmental Engineering Department, University of Illinois at Urbana – Champaign, Ahajbab2@Illinois.edu

Rahim F. Benekohal
Professor, Civil and Environmental Engineering Department, University of Illinois at Urbana – Champaign, Rbenekoh@Illinois.edu

## Abstract

To coordinate signals along a corridor or in a network, the conventional wisdom of using a common cycle length for all coordinated intersections has been widely used and has led to very desirable results. On the other hand, using variable cycle length for this purpose may result in good signal coordination as well, however, the effectiveness and appropriateness of this method remains unknown. In this study, we compare the effects of using a common cycle length, to the effects of using a variable cycle length on signal coordination in a network of oversaturated intersections. For this comparison, we have used our microscopic-simulation-based Intelligent Dynamic Signal Timing Optimization Procedure (IDSTOP) that uses Genetic Algorithms to find near-optimal signal timing parameter in a network of oversaturated intersections. IDSTOP determines near-optimal signal timing parameters on a network of oversaturated intersections with a) fixed cycle length policy, and b) variable cycle policy. In both cases, we have made sure that the network was working at its near-optimal condition. In the first policy, IDSTOP determined a common cycle of 144 seconds and offsets and green splits so that in each corridor the through movements were coordinated. For the common cycle policy, the network processed 4854 vehicles that traveled 482 vehicle-miles, and on average each vehicle experienced 1.27 minutes of delay. The analysis of offsets showed that through traffic movements were coordinated in each corridor of the network. In the second policy, IDSTOP determined cycle length that ranged from 136 to 156 seconds. This shows that the cycle lengths were not similar but the range of that was narrow. The network processed 5298 vehicles, that travelled 501 miles, and the delay was 1.23 minutes. With the variable cycle policy the network processed 8.4% vehicles more than the first policy, vehicle miles travelled were significantly increased, and the average delay per vehicle did not change significantly. The analysis of the offsets showed that through traffic movements were coordinated when needed. These figures show a significant improvement over the traditional common cycle policy. The finding, as an example of other conditions that were analyzed, indicated that using variable signal timing policy results in a considerably higher number of vehicles processed by the network as well as a significantly lower delay in the network compared to a common cycle length.

## Introduction

Traffic congestion in major US metropolitan areas costs $87 billion dollars annually [1]. This substantial cost plus other negative effects of traffic congestion, calls for practical methods of managing congestion in transportation networks. Transportation supply management has potential to effectively reduce congestion in a transportation network by determining signal timing parameters that provides maximum capacity.

Traffic signal coordination, when done properly, improves intersection traffic operation and safety. In traffic signal coordination, it is tried to synchronize two or more closely located traffic lights such that vehicles passing through them experience a minimum number of stops. This means that for two coordinated intersections, vehicles released from upstream intersection during the green signal of the coordinated phase, do not stop for a red light at the downstream intersection in most of the times. Most of the researchers believe that signal coordination is only possible when the cycle length of the coordinated signals are the same, or one is *"k"* times longer than the other (*k* is an integer number), and in fact, they have achieved very desirable results. If having coordinated movements in all cycles of the coordinated intersections is desired, the cycle length of coordinated intersections has to be equal.

However, one may raise this question that does having similar cycle lengths result in better network performance in terms of number of vehicles discharged, vehicle-mile travelled, average delay, and number of stops. In fact, it may be possible that using approximately equal cycle length instead of exactly equal cycle lengths in a transportation network results in a better network performance. We have designed this study to answer that question. In this study, for a case study network we find optimal signal timing in two different cases when a) the cycle lengths of different intersections are exactly equal, and b) when they are different. Then using four network performance indices, we compare the performance of the network under different signal timing parameters.

To find the near-optimal signal timing parameters we have used our microscopic-simulation-based Intelligent Dynamic Signal Timing Optimization Procedure (IDSTOP) that uses Genetic Algorithms to find near-optimal signal timing parameter in a transportation network with: a) common cycle policy, and b) variable cycle policy. In both cases, since IDSTOP found the near-optimal signal timing parameters, we were confident that the network was working in its near-optimal condition. The case study network had nine oversaturated intersections that were 500 feet apart with two-way streets and protected left turn phase.

The rest of this paper presents the methodology used to compare the performance of the network for fixed signal timing and variable signal timing, some information on formulation of IDSTOP and how it works, the case study network, findings, and finally the concluding remarks.

## Methodology

As mentioned before, using variable cycle length may result in a more efficient network performance than the common cycle length. To determine which policy results in a more efficient network performance, we use IDSTOP to determine near-optimal signal timing for each policy (i.e. one time for common cycle length policy and one time for variable cycle length policy) on a single

oversaturated transportation network when all geometric, roadway, and traffic conditions are remained the same. The IDSTOP and this transportation network will be explained later.

Since IDSTOP finds the near-optimal signal timing in the network for each policy, we are sure that the performance of the network is very close to its optimal condition. As a result, we will compare the effects of the best common signal timing policy to the effects of the best variable signal timing policy on the same network. To determine which cycle plan policy (common or variable) resulted in better signal coordination, four different Network Performance Indices (NPI) are used. These NPI are the total throughput, total vehicle-mile travelled in the network, average delay per vehicle, and total number of stops in the network.

Total throughput is sum of the number of vehicles discharged from each approach in each intersection of the network. As a result a larger value of throughput indicates a larger number of vehicles that are processed by the network and consequently a better network performance. Total vehicle-mile is the total mileage that is travelled in the network by all vehicles. As a result, a larger value for vehicle-mile shows that vehicles can move more in the network. Thus, network performance is better. A shorter average delay in the network indicates that vehicles have to stop shorter for the green signals that indicates a better network performance. Finally a less number of stops shows that vehicles have to stop less at the intersections for the green signal. This shows a better network performance and a better signal coordination in the network.

After finding the near-optimal signal timing parameters for the common and variable cycle length scenarios, the values of all above mentioned NPI are determined by microscopic simulation. Since the value of each NPI depends on a sequence of random numbers, a single simulation run does not result in reliable comparison. As a result based on the accepted error in estimation of each NPI, the total number of simulation runs is determined using Eq. 8:

$$\#\ of\ Simulation\ runs = \frac{S^2 z^2}{\varepsilon^2}$$

Where:

$S$: standard deviation of each NPI,

$z$: critical value of normal distribution (or student t distribution if N is less than 30) for a certain confidence level, and

$\varepsilon$: the acceptable error.

For each NPI, required number of runs is determined and the maximum of them is chosen as the number of runs. It should be noted that this process is an iterative process. In this study we limited the acceptable error in estimation of each NPI to 1% that resulted in 250 simulation runs. In the next sections IDSTOP formulation and the case study network will be described.

## IDSTOP Formulation

Several studies have addressed signal timing optimization in one intersection [2], [3], [4], two intersections [2], [4], [5], or three intersections [6]. Other studies found optimal signal timing for arterials without left turns [7] and with left turns [8]. Some studies determined optimal signal timing in a network of oversaturated intersections but, did not take left turns into account [9], [10]. Sun and Benekohal [11]

studied oversaturated networks of one-way streets that left turns were allowed, but that is obviously not applicable to a network of two-way streets. Finding signal timing parameters in oversaturated condition is an optimization problem. Previous studies have used delay minimization, system throughput maximization, or queue length minimization for this purpose. Minimizing delay could work effectively in undersaturated condition, however, in oversaturation, queues start to develop and eventually will reach the upstream intersections and will block them. This results in a significant drop in the capacity of the network and consequently a significant increase in delay. On the other hand in oversaturated condition, maximizing the number of vehicles that are processed by each intersection or the whole network could be more effective than minimizing delay. In addition, the queue lengths should be minimized in the network to ensure preventing queue spill over and upstream intersection blockage. Girianna and Benekohal [10] suggested maximizing system throughput as the objective function and penalized it by the product of the queue length in each link and the length of that link. In this study we use a quite similar objective function. It maximizes the sum of number of vehicles released by each link of the network at each time interval and penalizes it by the summation of number of vehicles in the queue in each link at each time interval. The objective function could be formulated as follows:

$$Max \quad \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{\varphi=2}^{\phi_{ti}} n_{i\varphi}^{t} - \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{\varphi=2}^{\phi_{ti}} \delta_{i\varphi}^{t}.\, q_{i\varphi}^{t} \qquad (1)$$

where:

*T*: number of study periods,
*N*: total number of intersections,
$\phi_{ti}$: number of phases at intersection *i*, at time period *t*,
$n_{i\varphi}^{t}$: total number of vehicles processed by intersection *i*, at time period *t*, in phase $\varphi$,
$q_{i\varphi}^{t}$: queue length at intersection i, at time period *t*, waiting to be served by phase $\varphi$, and
$\delta_{i\varphi}^{t}$: penalty for queue length at intersection *i*, at time period *t*, waiting to be served by phase $\varphi$.

This optimization problem has several constraints. The first one limits the maximum queue length in each link at each time interval to a certain number to avoid spill over. This number could be the length of each link or any number less than that. Equation (2) represents this constraint.

$$q_{i\varphi}^{t} \le Q_{i\varphi}^{t} \qquad (2)$$

Where:

$Q_{i\varphi}^{t}$: is the maximum allowed length of the queue at intersection *i*, at time period *t*, waiting to be served by phase $\varphi$.

The cycle length, green times for each phase, and the offset, should fall between minimum and maximum allowed values. Equations 3, 4, and 5 represent these constraints.

$$Cmin_{i}^{t} \le C_{i}^{t} \le Cmax_{i}^{t} \qquad (3)$$

$$gmin_{i\varphi}^{t} \le g_{i\varphi}^{t} \le gmax_{i\varphi}^{t} \qquad (4)$$

$$0 \le off_{i\varphi}^{t} \le C_{i}^{t} \qquad (5)$$

Where:

$C_i^t$: cycle length at intersection *i*, at time period *t*,
$Cmin_i^t$, $Cmax_i^t$: minimum and maximum cycle length at intersection *i*, at time period *t*,
$g_{i\varphi}^t$: green time for phase $\varphi$, at intersection *i*, at time period *t*,
$gmin_{i\varphi}^t$ and $gmax_{i\varphi}^t$: min and max of green time for phase $\varphi$, at intersection *i*, at time period *t*, and
$off_{i\varphi}^t$: offset of phase $\varphi$ (to a known time origin), at intersection *i*, at time period *t*.

The number of vehicles released from, as well as the queue length in each link of the network at each time interval depends on the signal timing in the network and are determined using a microscopic simulation approach. These relationships are formulated as follows:

$$n_{i\varphi}^t = f_1\left(off_{i\varphi}^t, g_{i\varphi}^t\right) \quad (6)$$

$$q_{i\varphi}^t = f_2\left(off_{i\varphi}^t, g_{i\varphi}^t\right) \quad (7)$$

Where:
$f_1$and $f_2$ indicate functions.

In solving this optimization problem, traditional methods such as exhaustive search, dynamic programming, approximate dynamic programming, and lagrangian relaxation may not result in optimal solution in a reasonable amount of time. As a result we decided to use a meta-heuristic approach to solve the problem and chose Genetic Algorithms. GAs are search techniques to find exact or approximate solutions to an optimization or a search problem. They are global search heuristics and are known to be less likely trapped in a local optimum. In addition, several techniques (such as niches) exist that reduce the chance of being trapped in a local optimum.

IDSTOP uses GA to solve the problem. The initial population is generated randomly to prevent any bias. To determine the fitness value of each individual in the population, CORSIM, a widely used microscopic traffic simulation package, is called. Based on its output, the value of objective function (Eq. 1) is determined which is the fitness value of that individual. GA directly takes constraints (3), (4), and (5) into account but, in case of constraint (2), the objective function is penalized when this constraint is violated. After determining the fitness value of the first population, using tournament selection two parents are selected, then crossed over, and mutated. This process of selection, crossover, and mutation is continued until enough individuals are created to form the new population. The fitness value of all members of the new population is determined and the next population is generated by using GA operators. This procedure is continued until the termination criteria are met.

For the variable-cycle plan case, IDSTOP finds the near-optimal signal timing parameters. In this case, the cycle length varies from each intersection to another and from each time interval to the other one. For the fixed-cycle plan case, the following constraint is added to IDSTOP. This constraint indicates that the cycle lengths of all intersections have to be exactly the same.

$$C_i^t = C_j^t \qquad for\ i,j = 1,2,\dots,N \qquad (8)$$

$$for\ t = 1,2,\dots,T$$

Where:

$C_i^t$: cycle length at intersection *i*, in time interval *t*, and *N* and *T* were introduced previously.

After adding this constraint, IDSTOP finds the optimal signal timing parameters for the same network under the same roadway and traffic condition.

## Case Study Network

Case study is a three by three rectangular network of nine intersections that are 500 feet apart. This network may be considered a small, or unrealistic network however, it shows the differences between the two policies. In addition, this network is very suitable for signal coordination due to its symmetric geometry, closely located intersections, and symmetric traffic demand.

Each street has a single through lane, a 200-feet long left turn pocket and a 150-feet long right turn pocket. Vehicles enter the network from 12 entry points that are shown by circles in Figure 1. A study period of 15 minutes is assumed where for a common-cycle plan, the cycle lengths at all intersections are the same while, for a variable-cycle plan they are not.

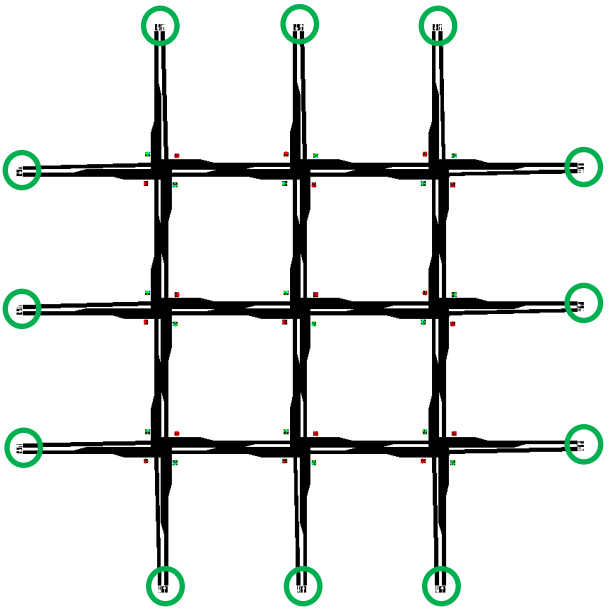


Fig. 1. Schematic view of the case study network. Entry points are marked with circles.

For both cycle plans we assumed that traffic enters the network from all twelve entry points with the rate of 1000 vehicles per hour that is just enough to make the network congested. These entry points are shown with green circles in Figure 1. We further assumed that at each intersection, 10% of vehicles turn left, 80% move through, and the remaining 10% turn right.

A four-phase signal plan is used with dual left turns signal, through signal and similar signals for the other directions as shown in Figure 2. Three seconds yellow signals and 1 second all-red signal is also assumed. Left turn signal may be omitted by the algorithm. If it is used, its minimum is 7 seconds and its maximum is 20 seconds. For through signals the minimum value of green signal is 15 and its maximum is 80 seconds. We used this wide range to ensure that even when the range is wide the algorithm is capable of finding near optimal values of signal timing parameters.

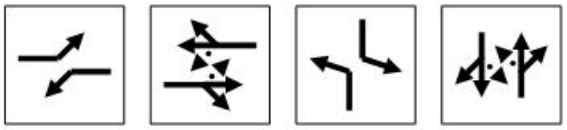


Fig. 2. Phase plan.

## Findings

In this section the results of our study will be presented. For a fixed-cycle plan, IDSTOP found a common cycle of 144 seconds. The offsets were almost zero meaning that at a certain time, all signals along a coordinated route became green. This happened for all coordinated routes of the network. In the network all through traffics were coordinated. This small value for offset was reasonable since the distance between the signals were only 500 ft.

For the variable cycle scenario, the cycle length ranged from 136 seconds to 156 seconds that showed 20 seconds difference between the minimum and the maximum of cycle length. Offsets where chosen so that for some part of the study period signals were coordinated along a path. Then the signals operated independently over that path while some other signals were coordinated.

As mentioned before, the network performance indices were determined by averaging their values over 250 simulation runs. For the common cycle scenario, average traffic throughput was 4854 vehicles. Using variable cycle length resulted in a throughput of 5298 vehicles that indicated 8.4% increase in the system throughput. This finding showed that using variable cycle length in the case study network could significantly increase the number of vehicles processed by each intersection of the network. We ran a student t-test to verify if the difference between the number of vehicles processed at each intersection for these two scenarios were significantly different or not, and the result of the test showed that system throughput for the variable cycle length scenario was significantly larger than that for the common cycle scenario.

Since the system throughput was larger for the variable cycle length scenario, we expected to observe the same result for total vehicle-mile travelled in the work zone. For the common cycle scenario, the vehicles travelled a total of 482 miles in the network. For the variable cycle scenario, the total vehicle mile travelled in the network was 501 miles. This shows that when a common cycle for all intersections of our case study was not used, vehicles travelled more in the network (with the same traffic assignment). This indicated that the network was less congested and consequently operated better.

For the common cycle strategy, the average delay per vehicle in the network was 1.27 minutes per vehicle. Using a variable cycle length strategy resulted in 1.23 minutes per vehicle average delay. This indicated that the average delay was reduced by 3.2%. This reduction in the average delay was not statistically significant. However, it showed that the performance of the network had improved since with the same amount of delay per vehicle, the case study network processed around 24.5% more vehicles.

For the common cycle strategy, on average each vehicle stopped 0.61 times in the network. Please notice that each vehicle traveled through three intersections on average. An average stop of 0.61

means that a vehicle did not stop at two intersections and 61% of the times it stopped at one of the intersections.

For variable cycle length, the average number of stops in the network was 0.63 times. This indicated that vehicles did not have to stop at two of the intersections at the network (in fact, they did not stop) and 34% of times they had to stop at one of the intersections. As a result, using a variable cycle plan resulted in slightly higher number of stops in the network compared to a common cycle plan. We ran proportion test and did not find this difference between the numbers of stops significant at a 95% confidence level.

## Conclusion

The findings of this study showed that on a small case study network that was suitable for signal coordination due to its symmetric geometry and traffic condition, using variable cycle lengths resulted in a better network performance than using a common cycle length. In the case study with a specific traffic demand, using variable cycle length increased the number of vehicles processed by the network by 8.4%. It significantly increased the total distance travelled in the network. A larger number of vehicles were processed by the network while average delay per vehicle stayed the same when variable cycle length was used. In addition, number of stops in the network slightly increased for the variable cycle strategy.

In summary, the findings of this study shows that in a specific network with a specific traffic demand, using variable cycle strategy can improve the performance of the network. To make sure if this is the case for different networks and traffic conditions, more analysis with different networks and different traffic demand is needed. A study with different traffic demand and different transportation networks is underway.

## Reference

[1] Urban Mobility Report, University Transportation Center for Mobility, Texas Transportation Institute, 2009.
[2] D. C. Gazis, "Optimum Control of a System of Oversaturated Intersections," *Operation Research*, vol. 12, pp. 815–831, 1964.
[3] D. C. Gazis, and R.B. Potts "the Oversaturated Intersection," *Proceedings2nd international symposium on the theory of road traffic flow* , London, pp. 221, 1965.
[4] P.G. Michalopoulos, and G. Stephanopoulos, "Optimal Control of oversaturated intersections: theoretical and practical considerations ," *Traffic Engineering and Control*, vol. 19, pp. 216–222, 1978
[5] C. Yuan, X. Yang, and F. Shen, "Fixed Cycle Strategy in Oversaturated Network Traffic Control," *Proceedings of the 6th World Congress on Intelligent Control and Automation,* pp. 8674–8678, 2006.
[6] H. L. Lo, and H.F. Chow, "Control Strategies for Oversaturated Traffic," *Journal of Transportation Engineering ASCE,* pp. 466–478, 2004.
[7] G. Abu-Lebdeh "Development of Dynamic Traffic Signal Control Procedures for Oversaturated Arterials and Genetic Algorithms Solutions". Ph.D. Dissertation, University of Illinois at Urbana Champaign, 1999.

[8] L. Zhang, Y. Yin, Y. Lou, "Robust Signal Timing for Arterials under Day-to-Day Demand Variations", TRB 89 DVD, 2010.
[9] H. Nagase, "Modelling and Optimal Control of Oversaturated Transportation Networks," *Applied Math Modelling,* pp. 101–108, 1980.
[10] M. Girianna, R. F. Benekohal, "Using Genetic Algorithms to Design Signal Coordination for Oversaturated Networks," *Intelligent Transportation Systems,* pp. 117–129, 2004.
[11] D. Sun, R. F. Benekohal "Multi-objective Traffic Signal Timing Optimization Using non-dominated Sorting Genetic Algorithm," *IEEE-20-7803-7848,* pp. 198–203, 2003.

# Evaluation of Wireless Magnetometers for Vehicle Detection at Signalized Intersections

Juan C. Medina[1], Ali Hajbabaie[2], Rahim F. Benekohal[3]

[1]PhD Candidate, Department of Civil and Environmental Engineering. University of Illinois. Urbana, IL. 61801; PH (217) 333-5967; email: jcmedina@illinois.edu
[2]PhD Candidate, Department of Civil and Environmental Engineering. University of Illinois. Urbana, IL. 61801; PH (217) 333-8988; email: ahajbab2@illinois.edu
[3]Professor, Department of Civil and Environmental Engineering. University of Illinois. Urbana, IL. 61801; PH (217) 244-6288; email: rbenekoh@illinois.edu

## ABSTRACT

Wireless magnetometers for vehicle detection were evaluated at a signalized intersection at both stop bar and advance locations. The studied approach had three lanes, and for each lane one sensor was installed at the stop bar and one more at the advance zones. Loop detectors were installed at the same location of the magnetometers and were used as a pointer to identify potential detection errors. Initially, data was collected after the system was installed by the manufacturer. Then, the system setup was adjusted by the manufacturer based on the analysis of the initial data, and data was collected again. Results show that the most frequent error at the stop bar was false calls (5.6%-7.6% due to vehicles in adjacent lanes, and additional 7.8%-9.6% due to single vehicles placing multiple calls), and at the advanced zones it was missed calls (0.9% to 10% per zone). Stuck-on calls and dropped calls were very rare.

## INTRODUCTION

New vehicle detection technologies have been increasingly used in the last decade in the U.S. These new systems include wireless magnetometers, which may provide more flexibility, ease of installation, and also less maintenance requirements compared to the widely used inductive loops. A typical installation of wireless magnetometers would consist of: wireless sensors; a contact closure card that allows the system to interface with the traffic controller; and an access point that provides wireless communication with the sensors and wired connection with the contact closure card. Thus, the access point serves as the 2-way-communication link between the cabinet and the sensors. The wireless sensors are embedded inside the pavement by drilling a hole of approximately 4 inches in diameter and 2 ¼ inches in depth. In addition, if the sensors are located out of the maximum communication range of the access point, wireless repeaters may create a bridge between the access point and the sensors to extend the usable range of the system. Repeaters allow the sensors to be as far it is needed, so that at a signalized intersection both stop bar and advance detection zones can be achieved.

Previous studies have evaluated some aspects of the performance of wireless magnetometers for vehicle detection, mostly for freeway applications. In 2006, the California Center for Innovative Transportation evaluated the Sensys system on freeways, finding that more than 99% of the records sent by the sensors were received by the access point within 0.5 seconds without loss of data, with a detection (and traffic measurement) accuracy of 95%, and miscounts less than 2% (mainly due to missed motorcycles and double counting of heavy vehicles). Also in 2006, the Texas Transportation Institute (TTI) evaluated the accuracy of the Sensys system for a freeway application, and found average count errors of +/- 1% in 15-min intervals, comparable to loop detectors. A different report from Florida A&M and Florida State University described, for a freeway application, average error counts of +/- 1% (between -4.1% and +2.5% in 15min counts) and average speed errors of -1.4% (between -3% and +0% for 15min counts). A separate study by the University of California at Berkeley (2006) studied the accuracy of speed reported by the Sensys system at about 150m from an intersection, and reported speed errors between 0.3% and 6.9% and frequently failing to detect motorcycles.

In addition, a study from Day et al. (2010) studied the performance of Sensys magnetometers placed at left-turn pockets of an actuated signalized intersection. This setup used four sensors adjacent to each other (along the roadway direction) in the same lane, and they also installed inductive loops as a base to compare the magnetometers. Data from 240 hours was analyzed, finding similar performance of both systems in terms of missed calls, but observed that wireless magnetometer had a higher tendency to generate false calls.

In this paper, an evaluation of the performance of wireless magnetometers manufactured by Sensys Networks is presented for the case of a signalized intersection. Different from previous research, a single detector was installed in each of the lanes of a selected approach at locations close to the stop bar and also at about 250 ft upstream from the stop bar, typical of advance detection. The magnetometers were installed centered inside the loop detectors in order to obtain similar detection areas from both systems. This allowed the automation of the initial analysis of the data by using a computer algorithm that compared loop detectors and magnetometers and could detect discrepancies that were classified as potential detection errors.

The performance evaluation was based on four types of detection error: false calls, missed calls, stuck-on calls, and dropped calls. These measures of performance encompass the most common detection errors from vehicle detection systems, providing a clear picture of the advantages and disadvantages of using wireless magnetometers at each detection zone.

Two sets of data were analyzed: 1) after the wireless system was initially configured by the manufacturer, and 2) after the manufacturer made modifications to improve the initial system performance. The performance of the system from the "initial" setup and the "modified" setup are described in detail for each zone, indicating the main sources of errors.

The remaining of the paper is organized as follows. The next section describes test site and the system setup, followed by the methodology and the description of the data used in the performance evaluation. Then, the results and

analysis are presented, and the final section includes the conclusions and closing remarks.

## FIELD TEST SETUP

The eastbound approach of the signalized intersection of Century Blvd and Veterans Pkwy, in Rantoul, IL, was selected for the evaluation of the wireless magnetometers. This approach has two left-turn lanes and a shared right-thru lane. An inductive loop (6ft x 6ft) was installed in each lane at stop bar and advance locations (about 250ft upstream), providing a total of six detection zones. At the center of each loop, a Sensys wireless magnetometer was installed by a representative from the sensor manufacturer, who also set the system configuration using their proprietary software. Thus, three sensors were installed at the stop bar locations and three sensors at the advance locations. The sensors were embedded in the pavement centered inside the loops, so that the detection area of the two detection technologies was as similar as possible.

In addition to the sensors, the installation of the Sensys system also required the use of one access point and two repeaters. One of the repeaters was installed on the mast arm of the traffic lights facing the vehicles on the opposite direction (in the receiving lanes) and linked the communications between the three stop bar detectors and the access point. A second repeater did it so for the advance detectors and was installed at a light pole.

A video camera was used to record images of the eastbound approach for later visual confirmation of the detection errors, and for the identification of their potential causes.

It is noted that all hardware required for data collection was housed in a separate cabinet form that used to operate the intersection. Thus, the data collection setup did not interfere with the normal operation of the traffic signals.

A diagram showing the location of the detection zones is included in Figure 1.

## METHODOLOGY

The evaluation of the magnetometers was performed using two types of data: activation and deactivation times of loops and magnetometers (called “timestamps”), and video images. The timestamps allowed the automation of the initial stages of the analysis using computer algorithms and also the use of large datasets. Timestamps were collected using an input/output (I/O) device to monitor vehicle presence as identified by each inductive loop and wireless detector. The I/O device verified the state of the 6 loops and the 6 sensors once every 50 milliseconds. Also, video images were fed as an input to a quad processor, along with a real-time graphical depiction of the status of the loops and sensors generated by I/O device. This graph provided an additional tool to visually confirm if a call took place in any zone. The recorded video images were also used as the ground truth to determine if detection errors had occurred.

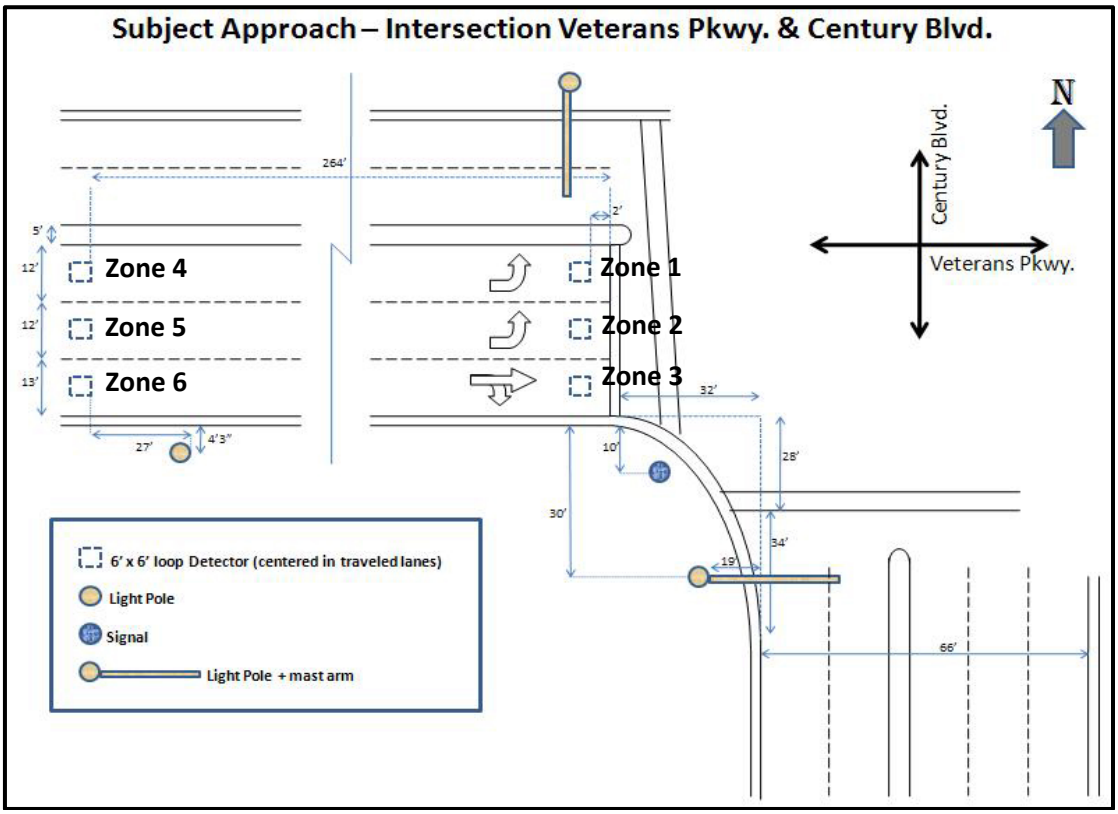



Figure 1. Diagram of signalized intersection and location of detection zones.

As mentioned above, four measures of performance were used to quantify the detection errors: false calls, missed calls, dropped calls, and stuck-on calls. These measures were estimated for each sensor separately by automatically detecting potential errors using the computer algorithms, and then by manually verifying them from the video images.

The computer code developed to accomplish the automated steps of the analysis read the timestamps from both loops and wireless sensors, establishing if there was a discrepancy between them. A time window of 2 seconds was used when comparing the activation and deactivation times of loops and sensors. Thus, a slightly different detection time was allowed, given the different characteristics of the two technologies. The four measures of performance were defined as follows:

- Missed Calls: Occurred when a wireless sensor failed to detect a vehicle. These errors could have adverse safety effects due to potential red light runners in cases where the corresponding phase was not called by the controller. In terms of timestamps, for every loop call if there was not a corresponding Sensys call, it was considered a potential missed call.

- False Calls: Are defined as calls placed by the wireless sensors when there was no vehicle present over the sensor or when multiple calls were created by the same vehicle. Thus, based on the cause of the false activations, false calls were divided into two subgroups: 1) multiple calls placed by the sensor for the same vehicle while it was occupying the detection area, and 2) calls caused by vehicles when they travelled in the adjacent lane (also called flickering false calls). False calls could have a negative effect in the operational efficiency of the intersection. In the

algorithm, for every call by a wireless sensor, if there was not a corresponding call from the loop detector, it was considered a potential false call.

- Dropped Calls: Occurred when a call by the wireless sensor was terminated while the vehicle was still present. If the sensor prematurely dropped the call placed to the controller, this may prevent the corresponding phase from being called, generating potential safety issues due to red light runners. In terms of timestamps, if the sensor call was terminated more than 5 seconds before the end of loop call, it was considered as a potential dropped call.

- Stuck-on Calls: Are defined as the calls that occurred when the wireless sensor indicated that the vehicle is still present, but the vehicle had departed. Stuck-on calls may affect operational efficiency of the signalized intersection. In the algorithm, if a sensor call continued to be active more than 10 seconds after the end of the loop call it was flagged as a potential stuck-on call.

## DATA DESCRIPTION

As mentioned above, the evaluation presented in this paper will be based on two sets of data: 1) data collected soon after the system was initially installed (called hereafter "initial setup"), and 2) data collected after the manufacturer made modifications to the system setup, based on the analysis of the "initial setup" (called hereafter "modified setup"). These modifications included an update to the firmware (in the sensors and the access points) and changing the operating mode of the stop bar sensors. A total of 26 hours of data was analyzed for each of the two system setups. Datasets were collected from several different days and at different times of day. It is noted that for both the initial and the modified setups datasets were selected from days with normal weather conditions (no rain, or snow) from the summer and fall season (preventing extreme temperatures).

The sample size for the initial setup was about 5700 vehicles for the three lanes together, and about 5065 for the modified setup. The estimated number of vehicles was obtained from loop detector counts at the stop bar zones, which at this location have resulted in counting errors lower than 1% (this information was obtained from separate datasets). The lane utilization was not even across the three lanes. The center lane is the most utilized, and it carried about 41% of the total number of vehicles, followed by the through lane (34%), and the left-most lane (25%).

## RESULTS AND ANALYSIS

The quantification of the detection errors and their description are presented in this section. Note that the zones are numbered following the numeration shown in Figure 1. Results from the initial setup and the modified setup will be shown simultaneously, for the reader to compare the differences between the two of them.

**Stop Bar Zones.** The detection error with highest frequency at the stop bar zones was false calls, both due to multiple calls caused by a single vehicle (flickering false calls) and due to vehicles in the adjacent lane. A general summary of the false calls for the three stop bar zones is shown in Table 1.

From Table 1, it is observed that false calls at the stop bar zones in the initial setup varied from 14.2% to 19.6%, being those from Zone 1 the highest. Most of the false calls were generated when a single vehicle placed multiple calls and only a small percentage of the errors were due to vehicles in the adjacent lanes.

Changes from the initial to the modified setup resulted in limited improvements for all three zones. False calls were reduced by at most 3% (in Zone 1) and ranged between 13.5% and 16.8%. However, the source of the errors changed significantly, and there was a higher incidence of vehicles in the adjacent lanes placing false calls. This, in turn, reduced the percentage of errors due to flickering false calls in the modified setup.

It is noted that both passenger cars and trucks placed false calls when traveling in the adjacent lanes. For example, out of the 17 false calls caused by adjacent vehicles in the initial setup in Zone 1, 5 of them were due to trucks and 12 due to passenger vehicles. Similarly, for Zone 1 in the modified setup, 34 false calls were placed by trucks and the remaining 77 were placed by passenger cars.

**Table 1. False Calls at the Stop Bar Detection Zones**

| Dataset | Detection Zone | Total Activations | | False Calls | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Magnetometers (a) | Loops (b) | Caused by Vehicles in Adjacent Lanes | | Multiple Calls From a Single Vehicle (Flickering False Calls) | | Total | |
| | | | | Frequency (c) | % (d)=(c)/(a) | Frequency (e) | % (f)=(e)/(a) | Frequency (g)=(c)/e) | % (h)=(g)/(a) |
| Initial setup | Zone 1 | 1796 | 1464 | 17 | 0.9% | 335 | 18.7% | 352 | 19.6% |
| | Zone 2 | 2649 | 2275 | 26 | 1.0% | 359 | 13.6% | 385 | 14.5% |
| | Zone 3 | 2268 | 1961 | 22 | 1.0% | 301 | 13.3% | 323 | 14.2% |
| Modified setup | Zone 1 | 1532 | 1311 | 111 | 7.2% | 147 | 9.6% | 258 | 16.8% |
| | Zone 2 | 2450 | 2108 | 185 | 7.6% | 191 | 7.8% | 376 | 15.3% |
| | Zone 3 | 1884 | 1646 | 105 | 5.6% | 150 | 8.0% | 255 | 13.5% |

Also, in some cases, false calls were observed even without any vehicle being near the detection zones. Posterior analyses of the video images showed no obvious reason for these calls being generated. The precise account of the types of vehicles causing false calls from the adjacent lanes at the stop bar zones is shown in Table 2.

In addition to the false calls, only a few occurrences were found for the other types of detection errors. They only represented a very small fraction of the total number of errors and are enumerated in Table 3. It is noted that all cases of missed calls at the stop bar for both the initial and modified setups were related to motorcycles, with a total of 5 occurrences. A similar situation was observed for the dropped calls, which were all due to motorcycles (3 occurrences). In addition to these errors, three stuck-on calls were found and each of them lasted between 25 seconds and up to 4 minutes.

**Advance Zones.** Detection errors in the advance zones were created by different situations than those described at the stop bar. At the advance zones, many of the approaching vehicles were in the process of changing lanes from the center lane to the left-most lane, increasing the chances for the vehicles to be missed. In fact, the great majority of the errors were the result of vehicles traveling between two of the lanes and being missed by both detection zones.

**Table 2. Detail False Calls Due to Vehicles in Adjacent Lanes**

| Dataset | Detection Zone | Total Activations | | False Calls Caused by Vehicles in Adjacent Lanes | | | |
|---|---|---|---|---|---|---|---|
| | | Magnetometers (a) | Loops (b) | Caused By Passenger Cars (c) | Caused By Trucks (d) | Other Causes (e) | Total (f)=(c)+(d)+(e) |
| Initial setup | Zone 1 | 1796 | 1464 | 12 | 5 | 0 | 17 |
| | Zone 2 | 2649 | 2275 | 14 | 12 | 0 | 26 |
| | Zone 3 | 2268 | 1961 | 12 | 10 | 0 | 22 |
| Modified setup | Zone 1 | 1532 | 1311 | 77* | 34 | 0 | 111 |
| | Zone 2 | 2450 | 2108 | 157 | 16 | 12** | 185 |
| | Zone 3 | 1884 | 1646 | 58 | 47 | 0 | 105 |

* Includes 4 vehicles traveling in the opposing direction (at other side of the median)

** Calls generated when no vehicles were at or near the detection zone

**Table 3. Missed, Stuck-on, and Dropped Calls at the Stop Bar Detection Zones**

| Dataset | Detection Zone | Total Activations | | Detection Errors | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Magnetometers (a) | Loops (b) | Missed Calls | | | | Stuck-on Calls | | Dropped Calls | |
| | | | | Passenger Car / Truck | | Motorcycle | | | | | |
| | | | | Frequency (c) | % (d)=(c)/(b) | Frequency (e) | % (f)=(e)/(b) | Frequency (g) | % (h)=(g)/(a) | Frequency (i) | % (j)=(i)/(a) |
| Initial setup | Zone 1 | 1796 | 1464 | 0 | 0.0% | 2 | 0.1% | 0 | 0.0% | 2* | 0.1% |
| | Zone 2 | 2649 | 2275 | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% |
| | Zone 3 | 2268 | 1961 | 0 | 0.0% | 1 | 0.1% | 2 | 0.1% | 0 | 0.0% |
| Modified setup | Zone 1 | 1532 | 1311 | 0 | 0.0% | 1 | 0.1% | 1 | 0.1% | 1* | 0.1% |
| | Zone 2 | 2450 | 2108 | 0 | 0.0% | 1 | 0.0% | 0 | 0.0% | 0 | 0.0% |
| | Zone 3 | 1884 | 1646 | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% |

* Calls dropped when motorcycles where stopped over the detection zone

This constitutes a serious detection error since advance detection zones are often set to provide dilemma zone protection. Missing a vehicle could potentially represent a safety concern if the corresponding phase is never called (increasing chances of red-light running) or if the signal controller does not take the approaching vehicle into account for terminating the green phase.

On the other hand, in many cases the two sensors closest to a single approaching vehicle traveling between lanes were activated, generating a "double call". It is noted that this event was not considered as a detection error since it was difficult to judge from the video images at what point the vehicle was over one of the detection zones or the other, or over both. Besides having a potential effect in the efficiency of the traffic signal, these events do not represent a major concern in terms of safety and were not considered for the performance evaluation.

A summary of the percentage of missed calls for each of the advance detection zones is shown in Table 4. It is observed that for the three advance zones

the frequency of missed vehicles was reduced from the initial to the modified setup. Also, for both setups Zone 4 had the highest percentage of missed vehicles, and Zone 6 the lowest. This is a direct consequence of vehicles changing lanes form the center to the left-most lane, as it was explained above. Most vehicles traveling in the through lane (Zone 6) were properly aligned inside the lane markings.

In Zone 6, however, note that most of the missed vehicles were traveling directly over the sensors, indicating a different type of detection error. Some cases of missed vehicles traveling directly over the sensors were also found in zones 4 and 5, which suggest a more general trend for all sensors. Thus, data shows that there is a probability for vehicles to be missed at the advance zones even though they are traveling over the sensor, and this probability is in the order of less than 1%.

Regarding other types of detection error, false calls were significantly less frequent than at stop bar zones, and for each of the zones they represented less than 2.5% of the total number of activations (see Table 5). A slight increase in the false calls was found in the modified setup compared to the initial setup, but the operational and safety implications of these changes are expected to be low.

**Table 4. Missed Calls at Advance Detection Zones**

| Dataset | Detection Zone | Total Activations | | Missed Calls | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Magnetometers (a) | Loops (b) | Vehicles Traveling Over Sensor | | Vehicles Traveling Between Two Lanes and Missed by Both Sensors | | Total | |
| | | | | Frequency (c) | % (d)=(c)/(b) | Frequency (e) | % (f)=(e)/(b) | Frequency (g)=(c)+(e) | % (h)=(g)/(b) |
| Initial setup | Zone 4 | 1158 | 1336 | 10 | 0.7% | 123 | 9.2% | 133 | 10.0% |
| | Zone 5 | 2438 | 2679 | 9 | 0.3% | 109 | 4.1% | 118 | 4.4% |
| | Zone 6 | 1928 | 2001 | 46 | 2.3% | 17 | 0.8% | 63 | 3.1% |
| Modified setup | Zone 4 | 1092 | 1189 | 3 | 0.3% | 65 | 5.5% | 68 | 5.7% |
| | Zone 5 | 2280 | 2471 | 8 | 0.3% | 71 | 2.9% | 79 | 3.2% |
| | Zone 6 | 1688 | 1697 | 10 | 0.6% | 6 | 0.4% | 16 | 0.9% |

**Table 5. False Calls at Advance Detection Zones**

| Dataset | Detection Zone | Total Activations | | False Calls | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Magnetometers (a) | Loops (b) | Caused by Vehicles in Adjacent Lanes | | Multiple Calls From a Single Vehicle (Flickering False Calls) | | Total | |
| | | | | Frequency (c) | % (d)=(c)/(a) | Frequency (e) | % (f)=(e)/(a) | Frequency (g)=(c)/e) | % (h)=(g)/(a) |
| Initial setup | Zone 4 | 1158 | 1336 | 7 | 0.6% | 6 | 0.5% | 13 | 1.1% |
| | Zone 5 | 2438 | 2679 | 2 | 0.1% | 23 | 0.9% | 25 | 1.0% |
| | Zone 6 | 1928 | 2001 | 5 | 0.3% | 9 | 0.5% | 14 | 0.7% |
| Modified setup | Zone 4 | 1092 | 1189 | 16 | 1.5% | 10 | 0.9% | 26 | 2.4% |
| | Zone 5 | 2280 | 2471 | 10 | 0.4% | 29 | 1.3% | 39 | 1.7% |
| | Zone 6 | 1688 | 1697 | 6 | 0.4% | 16 | 0.9% | 22 | 1.3% |

Lastly, there seems to be a tradeoff between the slight increase in the frequency of false calls and the reduction in the missed calls. More precisely, the

reduction in missed calls was more prominent than the increase in the false calls, which represents an overall improvement in the system.

## CONCLUSIONS

This paper presented an evaluation of the detection performance of wireless magnetometers at a signalized intersection. This study is different from previous research in several aspects: 1) the detection is based on a single magnetometer per detection zone, providing a detection area similar to that of an inductive loop; 2) a quantification of detection errors at both stop bar and advance zones is included and described in detail, identifying the main error types and their causes; and 3) datasets collected soon after the magnetometers were initially installed by the system manufacturer (called "initial setup"), are compared to those after the manufacturer adjusted the system parameters to improve the initial performance (called "modified setup").

In summary, false calls were the most common type of error at stop bar zones, where 5.6%-7.6% of the calls placed by the magnetometers were due to vehicles in adjacent lanes, and additional 7.8%-9.6% due to flickering false calls (a single vehicle placing multiple calls). There were only a few missed calls at the stop bar in both the initial and modified setups, and all of them occurred to motorcycles (5 times). The stuck-on and dropped calls were also rare and only three stuck-on calls and three dropped calls (to motorcycles) were found in the two setups combined.

At the advance zones, false calls were far less frequent than at the stop bar zones, and varied between 0.7% and 2.4% for the two setups. In contrast, missed calls were higher than false calls and ranged between 3.1% and 10.0% in the initial setup, and between 0.9% and 5.7% in the modified setup. It is noted that missed calls for all three advance zones decreased in the modified setup. Most missed calls were due to vehicles traveling between lanes, but also motorcycles and vehicles were missed while traveling directly over the sensors. No stuck-on or dropped calls were observed at advance zones.

Overall, the changes in the system setup made by the manufacturer improved the performance as they decreased false calls at the stop bar and missed calls at the advance zones. However, a significant number of errors were observed even after the modifications. It is noted that at the advance detection zones, the percentage of missed calls could be reduced if the vehicles traveled well aligned inside the lanes at the sensor location, thus relocating them may have positive effects. On the other hand, reducing the sensitivity to try to decrease false calls at the stop bar may result in increased missed calls, which may in turn increase safety concerns.

Future research to evaluate the precise operational and safety effects of detection errors by the magnetometers is recommended. Different lane configurations and traffic volumes may also have an important effect on the system performance and are recommended to be studied.

An evaluation of the effects of winter and rain conditions in the performance of the wireless magnetometers in underway. Also, the analysis of additional sensors located a few feet upstream of the stop bar zones is being conducted, trying to further improve the system performance by reducing flickering false calls.

## REFERENCES

ARRB Consulting/La Trobe University (2007). Evaluation of Sensys Networks Equipment.

California Center for Innovative Transportation (2006). Evaluation of Wireless Traffic Sensors by Sensys Networks, Inc. Final Report. Berkeley, CA.

Day C. M., H. Premachandra,. T. M. Brennan, J. R. Sturdevant, D. M. Bullock. (2010). "Operational Evaluation of Wireless Magnetometer Vehicle Detectors at a Signalized Intersection", TRB 89th Annual Meeting Compendium of Papers DVD, Washington DC

Florida A&M University/ Florida State University (2007). Accuracy Assessment of the Sensys Wireless Vehicle Detection System.

Texas Transportation Institute (2007). Count Accuracy Assessment of the Sensys Wireless Vehicle Detection System.

# Genetic Algorithm Methodology to Optimize Signal Timing For Vehicle and Pedestrian Delays

Zengyi Yang[1] and Rahim F. Benekohal[2]

## ABSTRACT

There are two objectives for this research. One is to develop an effective procedure to optimize intersection signal timing by minimizing total delay for both vehicles and pedestrians. The second objective is to establish guidance for the pedestrian crossing phase pattern (two-way or scramble) and for the length of WALK phase when a scramble crossing is used.

An optimization procedure for signal plans in an isolated intersection was developed. The procedure yielded up to four phases for vehicles with either the two-way or scramble pedestrian crossing phase. A simple Genetic Algorithm (GA) was used in finding suitable signal plans because of the existence of a very large solution searching space. Furthermore, a look-up table was generated to guide the decision between two-way and scramble phases. The guidance considered different combinations of vehicle and pedestrian volumes.

Compared with the Highway Capacity Software (HCS) GA function, the proposed procedure has the same accuracy and slightly better capability when pedestrian delay is not considered. However, when pedestrian delay is considered in signal timing, such a comparison could not be made because HCS does not compute delay for pedestrians while the proposed GA procedure does.

## INTRODUCTION

Traffic signals generally aim to minimize average vehicle delay, but pedestrian delay is not taken into account. Such a strategy is reasonable for rural areas or highways where very few pedestrians interfere with vehicular traffic. However, in a central business district where many pedestrians walk, the strategy that only optimizes vehicle flow would not suitable because pedestrian delay is ignored.

[1]Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 205 N Mathews Ave, Urbana, IL 61801; PH (217) 714-6659; e-mail: zengyi.yang@gmail.com

[2]Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 205 N Mathews Ave, Urbana, IL 61801; PH (217) 355-6006; e-mail: rbenekoh@illinois.edu

Ignoring pedestrian delay may not promote orderly use of crosswalks and may lead to more conflict between pedestrians and vehicles. Therefore, when pedestrian flow has a significant influence on vehicle flow, the signal plans should be optimized considering delay for both groups. Traffic signal plan optimization is to make a trade-off between vehicle and pedestrian delays by considering delay for all of the travelers.

Pedestrians may be allocated in an exclusive pedestrian phase (scramble) or allowed to move during a vehicular phase (two-way crossing). The scramble and two-way crossing patterns have advantages and disadvantages. The scramble crossing may cause longer waiting time for through traffic, turning traffic, and pedestrians, but it can provide additional safety benefits for pedestrians. In an intersection with high volumes of right-turn traffic and pedestrians, the scramble crossing can also eliminate the conflict between them and improve the capacity of the right-turn lane group (since right-turn vehicles would not block the intersection while waiting for pedestrians). Nevertheless, in an intersection with extremely heavy pedestrian flow, a major enforcement problem could occur because it would be difficult to clear pedestrians out of the intersection at the end of the scramble phase which may cause extra waiting time for vehicles. Thus, some traffic engineers discourage the usage of the pedestrian scramble except for unusually high pedestrian intersections. In contrast, the two-way crossing reduces delay for through traffic, turning traffic, and pedestrians compared to the scramble crossing, but it might not be beneficial in a high volume intersection especially with high volumes of right-turn traffic and pedestrians.

Limited studies have been undertaken to establish criteria to decide which crossing pattern is more suitable for a given condition. Currently, the two-way crossing pattern is the "rule" and the scramble crossing is the exception, but that is partially due to the lack of procedures to consider delay for both pedestrians and vehicles. For example, the Highway Capacity Manual (2000) neither discusses the scramble crossing nor considers pedestrian delay in signal timing. This paper provides a procedure to consider both pedestrian and vehicle delays and shows when the scramble or two-way crossing is more appropriate.

## LITERATURE REVIEW

Few studies were done to investigate the balance between pedestrian delay and vehicle delay either at an isolated intersection or in a network.

Noland (1996) analyzed the signal timing based on travel time costs of both pedestrians and vehicles at isolated intersections with high pedestrian demand. From an economic perspective, when the ratio of pedestrians to automobiles is high, he maintained that pedestrians should be favored by decreasing the automobile green phase length, increasing the pedestrian green phase length, or by alternative strategies such as reducing the major road width and closing the selected street to vehicular traffic at certain peak hours.

Bhattacharya, et al. (2005) proposed a deterministic model that incorporated both pedestrian and vehicle delay in a signal coordination plan. The authors analyzed the running results of the model on a hypothesized five-intersection arterial with various offsets and found that the best offsets for vehicles and pedestrians along the arterial were not necessarily the same. In order to minimize total pedestrian and vehicle user cost, an optimal signal coordination plan could be achieved by balancing pedestrian and vehicle delay.

Ishaque, et al. (2007) studied trade-offs between pedestrian and vehicular traffic in a hypothetical network by a VISSIM microsimulation model. Aiming at optimizing the average travel cost per person in all modes of the network, they found that shorter cycle lengths were beneficial for pedestrians. Signal plans that were advantageous to vehicles could be disadvantageous to pedestrians. Based on different proportions of pedestrians to vehicle users and different pedestrian time values, the suitability of three different pedestrian phase types was analyzed, so that the optimal network performance could be achieved for all road users. However, the signal plan in the study was limited to double-phase vehicle signal plans with single-exclusive, double-exclusive, or staggered pedestrian crossing phase(s).

Li, et al. (2009) developed a traffic signal optimization strategy, programmed in Matlab, for an individual intersection to minimize weighted total vehicle and pedestrian delays. The total vehicle and sidewalk pedestrian delays were calculated based on their deterministic queuing model respectively. The total crosswalk pedestrian delay was calculated based on an empirical pedestrian speed model, which considered interactions between pedestrian platoons. According to a case study at a Japanese intersection, the proposed model improved the average person delay by 10% without changing existing cycle lengths, and the further improvement could reach 44% with additional cycle length optimization.

So far, the research has rarely developed optimization for signal plans including pedestrian phases based on minimal pedestrian and vehicle delays. Therefore, it would be useful to explore a procedure that considers both delays and determines appropriate pedestrian crossing patterns.

## METHODOLOGY DEVELOPMENT

The objective function model of the signal optimization procedure is composed of two major parts: total vehicle user time and total pedestrian user time.

$$\text{Total User Time} = n_v \cdot (\text{Total Vehicle Delay}) + K \cdot (\text{Total Pedestrian Delay})$$

Where K = relative time value of a pedestrian compared with a passenger car;

$n_v$ = average vehicle occupancy per passenger car.

The total vehicle delay is acquired by summing up the products of the corresponding average delays and vehicle volumes. The average delay per passenger

car is calculated from the control delay model in Highway Capacity Manual (2000), which includes uniform delay, incremental delay, and initial queue delay.

The total pedestrian delay is acquired by summing up the products of the corresponding average delays and pedestrian volumes. The average pedestrian delay is calculated from the model proposed by Virkler (1998). The model was based on the assumption that all pedestrians arrived randomly, i.e. pedestrians who arrived at the green signal entered the intersection without any delay, and pedestrian flow arrived uniformly at the red signal. It was also assumed in the model that the cycle length was constant and no pedestrian actuation was applied in the intersection. Moreover, the model considered the compliance effect of pedestrians. In order to avoid waiting for the next WALK interval, certain pedestrians increased their waking speed and began their crossing without a Walk indication. The majority of these behaviors happened during the flashing DONT WALK phase. According to Virkler's research (1998), 69 percent of pedestrians arriving at the curb during these periods would enter crosswalks. Therefore, 0.69 was used as an adjustment factor of pedestrian effective green time.

The average vehicle occupancy 1.22 ($n_v$=1.22) was used in this study, on the basis of the traffic condition observation by Bhattacharya, et al. (2005).

According to the research of Ishaque, et al. (2007), the relative time value of a pedestrian compared with a passenger car (K) could range from 0 to 3 in most cases. Bhattacharya, et al. (2005) recommended K value of 2. Therefore, without describing the value of K, one unit of pedestrian delay was set as two units of vehicle delay (K=2) in a test of this paper.

Genetic Algorithm was chosen to solve this signal optimization problem, because it has excellent capability to solve problems with complex objective functions and large solution searching space (276,824,064). The searching space was calculated according to the basic variable setting. It was assumed that the opposite direction effective green time equaled each other and that the right turn effective green time equaled the green time of the through movement in the same approach. Therefore, four effective green time variables of vehicle lane groups 2, 5, 1, and 4 are basic variables for the optimization procedure. Based on a similar assumption, two WALK phase durations of pedestrian groups 1 and 2 are also basic variables. The notation of lane group and pedestrian group indexes is illustrated in Figure 1. In addition, the index of the pedestrian crossing pattern is also a basic variable.

The cycle length of the signal plan can be calculated based on the variables shown above and the parameters shown below. The yellow interval is 3 seconds for each phase. The red clearance interval with pedestrians moving in the previous phase is 2 seconds. Otherwise, it is 1 second. The flashing DONT WALK interval equals the crosswalk length divided by the average pedestrian speed (4 ft/s). If the total duration of the WALK phase and the flashing DONT WALK phase is smaller than the pedestrian minimum green time, Gp from HCM 2000, the WALK phase duration is

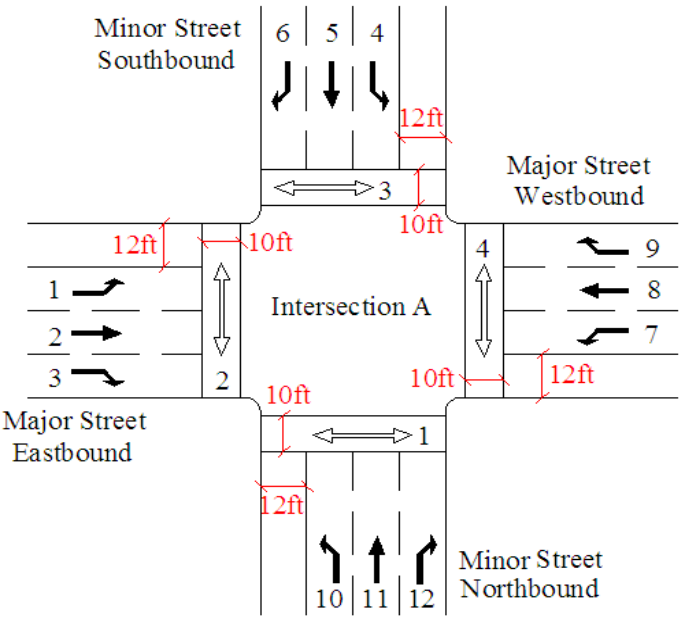



Figure 1. The Diagram of Intersection A

reset as the difference between the pedestrian minimum green time and the flashing DONT WALK interval.

Each signal plan is considered as an “individual” in the GA. The basic variables of the signal plan are encoded into a chromosome that represents the individual. The whole group of individuals is considered as a “population”. In general, the minimal total user time of the population gradually reduces throughout generations. According to the 500-generation test runs of the GA procedure, 50 generations are sufficient for this signal plan optimization (Yang, 2010).

The GA operators employed in this study are pairwise tournament selection without replacement, single point crossover, and mutation. The selection operator involves running “tournaments” among all the individuals, and the “winner” of each “tournament” (the individual with lower total user time) is selected for the crossover and mutation proceeding. Since the tournament selection is pairwise, the crossover probability is 50%. The crossover and mutation operators are applied to generate the population of the next generation from the individuals selected earlier in the “tournaments”. In order to ensure minimum diversity in the population and successful local search, the mutation probability is set as 3.23% and the population size is 31.

Based on all the definitions above, GA optimization procedure can be run after 12 vehicle volumes, 4 pedestrian volumes, and 12 initial vehicle queues are input into the procedure. The procedure result would be the signal plan proposed by the individual with the minimal total user time in the final generation population. The procedure determines whether each phase would be displayed and optimizes each phase length. However, the phase types and sequence of the basic signal plan (Figure

2) would not be changed.

According to the verification tests of the GA procedure (Yang, 2010), signal plans selected by the GA procedure in various volume combinations were reasonable. Phase lengths increased with the growth of the corresponding critical vehicle or pedestrian volumes. The pedestrian crossing type changed from a two-way crossing to a scramble crossing when the corresponding pedestrian and vehicle volumes grew considerably.

Furthermore, the total delay comparison analysis showed that the differences of the means from the GA procedure and the means from the HCS GA function were not significant at the significance level 0.05 (Yang, 2010). Minor differences occasionally occurred, probably because rounding up settings of HCS and the GA procedure are distinct from each other during calculation.

## GA EFFECTIVENESS OF THE PROPOSED PROCEDURE AND HCS

HCS Software has a GA optimization function for signal plan setting. Since HCS does not include any pedestrian delay calculation, the comparison only involves vehicles. In order to analyze the GA effectiveness of the proposed procedure, two tests were run respectively with Traffic Patterns A and B (Table 1) at Intersection A by both the proposed procedure and HCS GA.

In both tests, the comparison analysis results (Table 2) show that the proposed GA procedure performs relatively better than the HCS GA function either with or without the elitist method. The elitist method is to ensure that the signal plan with the minimal total user time in each generation would be reserved for the next generation, so that the most appropriate signal plan would not be lost throughout generations. When there is no pedestrian at the intersection, the signal plans recommended by the proposed GA procedure with an elitist method result in delay values that are lower than or at least equivalent to the delay values from the HCS GA optimized signal plans. Nevertheless, when pedestrian delay is considered in signal timing, such a comparison cannot be made because HCS does not consider pedestrian delay, while the proposed GA procedure considers both vehicles and pedestrians.

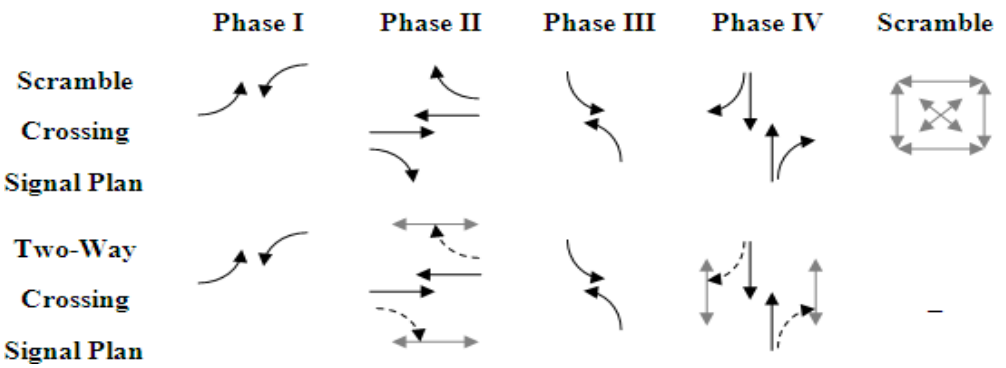



Figure 2. The Basic Signal Plan of Intersection A

Table 1. Vehicle Volumes, Initial Vehicle Queues, and Pedestrian Volumes Of Traffic Patterns A and B

| | Traffic Pattern | Vehicle and Pedestrian Movement Index | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Vehicle Volume (100 veh/h) | A | 2 | 2 | 2 | 2 | 4 | 2 | 2 | 4 | 2 | 2 | 4 | 2 |
| | B | 2 | 8 | 2 | 2 | 4 | 2 | 2 | 4 | 2 | 2 | 4 | 2 |
| Initial Vehicle Queue (veh) | A | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | B | 5 | 10 | 5 | 5 | 10 | 5 | 5 | 10 | 5 | 5 | 10 | 5 |
| Pedestrian Volume (100 ped/h) | A & B | 0 | 0 | 0 | 0 | - | - | - | - | - | - | - | - |

Table 2. The Signal Plan and Intersection Delay of Traffic Patterns A and B By the HCS GA Function and the Proposed GA Procedure

| Result Source | HCS GA | | | | GA Procedure | | | |
|---|---|---|---|---|---|---|---|---|
| With the Elitist Method? | No | | Yes | | No | | Yes | |
| Traffic Pattern | A | B | A | B | A | B | A | B |
| Cycle Length (s) | 97 | 174 | 107 | 155 | 95 | 162 | 113 | 166 |
| Phase I (s) | 14 | 28 | 15 | 19 | 14 | 21 | 17 | 21 |
| Phase II (s) | 24 | 62 | 28 | 57 | 24 | 62 | 29 | 64 |
| Phase III (s) | 16 | 20 | 15 | 20 | 13 | 21 | 17 | 21 |
| Phase IV (s) | 26 | 45 | 30 | 40 | 26 | 40 | 32 | 42 |
| Intersection Delay (s/veh) | 66 | 191 | 64 | 187 | 65 | 188 | 64 | 187 |

Moreover, additional tests were also run with the modified GA parameters to analyze the impact of changing the GA parameters on the optimization result in the HCS GA. In the extra two runs with Traffic Patterns A and B, population size, crossover probability, and mutation probability were modified as 31, 50%, and 3.2% to be compared with their default values of 10, 30%, and 4.0% in the earlier tests.

According to the test results (Table 3), the intersection delays of the results from the proposed procedure at Intersection A with Traffic Patterns A and B are 120% and 62% less respectively than those from the HCS GA with the modified GA parameters. In the runs with the modified parameters, although vehicles moving in the direction with higher traffic flow experience less delay, vehicles in the lower traffic direction generally suffer even longer delay, and longer cycle lengths and phase lengths tend to be favored.

Therefore, it is important to note that the critical GA parameters should remain the default setting in the HCS GA. Since the chromosome structure in the HCS GA function is unknown, it is hard to determine the suitable critical GA parameters to ensure both the evolution diversity and stability. Thus, the default parameters, which

are determined based on the HCS GA chromosome structure, lead to better performance than the modified parameters which do not satisfy the diversity and stability requirement. For the future development of HCS, the software will be easier to utilize if users are either not able to change the critical GA parameters or able to access the critical GA information.

## APPROPRIATE PEDESTRIAN CROSSING PHASE

The proposed GA procedure can assist a traffic engineer to optimize the pedestrian crossing phase at an intersection. If a guide table is generated, it will be easier and quicker for the engineer to reach suitable solutions under different traffic patterns.

In order to evaluate this application of the proposed procedure, a guide table was generated for Intersection A with a traffic pattern shown in Table 4. Eastbound right-turn vehicle volume (X) and eastbound pedestrian volume (Y) varied.

The guide table (Table 5) shows which pedestrian crossing pattern is more

Table 3. The Signal Plan and Intersection Delay with Traffic Patterns A and B By the HCS GA with Modified GA Parameters and the Proposed GA Procedure

| Traffic Pattern | A | | B | |
|---|---|---|---|---|
| Optimization Method | HCS GA with Modified GA Parameters | GA Procedure | HCS GA with Modified GA Parameters | GA Procedure |
| Cycle Length (s) | 114 | 95 | 150 | 162 |
| Phase I (s) | 32 | 14 | 26 | 21 |
| Phase II (s) | 15 | 24 | 75 | 62 |
| Phase III (s) | 19 | 13 | 11 | 21 |
| Phase IV (s) | 31 | 26 | 20 | 40 |
| Intersection Delay (s/veh) | 142 | 65 | 313 | 188 |

Table 4. Vehicle Volumes, Initial Vehicle Queues, and Pedestrian Volumes Of the Traffic Pattern for the Guide Table

| | Vehicle and Pedestrian Movement Index | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Vehicle Volume (100 veh/h) | 2 | 4 | X | 2 | 4 | 2 | 2 | 4 | 2 | 2 | 4 | 2 |
| Initial Vehicle Queue (veh) | 5 | 10 | 5 | 5 | 10 | 5 | 5 | 10 | 5 | 5 | 10 | 5 |
| Pedestrian Volume (100 ped/h) | Y | 2 | 2 | 2 | - | - | - | - | - | - | - | - |

Table 5. The Guide Table of Scramble WALK Phase Lengths (s) at Intersection A

| Y \ X | 700 | 800 | 900 | 1000 |
|---|---|---|---|---|
| 1400 | T | T | T | 7* |
| 1500 | T | T | 8* | 8 |
| 1600 | T | T | 9* | 9 |
| 1700 | T | 10* | 10 | 10 |
| 1800 | T | 11* | 10 | 10 |
| 1900 | 12* | 11 | 11 | 11 |
| 2000 | 12* | 12 | 12 | 12 |
| 2100 | 14 | 13 | 12 | 12 |
| 2200 | 14 | 14 | 13 | 13 |

appropriate and how the pedestrian scramble WALK phase length changes (when the scramble crossing is recommended). The scramble WALK phase length was acquired from the median value of five minimal user time values by different random seeds and was rounded up to the closest integer. "T" indicates that the two-way crossing is suggested in all five of the random seed runs. "*" indicates that the scramble crossing is suggested in 1, 2, or 3 runs out of 5 random seed runs. In such a situation, it depends on the traffic engineer's personal judgment to decide which pedestrian crossing pattern is more appropriate. The decision would mainly be influenced by how the engineer evaluates pedestrian time value compared with passenger car time value, which can be interpreted as the relative time value (K). In the remaining scenarios, the scramble crossing is suggested in 4 or 5 runs out of 5 random seed runs.

As shown in Table 5, the proposed GA procedure selects suitable scramble WALK phase lengths by considering both pedestrian and vehicle delay. With the increment of the eastbound pedestrian volume or right-turn vehicle volume, the pedestrian crossing pattern switches from a two-way crossing to a scramble crossing. When all the vehicle and pedestrian volumes (except the eastbound pedestrian volume) are fixed, the scramble WALK phase length goes up with the growth of the eastbound pedestrian volume. When the pedestrian and vehicle volumes (except the eastbound right-turn vehicle volume) are fixed, the scramble WALK phase length remains relatively stable with the growth of the eastbound right-turn vehicle volume, such that the right-turn vehicles do not experience longer waiting time.

## CONCLUSIONS

A GA optimization procedure was developed to optimize signal timing of an individual intersection by minimizing total user time which considered both vehicle and pedestrian delays. Compared with the HCS GA function, the proposed procedure has the same accuracy and slightly better capability. The signal plans recommended by the proposed procedure can result in delay values that are slightly less than or

equal to the delay values from the HCS GA optimized signal plans, when there is no pedestrian at the intersection.

However, HCS neither considers pedestrian delay nor discusses the scramble crossing. Therefore, a comparison cannot be made between the proposed GA procedure and the HCS GA function when pedestrian delay is considered. Apart from the GA procedure's ability to consider pedestrian delay, it can realize real time signal plan optimization by shortening the analysis period to cycle length.

The proposed GA procedure can also generate guide tables for intersections with different geometric designs and traffic patterns. With such guide tables, it would be easier and quicker for a traffic engineer to determine which pedestrian crossing pattern is more appropriate and how the pedestrian scramble WALK phase length changes (when the scramble crossing is recommended).

## REFERENCES

Bhattacharya, P., Virkler, M. R. (2005). "Optimization for Pedestrian and Vehicular Delay in a Signal Network". In *Transportation Research Record 1939,* TRB, National Research Council, Washington, D.C., pp. 115–122.

Ishaque, M. M., Noland, R.B. (2007). "Trade-offs between Vehicular and Pedestrian Traffic Using Micro-Simulation Methods". In *Transportation Policy 14*, Elsevier Ltd., pp. 124-138.

Li, M., Alhajyassen, W. K. M., Nakamura, H. (2009). "A Traffic Signal Optimization Strategy Considering Both Vehicular and Pedestrian Flows". *Presented at the 89th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Noland, R. B. (1998). "Pedestrian Travel Times and Motor Vehicle Traffic Signals". In *Transportation Research Record 1553*, TRB, National Research Council, Washington, D.C., pp. 28–33.

Transportation Research Board. (2000). "Highway Capacity Manual". National Research Council, Washington, D.C.

Virkler, M. R. (1998). "Pedestrian Compliance Effects on Signal Delay". In *Transportation Research Record 1636,* TRB, National Research Council, Washington, D.C., pp. 88–91.

Yang, Z. (2010). Signal Timing Optimization Based on Minimizing Vehicle and Pedestrian Delay by Genetic Algorithm. M.S. thesis. University of Illinois at Urbana-Champaign.

# COMPARISON OF SIMULATION-BASED DYNAMIC TRAFFIC ASSIGNMENT TOOLS FOR OPERATIONS MANAGEMENT

Ramachandran Balakrishna*, Daniel Morgan, Qi Yang and Howard Slavin
Caliper Corporation

* *Corresponding author [rama@caliper.com]*

## Abstract

Numerous traffic simulation tools are available to assist in the efficient allocation of resources while operating and managing traffic networks. Their model fidelity ranges from detailed (microscopic) to aggregate (mesoscopic) and intermediate (anisotropic mesoscopic), each approach differing in the level of detail of their component demand and supply models. While each tool may have been demonstrated on specific datasets, these data are generally too varied to allow meaningful comparisons of use to practitioners. Objective evaluations of such tools are often lacking, leading to ambiguity about their modeling assumptions, accuracy, feature sufficiency, running time and scalability. In this paper, we apply three popular traffic simulation tools on a common network and demand data. TransModeler (microscopic), DYNASMART-P (mesoscopic) and DynusT (anisotropic mesoscopic) are tested on a real, urban region in Eureka, CA. Their performance over a range of factors are studied, documented and analyzed from both theoretical and empirical perspectives.

## Introduction

Metropolitan planning organizations (MPOs) and Traffic Management Centers (TMCs) are increasingly seeing the need for high-fidelity traffic models with a temporal resolution that captures day-to-day and within-day dynamics, queuing and congestion patterns. These capabilities are necessary for short-term planning (e.g. work zone scheduling or evacuation planning) and operations management (e.g. incident response). Static approaches are tailored towards long-term processes involving significant changes in land use, residential location choices and auto ownership decisions, and are not equipped to handle within-day dynamics at a level necessary for day-to-day operations. Dynamic Traffic Assignment (DTA) is gaining popularity with its potential to accommodate changes in travel demand and network supply over very short intervals such as 5 to 15 minutes, and the ability to model the spatial and temporal results of their interactions. These interactions are often captured through behavioral models that predict individual drivers' route and lane choices. The advent of powerful computers is accelerating the interest in applying DTA to medium and large networks.

Several DTA approaches based on analytical and simulation techniques have been proposed, implemented and tested in both academic and commercial settings. Peeta and Ziliaskopoulos (2001) provide a comprehensive conceptual review of these approaches, several of which have been developed into tools that are available to modelers. These include DynaMIT (Ben-Akiva et al., 2002), DYNASMART (Mahmassani, 2002), DynusT (UA, 2009), Dynameq (Tian et. al., 2007), TransCAD (Caliper, 2009a) and TransModeler (Caliper, 2009b). However, examples of rigorous field tests are few and on various datasets that differ in their spatial extent, network structure, demand levels, etc. Further, the modeling differences and default parameters across tools make it difficult to objectively evaluate the results published in the literature. In this paper, we present a conceptual analysis comparing the

capabilities and modeling approaches of three DTA tools: DYNSMART-P, DynusT and TransModeler. We apply these tools on a common dataset so that their outputs can be compared, and their performance evaluated.

**Simulation-based dynamic traffic assignment**

As in static planning models, travel demand for dynamic models is specified through origin-destination (OD) trip matrices. However, each matrix contains the trips that depart within a very short time interval, usually between 5 and 15 minutes. DTA packages are built on the premise that a dynamic equilibrium exists in the real world. The definition of this equilibrium is a direct extension of Wardrop's principle along the temporal dimension. Consequently, all used network paths between an OD pair have the same, minimum travel time at equilibrium, *for a given departure time (interval)*. A simulation model thus calibrated may then be used to analyze a variety of scenarios involving perturbations such as incidents and work zones.

In this paper, we focus on DTA packages that are based on traffic simulation principles. Such approaches are broadly classified as microscopic, mesoscopic and macroscopic depending on how they each represent network supply and vehicle movements. A microscopic DTA employs detailed models of a vehicle's interactions with those around it. Such interactions take the form of car following and lane changing maneuvers, with drivers determining their speed, acceleration and lane based on individual desired speed preferences and path choices. Macroscopic models either treat traffic as a fluid, or use volume-delay functions (such as those used in static assignment methods) to capture the interplay between congestion levels and link traversal times. Mesoscopic models provide a theoretical middle ground in which the link performance functions are based on the fundamental diagram typically through speed-density or speed-flow equations.

A DTA's demand input is disaggregated into trips that are loaded onto the network through route selection logic based on either historical or updated travel times. At the end of the simulation, the outputs are aggregated and used to compute the input travel times for the network loading in the next iteration. This is essentially a fixed point problem of the form:

$$x = f(x)$$

where $x$ are the desired congested travel times specified by departure time, and $f(\bullet)$ is the simulation loading. A detailed treatment of the solution of fixed point problems in the transportation context can be found in Bottom (2000).

DTA implementations can vary widely in their modeling assumptions, features and capabilities. The remainder of this section briefly summarizes these differences for three packages selected for their different approaches to modeling traffic dynamics.

*DYNASMART-P* uses a mesoscopic simulation approach in which vehicles are propagated on a link-node network representation. Routes are assigned to vehicles based on a k-shortest path algorithm, with a provision for some vehicles to make en-route switches. The software is based on the principle of bounded rationality (Simon, 1957), so that an alternative route must provide some minimum improvement in order for drivers to switch routes. Vehicles are propagated along their routes according

to a queuing model with link output capacities acting as constraints. Vehicle speeds are set using aggregate speed-density relationships that follow a modified version of the Greenshields formula. The algorithm is expected to move towards a dynamic equilibrium by feeding the output travel times back into the k-shortest path algorithm, so that vehicles may re-evaluate their route selections.

*DynusT* is a mesoscopic DTA package that shares many modeling elements with DYNASMART-P. Its key difference, though, is its Anisotropic Mesoscopic Simulation (AMS) for network loading. In this approach, each vehicle's speed is set as a function of the speeds of the vehicles in its Speed Influencing Region (SIR). This region is defined to include the vehicle immediately downstream of the current vehicle, and up to one vehicle in front of but on the adjacent lanes to the left and right. The AMS approach thus has a more microscopic focus than traditional mesoscopic models.

*TransModeler* is a 4D geographic information system (GIS), in which the temporal dimension is added to a 3D representation of the transportation system. The network, vehicles and other infrastructure (including control devices, message signs, etc.) are stored as geographic layers that are updated as the simulation progresses. TransModeler's microscopic loading is based on car following and lane changing logic, and the operation of traffic control infrastructure is simulated in great detail. TransModeler also has the ability to perform DTA at mesoscopic and macroscopic fidelities, and is unique as a hybrid simulation that allows sets of segments to be modeled at different fidelities (i.e. microscopic, mesoscopic or macroscopic) *within the same simulation run*. Large networks at the regional scale may therefore be simulated without having to scale back the study area for computational reasons. The speed calculations in the mesoscopic fidelity are captured through speed-density relationships defined for various facility types (e.g. freeway, expressway, ramp, urban street). A vehicle's speed in the macroscopic fidelity is calculated from a volume-delay function (VDF) similar to those deployed in regional planning models. The DTA functionality can be based on any fidelity desired by the modeler, including hybrid simulation.

All three packages outlined above employ an iterative algorithm to attempt to solve for dynamic equilibrium. Typically, the outputs from one iteration feed back into the next after an update of either route flows or network travel times. For illustrative purposes, the travel time updating mechanism in TransModeler is shown in Figure 1.

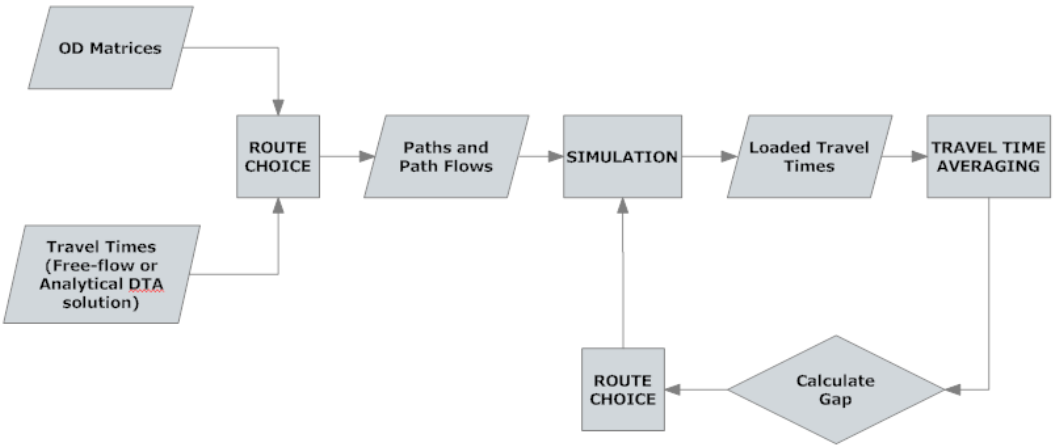



**Figure 1. Travel Time Updating in the TransModeler DTA**

In the above flowchart, a time-dependent series of OD matrices and any available travel time estimates (congested or free-flow) are input to a route choice model that splits the OD flows into path flows. These flows are simulated on the network and the experienced (loaded) travel times are logged. A travel time averaging function is then applied before re-evaluating drivers' route choices. If the travel times have not converged (i.e. a solution to the fixed point problem has not been identified), the new route choices (and hence path flows) are simulated in the next iteration. The process is repeated until the relative gap between the input and output travel times for the iteration falls below a user-specified value. The following travel time averaging scheme is used:

$$x_{i+1} = (1-\alpha_i)x_i + \alpha_i f(x_i)$$

The choice of the factor $\alpha_i$ will determine the type of averaging, such as the Method of Successive Averages (MSA), Polyak averaging or fixed-factor averaging (see Balakrishna et. al. (2009) for details). TransModeler's DTA is evaluated using the relative gap measure, calculated for each departure time interval and DTA iteration:

$$Gap^{\tau} = \frac{\sum_{i \in I} \sum_{k \in K_i} f_k^{\tau} t_k^{\tau} - \sum_{i \in I} d_i^{\tau} t_{\min,i}^{\tau}}{\sum_{i \in I} d_i^{\tau} t_{\min,i}^{\tau}}$$

where:

| | | |
|---|---|---|
| $Gap^{\tau}$ | = | Relative gap in time interval $\tau$; |
| $I$ | = | Set of all O-D pairs $i$; |
| $K_i$ | = | Set of paths used by trips traveling between O-D pair $i$; |
| $f_k^{\tau}$ | = | Number of trips taking path $k$ (i.e. path flow) in time interval $\tau$; |
| $t_k^{\tau}$ | = | Travel time on path $k$ in time interval $\tau$; |
| $d_i^{\tau}$ | = | Demand departing in time interval $\tau$; |
| $t_{\min,i}^{\tau}$ | = | Travel time on shortest path between O-D pair $i$ in time interval $\tau$; |

## Methodology

The comparison of different DTA packages is complicated by several factors. First, descriptions of existing applications as reported in the literature are largely not directly comparable since they are obtained from different networks, datasets and input parameters. Second, the modeling approaches and (default) model parameters often vary widely across the packages. Third, it is not obvious if sufficient model calibration has been carried out in each case, using real-world traffic data. Finally, the definition of performance measures and convergence thresholds can vary across packages and studies. In this study, we propose to reduce the effects of many of the shortcomings outlined above.

A key aspect of our methodology is the selection of a common network location for testing purposes. The network from Eureka, CA (Figure 2) was chosen to represent a dense, urban network with actuated signal control and numerous stop signs in addition to freeway corridors. The area also has significant pedestrian traffic. This dataset will be used as the backdrop to our discussion of the study methodology.

The TransCAD planning model for the Greater Eureka Area was used as the basis for generating the computer representation of the network in each of the three packages. Since TransModeler shares the same native GIS platform as TransCAD, the transition from the planning model to a simulation model was seamless. The network in Figure 2 contains 1461 nodes, 2298 links and 417 zones.

A small sub-area (Figure 3) made of 80 nodes, 99 links and 35 zones was identified for initial testing. The intersections in this sub-area were all controlled either by stop signs or actuated signal logic. TransCAD's sub-area analysis feature was used to create a sub-OD matrix of trips passing through the test network.

Both DYNASMART-P and DynusT provide importers for (TransCAD) GIS formats. These features were used to first export the TransCAD node, link and zone data to Microsoft Excel worksheets and text geography (*.geo) files before reading them into the DTA packages. It must be acknowledged that DYNASMART-P and DynusT are based on node-link networks rather than a GIS engine. Two-way links in the GIS had to be manually coded as two separate directed links before importing into DYNASMART-P and DynusT. These aspects are time-consuming and error-prone when developing models for large, urban regions. The resulting network is displayed in Figure 4.

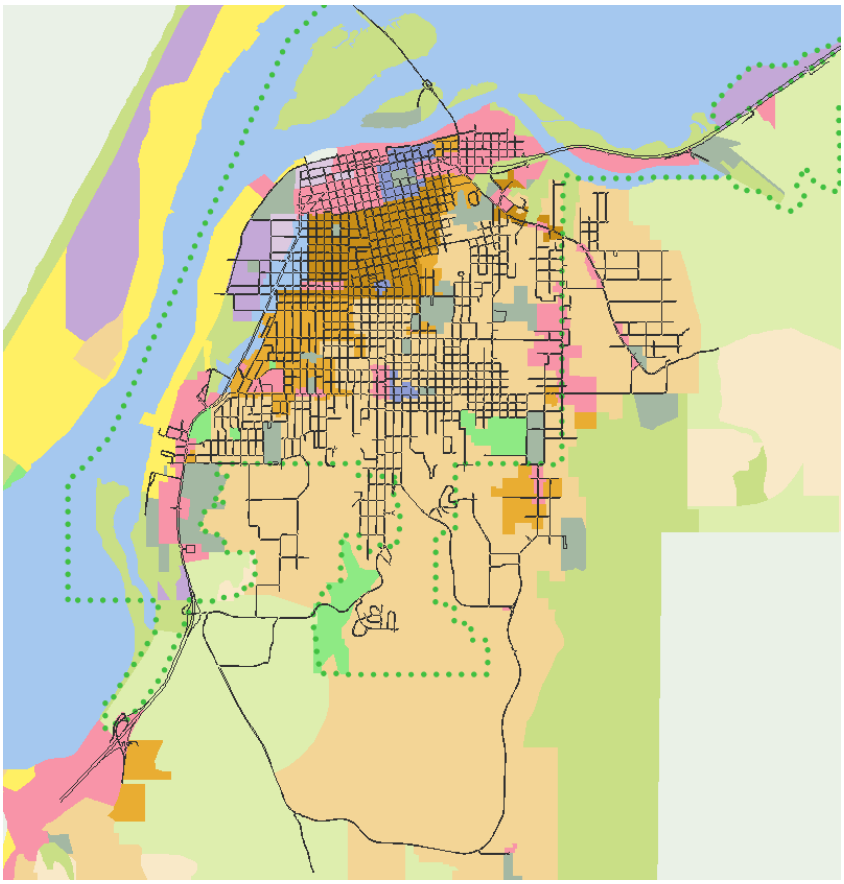


**Figure 2. The Greater Eureka Area in TransModeler**

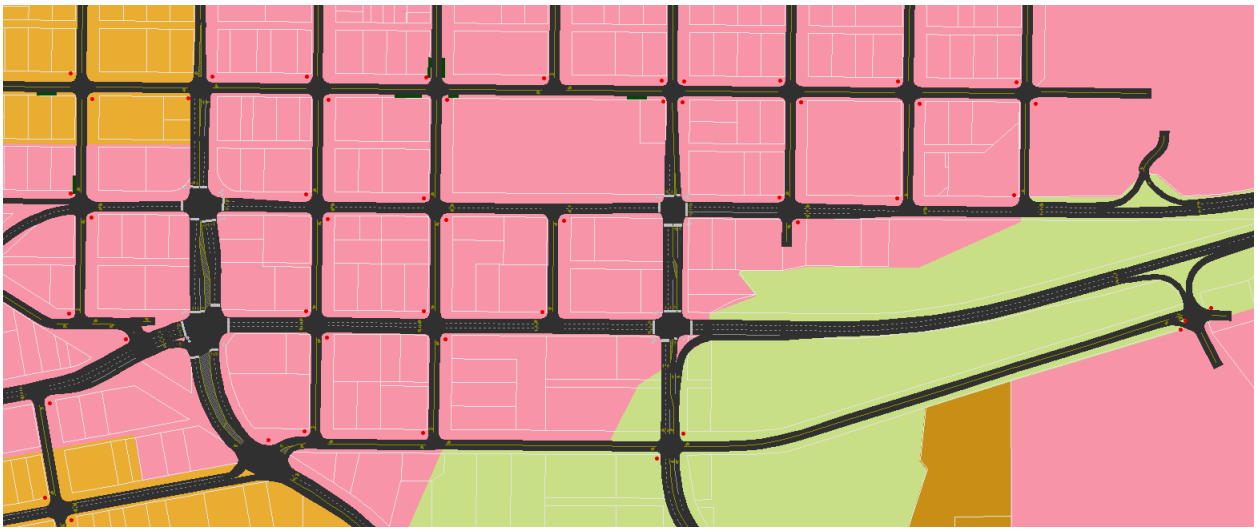


**Figure 3. TransModeler Sub-Area used for Initial Tests**

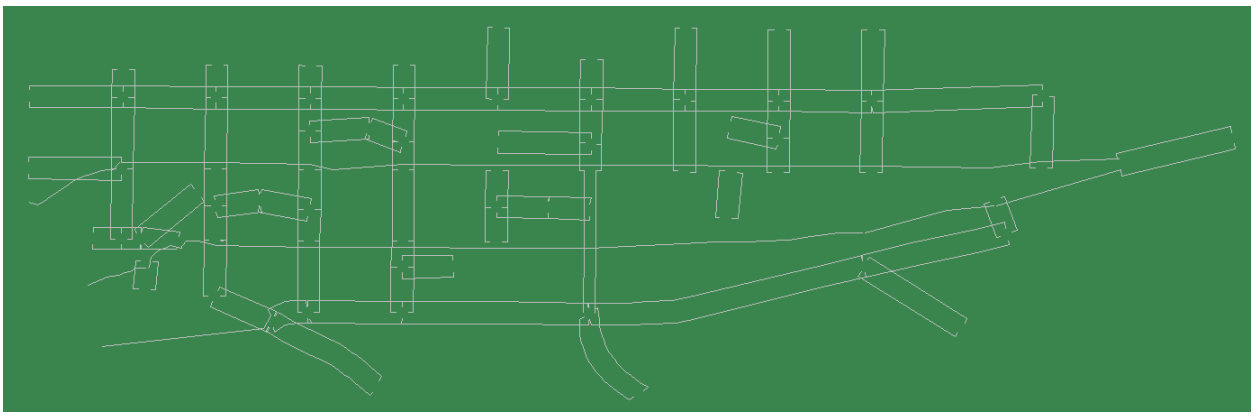


**Figure 4. DYNSMART-P and DynusT Sub-Area used for Initial Tests**

All three tools selected for this study are based on dynamic traffic simulation. However, their core abilities from a practitioner's viewpoint are rather different. For example, the tools vary in the fidelity of their handling of ITS strategies. TransModeler simulates the operation of complex signal timing logic, toll plazas, lane use signs, ramp meters, HOV (high occupancy vehicle) and HOT (high occupancy toll) lanes, etc. Figure 5 illustrates the flexibility and level of detail in modeling actuated signals. In addition, simulated drivers make lane choices based on driver characteristics (e.g. aggressiveness in seeking gaps for lane changing) and prevailing traffic conditions (e.g. speed of vehicles in the vicinity). This approach results in highly realistic reproductions of traffic phenomena, such as the queuing due to a single vehicle slowing down.

DYNASMART-P and DynusT are restricted in their handling of such scenarios. While they can incorporate the impacts of HOV vehicles and basic traffic control, they are unable to handle routinely deployed technologies such as ring-barrier traffic actuated control, or special geometric configurations such as three-way stop control or intersections with more than four legs. Figure 6 shows the maximum extent of their actuated control capabilities, which approximates TransModeler's cycle length of 93.6 seconds with a simpler logic covering 121 seconds. Their internal network representations are also not

as precise as TransModeler's GIS, and this propagates modeling errors due to inaccuracies in link lengths, vehicle storage capacities under queuing, etc. The above meso models also impose strict minimum link length constraints based on a link's maximum speed and the length of the simulation's basic time step. Both user manuals acknowledge that the presence of several short links can significantly distort the internal network representation used by the model. This can be a cause for concern, as urban networks are inevitably made up of short links between intersections, as well as short ramps. Given these limitations, the Eureka network and control data were incorporated into DYNASMART-P and DynusT to the best allowed extent.

Since the different DTA packages may employ different model paradigms to capture the same real-world traffic phenomena, the default model parameters provided with each DTA package must be analyzed and compared to check if they represent similar processes. While DYNASMART-P and DynusT are reasonably close in their use of speed-density relationships, comparison across fidelities can be harder: TransModeler's microscopic fidelity uses driver behavior mechanics such as car following and weaving, while the mesoscopic models rely on more aggregate relationships. The applicability of the default model parameters supplied with each DTA package will be tested empirically against field measurements to ascertain if the models need further calibration before a comparison can be effective. The use of field data provides an objective baseline that is exogenous to the models themselves.

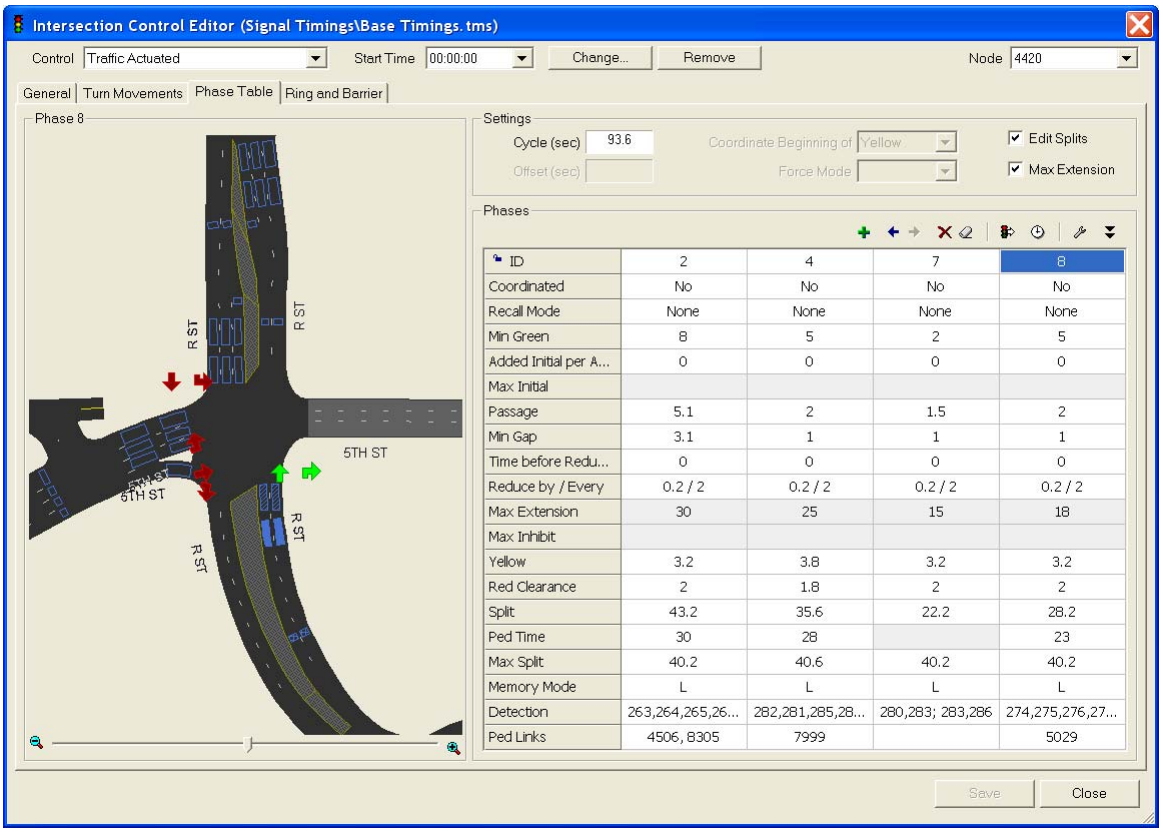



**Figure 5. Actuated Signal Control Interface in TransModeler**

Though all three packages have graphical interfaces to simulation model output, the visual results are not directly comparable. TransModeler uses its GIS engine to allow the user to color-code the vehicles, links and other elements even while the simulation is running. This can be a powerful feature to

visualize network and vehicle performance from the perspectives of virtually any simulation output such as speed, acceleration, flow, etc. In addition, multiple themes can be set so that each layer is color-coded simultaneously (for example, the links can be coded to show average speed while the vehicles are coded to identify their next turn movements). The meso models in this study lack this flexibility, and are compounded by their somewhat unintuitive choice of outputs. DynusT, for example, color-codes its link volumes by the number of vehicles on the link (rather than the standard vehicles per hour). The speed values are shown as the percentage of their maximum values, while queue lengths are shown as the percentage of link length (rather than a distance measure). Densities, however, are displayed in the standard unit of vehicles/lane-mile. These non-standard unit selections make it hard to understand the correlations between the various color maps (for example, high-density links may have a wide range of colors in the speed map). Further, vehicles are shown as uniformly scaled dots on a center-line network, so that vehicles on adjacent lanes at the same location are collapsed into a single dot. The user must therefore open several ASCII text files even to check for basic reasonableness.

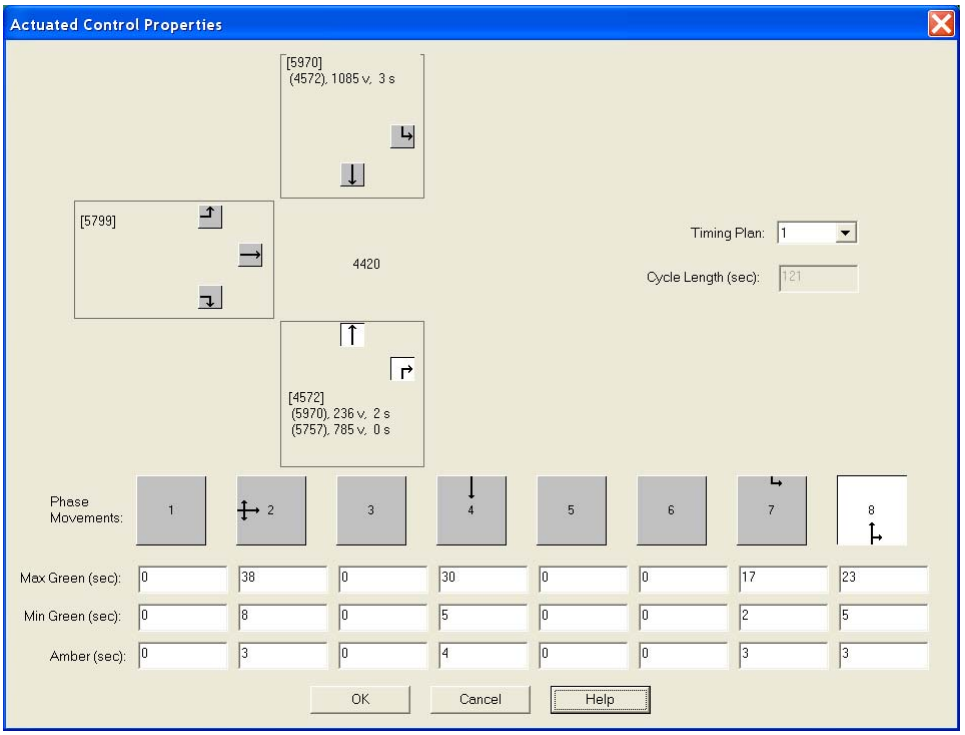



**Figure 6. Actuated Control Interface in DYNASMART-P and DynusT**

## Results

The development of mesoscopic simulation has largely been motivated by the perceived inability of microscopic simulation to yield reasonably fast results when executed iteratively (as in a DTA) on medium and large networks. As more computing power becomes affordable and available on desktops, it is useful to re-evaluate this motivation to ascertain if the more realistic microscopic processes can be retained for DTA. As part of this study, a comparison between the microscopic, mesoscopic and anisotropic network loading models is currently being performed on the full Eureka network. Both the accuracy of the model outputs and the computational effort required by each of these approaches are being evaluated to draw conclusions about the tradeoffs involved. At issue is the extent of accuracy loss

when moving from microscopic to mesoscopic models, and the inherent value of the accompanying running time savings.

Figure 7 illustrates the convergence properties of TransModeler's DTA based on a microscopic network loading process, for the entire Greater Eureka network depicted in Figure 2. It can be seen that the system moves consistently towards an equilibrium. These results were computed on a 6-core i7 Intel processor (3.33 GHz) with 24 GB of physical memory, a hardware specification that is an affordable desktop option today.

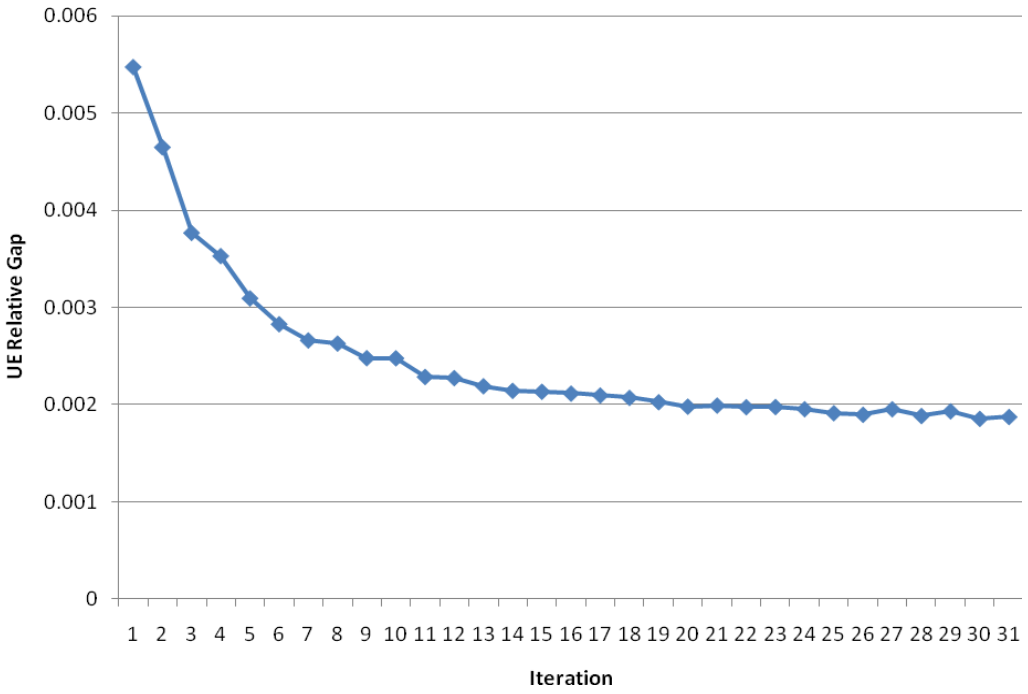



**Figure 7. Relative Gap Convergence for Eureka**

The importing of the full Eureka network (shown in Figure 2) into the mostly common format shared by DYNASMART-P and DynusT was more involved than anticipated. While the data importer program NEXTA is provided for the conversion of GIS datasets, there are time-consuming pre-processes that must be executed to first export all the GIS data into Microsoft Excel worksheets with rigid requirements. This caused loss of significant information regarding the network and zone structures. Figure 8 illustrates the imported network, while Figure 9 contains two close-up examples of disconnected links.

A trial DynusT run on the imported dataset failed because several destination zones were not associated with any links. This happened as a result of the Excel template's limitation of specifying exactly one zone centroid ID for any given link, while the planning dataset and TransModeler database often had multiple zones connected to different loading points on the same link. Additional manual work is thus needed in order to assign more than one zone centroid ID to any given link or node in the network. We are currently attempting to further study the structure of the DynusT input text files (generated automatically by the NEXTA conversion program) so that we may edit the contents directly to better reflect the original GIS database.

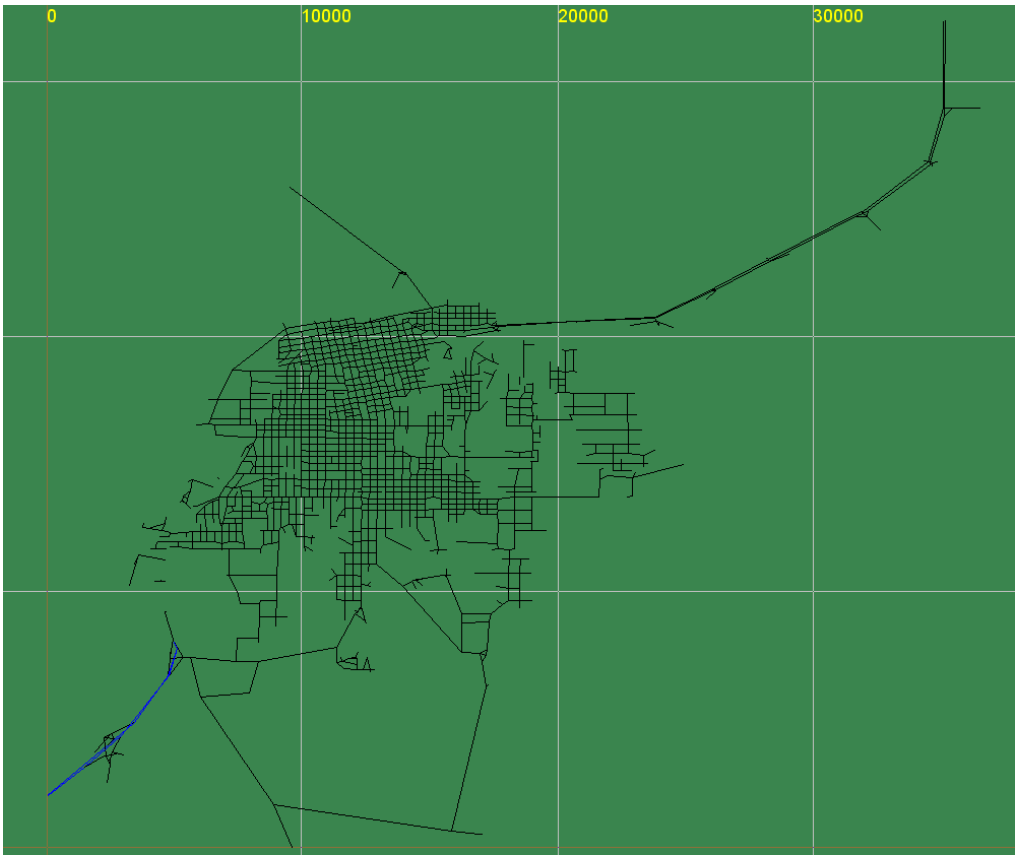



**Figure 8. The Greater Eureka Network in DYNASMART-P and DynusT**

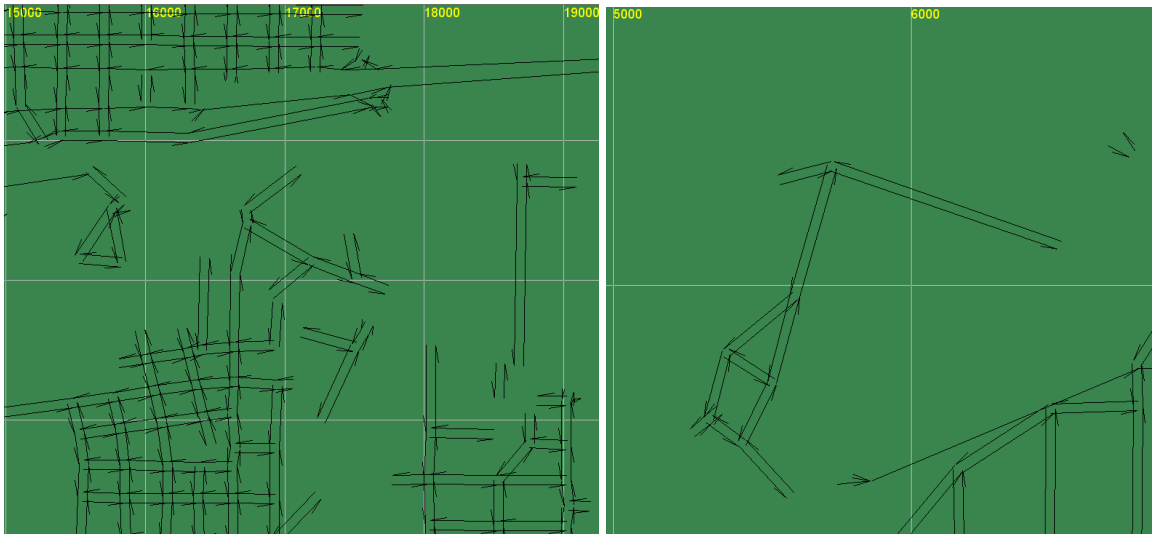



**Figure 9. Sample Locations Requiring Detailed Manual Intervention**

## Conclusion

This paper briefly summarizes the demand and supply modeling concepts of three simulation-based DTA tools: DYNASMART-P, DynusT and TransModeler. The focus of this study is to (a) understand

the basic theoretical differences in these tools, (b) compare their modeling capabilities, and (c) empirically explore their ability to replicate the real world. Rigorous numerical testing is currently underway to compare the different approaches in terms of their ability to reliably capture dynamic traffic conditions and equilibrium. A common network and demand data from Eureka, CA are being used so that the comparisons are performed with an objective baseline. The complete set of numerical results is expected to be available within two months.

## References


1. Balakrishna, R., D. Morgan, H. Slavin and Q. Yang (2009) "Large-Scale Traffic Simulation Tools for Planning and Operations Management." 12th IFAC Symposium on Control in Transportation Systems, Redondo Beach, California.
2. Ben-Akiva, M., M. Bierlaire, H. N. Koutsopoulos and R. Mishalani (2002) "Real-Time Simulation of Traffic Demand-Supply Interactions within DynaMIT." In Gendreau, M. and P. Marcotte (eds) Transportation and Network Analysis: Miscellenea in honor of Michael Florian, Kluwer, pp. 19-36.
3. Bottom, J. (2000) "Consistent Anticipatory Route Guidance." Ph.D. dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology.
4. Caliper (2009a) "TransCAD Transportation Planning Software: Travel Demand Modeling with TransCAD – User's Guide."
5. Caliper (2009b) "TransModeler Traffic Simulation Software – Version 2.5 User's Guide."
6. Mahmassani, H. S. (2002) "Dynamic Network Traffic Assignment and Simulation Methodology for Advanced System Management Applications." 81st annual meeting of the Transportation Research Board.
7. Mahmassani, H. S. and H. Sbayti (2006) "DYNASMART-P Version 1.2 User's Guide".
   Peeta, S. and A. Ziliaskopoulos (2001) "Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future." Networks and Spatial Economics, Vol. 1, No. 3/4, pp. 233-266.
8. Simon, H. (1957) "Models of Man." Wiley, New York.
9. Tian, X., M. Mahut, M. Jha and M. A. Florian (2007) "Dynameq Application to Evaluating the Impact of Freeway Reconstruction." 86th annual meeting of the Transportation Research Board.
10. UA (2009) "DynusT ONLINE USER'S MANUAL (version 2.0.0.1 Beta)". Accessed on 28th October, 2009.

# Comparing data from mobile and static traffic sensors for travel time assessment

**Nicolas Saunier and Catherine Morency**

Department of civil, geological and mining engineering, École Polytechnique de Montréal, P.O. Box 6079, Station Centre-Ville, Montreal, PQ, CANADA, H3C 3A7; emails: nicolas.saunier@polymtl.ca, cmorency@polymtl.ca

## ABSTRACT

Travel time and speed measures on road networks provide key information to identify critical spots of congestion and evaluate the scale of this phenomenon across an urban area. Many technologies are currently available to measure travel time and speed, but each has its limitations. As part of a wider project aiming to develop travel time reliability indicators, this paper presents a comparison and validation of data collected through four different non-intrusive techniques: floating cars, GPS traces, Bluetooth detections and video processing. First, some background information regarding the project and the modeling of travel times and speed on highway networks is provided. Then, a comparison between the various sources of data is presented. Finally, the characteristics of the various data sources are discussed based on the relevance and the availability of the sources.

## INTRODUCTION

With the increasing problems of traffic congestion in urban areas, transportation planners need better tools to assess its evolution through the years. Travel time measures on road networks provide key information to identify critical spots of congestion and evaluate the scale of this phenomenon across the area. Many technologies are currently available to measure travel time at specific locations: license plate matching, loop detectors, Bluetooth (BT) device matching, and video processing. Traditionally, at a larger scale, travel time is estimated using mandated floating cars (FC) that run on specific routes to gather data on the road network. Nowadays, ad-hoc, or randomly collected, data can also be obtained from GPS devices aboard individual cars and commercial vehicles. In addition, interest for video data is growing since various computer vision techniques permit automated traffic monitoring and data collection. However, each of these data collection methods has limitations that must be addressed through the right aggregation method.

As part of a wider project aiming to develop travel time reliability indicators, this paper presents a comparison and validation of data collected through four different non-intrusive techniques: FC sample, GPS traces, BT detections and video-based traffic data. These sources are available for various locations of the Montreal highway network and time spans: the first two mobile sources, GPS and FC, provide traffic data on large spatial areas, but limited in time, while BT devices and video data may be collected continuously, but from static sensors at specific locations.

First, some background information on the research project and available sets of data is presented. Then, the analysis framework is presented along with a statistical description of the travel times and spot speed data extracted. The next section

presents the possible comparisons between travel condition patterns identified in the various datasets. The conclusion and perspectives are finally provided.

## BACKGROUND AND INFORMATION SYSTEM

### Project: assessing the reliability of the highway network

In 2008, the Quebec Ministry of Transportation (MTQ) mandated Polytechnique to assess the potentialities of historical FC data to provide relevant estimates of the reliability of the Greater Montreal Area highway network. Six years of FC data were analyzed and used to model the evolution of travel times and variability of travel times over time. Also part of the project was the evaluation of the available technologies to monitor the evolution of travel times and/or speed on the main highway network. Experimentations are currently conducted to compare outputs of various data collection methods. The purpose of this comparison process is to formulate recommendations with respect to the value and relevance of available datasets, tools or technologies to provide critical information for the strategic planning of the network by decision makers.

### Evaluating congestion on road networks

Better understanding the use of transportation network is a key factor for the enhancement of transportation planners and models. Hence, more and more studies are conducted to identify the various negative drawbacks of increasing congestion in urban areas. Issues related to sustainable development are actually providing new incentives to better understand and monitor congestion on transportation networks. For instance, studies on environmental impacts of congestion (Nesamani et al. 2005) as well as on the modification of activity rhythms of households and individuals, namely the reduction in shopping related trips (Schmöcker et al. 2006), are being conducted. Also, numerous research works are linked to the definition of sustainability and the various indicators to assess its level. Litman (2008) provides a very interesting view as well as an extensive list of indicators assisting the objective assessment of sustainability. Indicators such as commuting travel time or delays due to congestion are set as economical indicators in the evaluation framework.

Along with the increasing need to better understand congestion problems is the increasing availability of information technology that output multitudes of data on the movements of objects (people, vehicle, phones and other devices). This data can be processed to contribute to the measurement of critical indicators such as travel times or speeds. The emergence of services provided by companies (such as INRIX and Google) relying on distributed data collection and aggregation is an obvious witness of the spreading availability of multiple layers of transportation-related data and of the value of such information.

In the United States, the recurrent reports by Schrank and Lomax (2003, 2005, 2007) illustrate the relevance of estimating comparable congestion-related indicators for the main areas of this country. In the same perspective, FHWA (2004) is pursuing a continuous effort to develop and estimate indicators through its Mobility Monitoring Program. This program relies on the data outputted from detectors available in some thirty areas. Another interesting contribution is from the

Pennsylvania DOT (Szekeres and Heckman 2005) that implemented a Congestion Management System at the state level.

## Information system

### *Study area*

The study relies on various datasets gathered for routes and highway segments of the Greater Montreal Area (GMA). The GMA accounts for more than half of the Québec population and is the second largest area in Canada, after the Toronto Area. The most congested and intensely used highway segment of the Province of Québec is located on the Montreal Island (Highway 40, between the two connections with Highway 15). Many urban highways of the region act as main through corridors and face increasing congestion level.

### *Floating cars*

This research project first started with the systematic processing of FC data that were collected at regular intervals by the MTQ to assess travel times over the main highway corridors. From 1998 to 2004 (no data collection in 2003), 29,229 FC routes were collected (over 51 months). The data collection process was articulated around 55 different routes for a total of around 800 kilometers of road segments.

### *GPS traces*

The GPS traces are made available by the Québec carsharing company (Communauto inc.) who owns in the GMA a fleet of 1000 cars from which 400 are equipped with GPS devices. The shared cars can be used by the 20,000 members of the area to either travel short or long distances. For the current project, all the data from the 2009 years was made available. More than 13 millions of data points were made available with an average sampling period of 5 minutes.

### *Video data*

The Québec highways, especially in urban areas, are well covered with traffic cameras, for use in traffic management centers and for traveler information through their display on the MTQ website. In addition, the MTQ regularly carries out traffic studies that require video data collection, typically for manual data extraction or verification. This work demonstrates that such data can be analyzed to extract traffic data.

### *BT devices*

MTQ has also collected travel times data from three experimental BT devices in operation for a 2 weeks period in the Greater Montréal Area. The experience was not as straightforward as expected since one of the device stopped working early in the process but they still manage to collect 5318 positive matches between the two remaining devices (Fournier 2009). Travel times data is derived from the matches.

## METHODOLOGY AND STATISTICAL DESCRIPTION

### Modeling the distribution of travel times using FC data

The first stage of the project was to propose a model of travel time distribution based solely on FC historical data. Since previous modeling attempts had failed to draw significant results, it was decided to transform the initial data source into travel time data over road segments of constant one-kilometer length. Each FC route was automatically cut into one-kilometer segments and travel times estimated for each.

The second stage of the modeling process was to create clusters of segments based on the similarity of frequency distribution of travel times. Various classifications were tested using a k-means algorithm and segments were finally assigned to one of eight clusters (two distinct sets of 8 clusters for two periods of the day) (Loustau et al., 2010 a and b). For each cluster, a modeled frequency distribution of travel times is derived using three additive log-normal laws. These models are used for comparison purposes in the following sections.

### GPS traces to spot speed data

With their limited temporal resolution (one point every 4-5 min), the current GPS traces are not suitable to provide accurate tracking of the cars over the network. Hence, it was decided to extract spot speed data from every point by matching to road segments using proximity and heading similarity. Every GPS point is linked to the nearest segment using GIS proximity function and then extracted for further analysis if the two following criteria are respected:

- $dist(GPS(t) - RS) < 25\ m$
- $|A(GPS(t) - A(RS)| < 10\ degrees$

where *GPS(t)* and *RS* stand respectively for the GPS point at time *t* and the nearest point on the nearest road segment, and *A()* is the angle of the vehicle heading or road segment with respect to a common direction, e.g. the north.

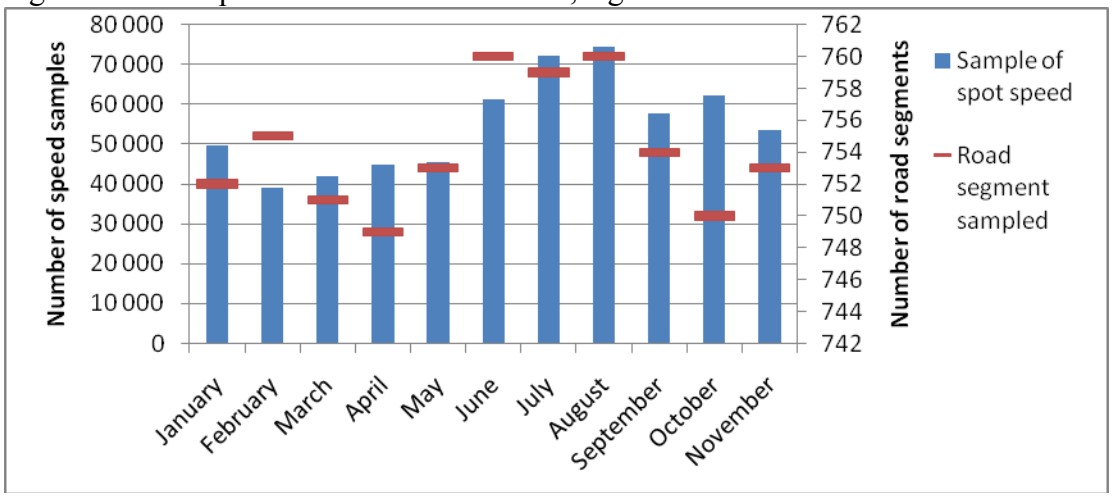



**Figure 1. Description of the monthly sample available using GPS traces from shared cars.**

Figure 1 summarizes the spot speed data extracted from the GPS traces over the 2009 year (11 months). The GPS traces provide between 39,000 and 74,000 spot speed data matched on the road segments that are typically surveyed using FC. The highest samples (both in terms of spot speed data and road segments with at least one

observation) are observed during the summer months where the carsharing service has its highest level of usage.

### BT detection to travel times

Because of the battery failing of one of the BT sensors, matches between the two stations can be obtained only for about 6 days, from September 17th to 22nd 2009. Data post-processing is necessary to derive the travel times (Haghani et al. 2009). The first detection at each station was used and very large travel times were discarded (more than 1000 s for the 1.555 km, or, given the mean travel time $\bar{t}$ over 6 min and the corresponding standard error of the mean $\sigma$, travel times $t$ such that $|t - \bar{t}| \geq 1.5\sigma$). For aggregate results, the travel times were averaged per 6 min intervals. Travel times are then converted into speed data to allow comparison with spot speed data extracted from GPS traces, using the simplest hypothesis of constant speed between the two devices. Figure 2 presents the daily speed patterns obtained from the BT recognition process in one direction. We clearly see the differences in travel conditions between weekdays and week-end days as well as the anticipated peak period on Friday afternoons. Comparisons will hence be conducted separately for weekdays and weekends.

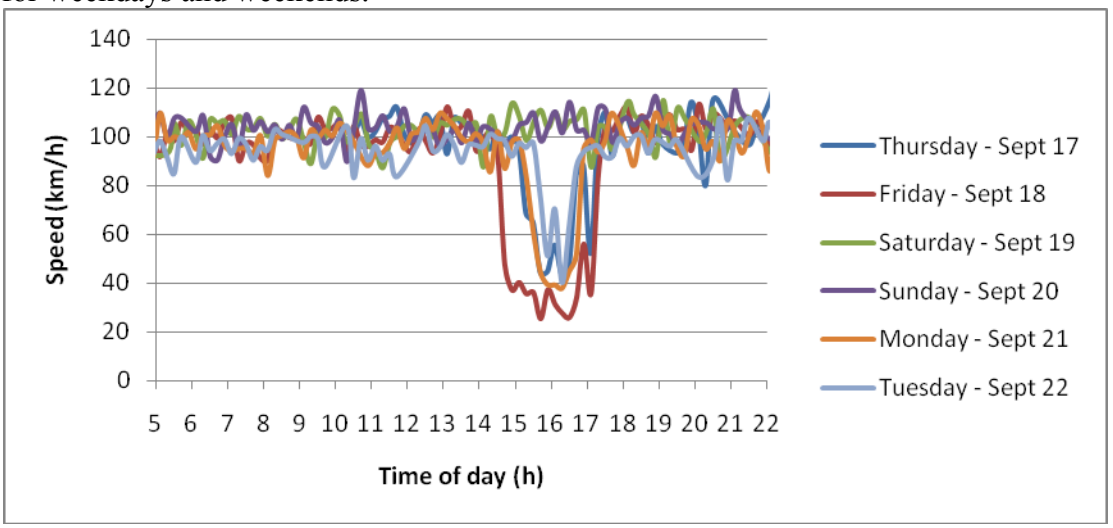



**Figure 2. Pattern of mean speed extracted from 6 days of BT recognition data.**

### Video sequences to spot speed data

Although video cameras may cover large areas, a few tens to hundreds of meters, depending on their resolution and installation, these areas are limited with respect to the scale of the road network. Multiple cameras could cover long corridors, but are not available in this work. Spot speed, counts (flow) and density may be extracted for a specific location covered by a video camera.

Various computer vision techniques have been proposed to analyze movement in video data. Although extracting speed data does not require in itself to track all vehicles, this work makes use of a feature-based road user tracking system developed previously at the University of British Columbia for automated road safety analysis (Saunier et al. 2006, 2008 and 2010). Feature-based tracking is robust to partial occlusions and does not require any initialization step. One requirement is to be able to convert measurements done in the image space to measurements in real world

coordinates, typically at the ground level. A robust camera calibration tool for urban traffic scenes developed previously was successfully used for this purpose (Ismail et al. 2010).

Once vehicle trajectories are extracted from video data, speeds are computed by a moving average of the vehicle displacements. At each location, two lines were identified between which vehicle speeds were averaged to yield a spot speed observation. Obvious outliers with very large mean speed, speed standard deviation and mean acceleration were discarded. The lane of each vehicle was roughly estimated based on their average lane positions; average speeds and 95 % confidence interval are computed for 5 min intervals. This process was applied to three locations and results for a busy tunnel entrance near the downtown area are shown in Figure 3. Spatial and temporal speed variations can be easily identified, especially the slowdowns on the exit of the highway and some spillover on other lanes. As can be expected, speeds are increasing from the rightmost to the leftmost lane.

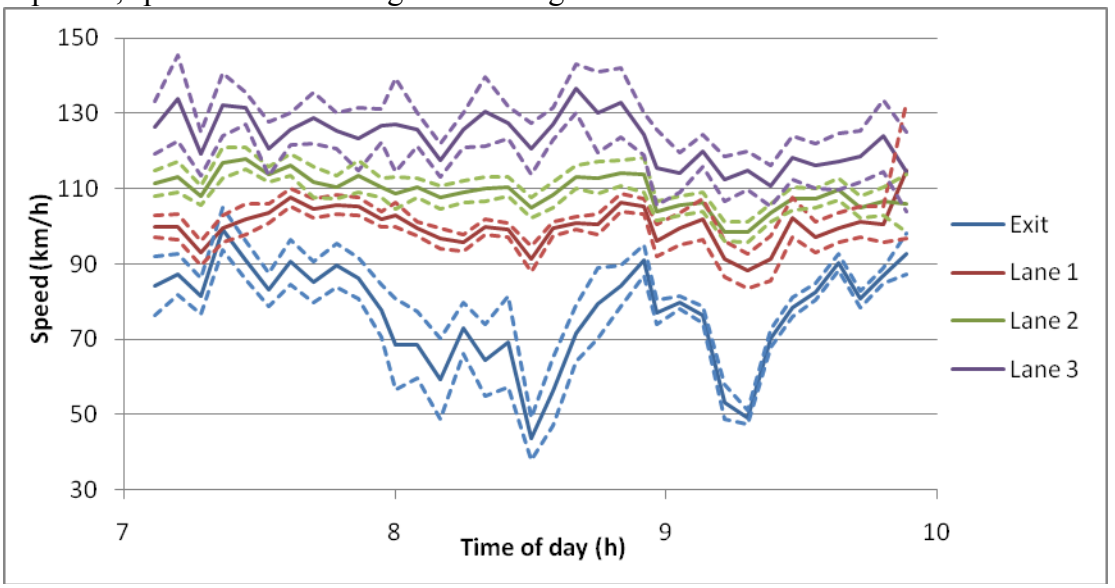



**Figure 3. Average speeds and 95 % confidence interval (dotted lines) extracted from video collected at the entrance of Ville-Marie Tunnel in September 2009.**

## COMPARISONS: RESULTS AND ANALYSIS

### Comparing GPS to FC travel times

Travel time estimations from FC data, based on one-kilometer segments, are compared with travel time estimation from GPS traces (a historical set of smaller size – 2006-2008 – as well as recent sets of 2009). Given the low temporal resolution of the GPS traces, spot speeds extracted from the traces are converted to travel time under the hypothesis of constant speed over 1 km. Figure 4 presents the comparison between travel times over highway 40 that namely contains the most congested segment of the Québec Province (between segment 22 and segment 37). All three sets of data manage to reveal the highly congested segments (all day periods pooled). Actually, the correlation between all series is very high:

- 0.833 between FC and historical GPS traces and from 0.800 to 0.852 between FC and monthly 2009 GPS data

- around 0.950 between historical and 2009 GPS estimates
- at least 0.970 between the 2009 patterns except for the July month for which correlation with fall or winter months is around 0.940

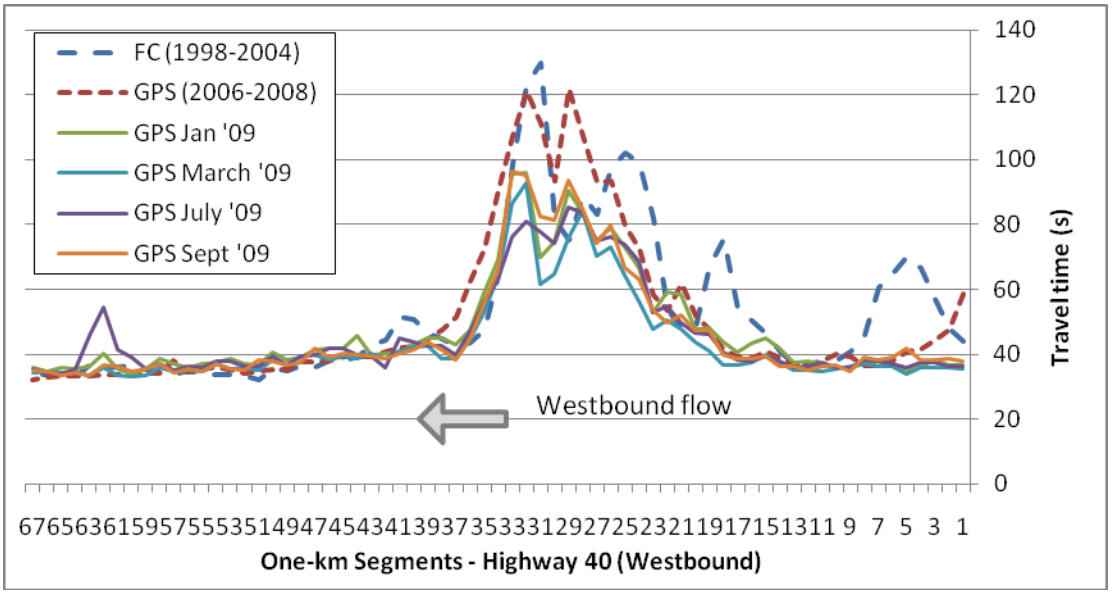



Figure 4. Average travel times estimation using the GPS and FC datasets for highway 40.

## Comparing GPS to video-based spot speeds

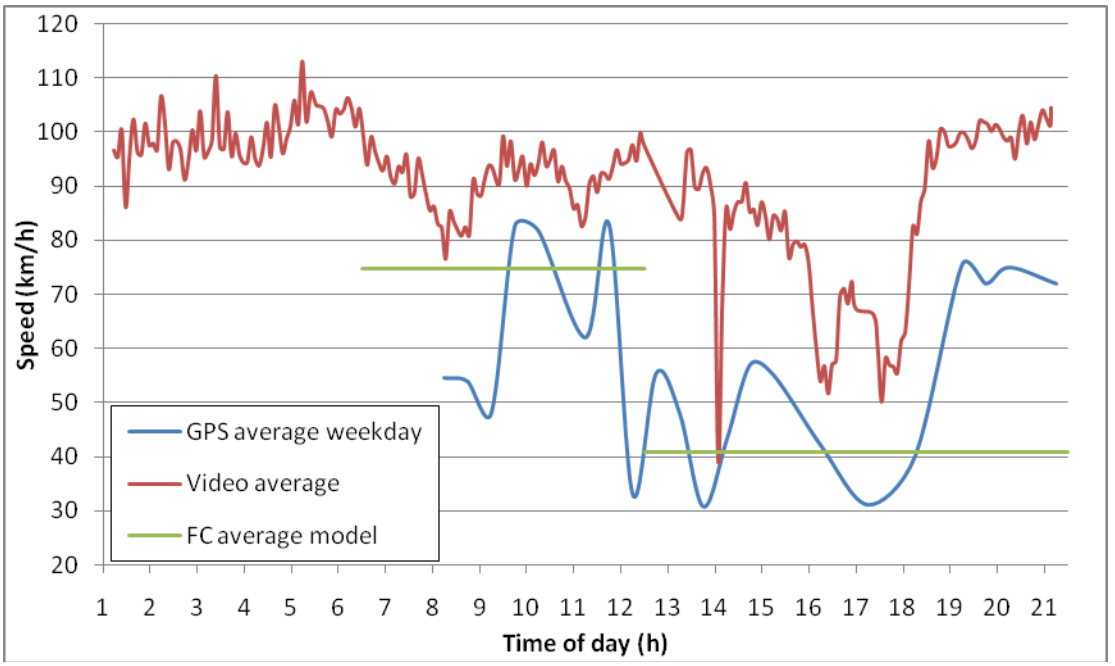



Figure 5. Video-based average speed, GPS speed and VF model average speed (respectively for morning and peak time periods) at the entrance of the Bridge-tunnel Louis-Hippolyte-Lafontaine.

Video data was collected at given locations for limited periods of time (all less than a day). The spot speeds extracted from the longest video recording (1h00 to 21h00 in November 2008) were compared to GPS traces (weekdays of October and November 2009) at the same location and a model estimated from FC data (1998 to 2004) (Figure 5). It is impossible to draw strong conclusions, but the trend is similar

between the different sources of data: fairly fluid traffic conditions in the morning with some slowdowns around 8-9h00 and congestion during the afternoon peak hours, recovering after 18-19h00. The slowdowns detected in the video data were manually verified. The systematically higher speeds obtained from video data compared to the other sources may be caused by perspective.

### Comparing GPS to BT devices

The experiment with BT data was unfortunately done in an area that was not covered by FC. Since only a limited number of GPS spot speed points were perfectly matched in space and time to BT recognition, GPS data from the fall period were processed to produce average patterns of speed over the studied segment (Figure 6).

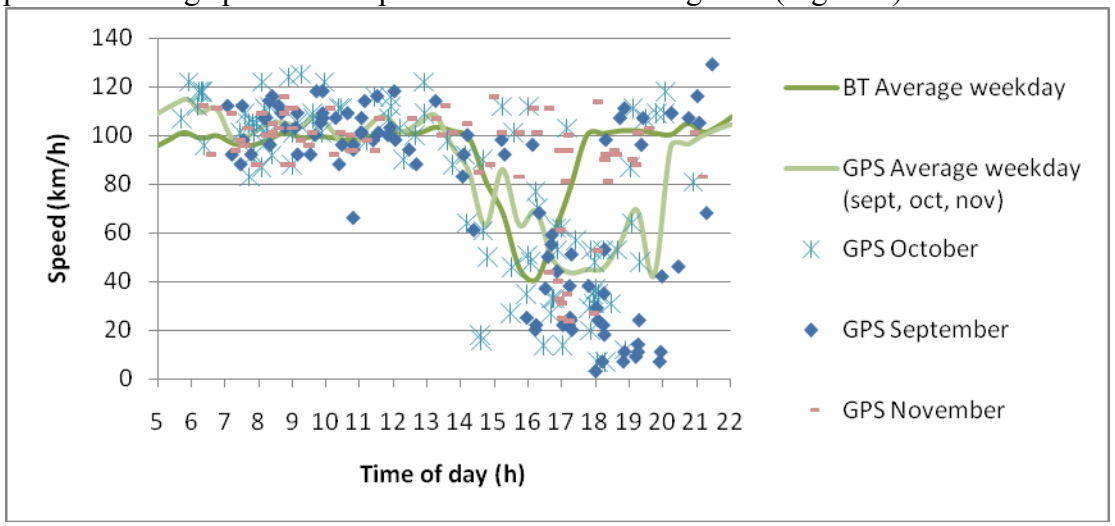



**Figure 6. BT and GPS average weekday speeds near Sainte-Julie in September 2009.**

Correlation between the two average speed patterns obtained with the two technologies is currently not convincing (Pearson r coefficient of 0.431): both technologies agree on non-congested conditions but the peak period is not similarly assessed. However, the three months of GPS data seem to provide similar weekday patterns of travel speed on the segment ($r = 0.660$ between September and October, $r = 0.674$ between September and November and $r=0.623$ between October and November). Uncertainties related to the recognition of the moving object by the BT device or measurement errors of the distance between the two devices could explain some observed differences.

## CONCLUSION AND PERSPECTIVES

Assessing the travel conditions on the road network is a critical task for transportation planners and operators. On the one hand, they need to confirm whether conditions are getting better or worse and provide the relevant information to the users to ease their travels. On the other hand, they need to feed strategic planning with relevant travel time estimations to allow for comprehensive forecasts of future travel conditions, according to various interventions. There was a time when the lack of data limited the development of more precise and extended evaluation of travel conditions. Nowadays, data is becoming easily available via various sources and non-intrusive sensors but their coherent processing and integration is not simple. Every type of data

can potentially contribute to increase the understanding of travel conditions but comes also with inherent limitations reflected in their mutual comparisons.

This paper acts as a first step into the integration of travel times and speed data for the GMA highway network. It has provided preliminary demonstration of the variability of the travel patterns observable on typical road segments. The characteristics of the data sources are summarized in Table 1. Large spatial coverage is useful to model typical travel conditions on various routes, but temporal coverage is required to inform on their variability and evolution. A recommendation is then to supplement costly or uncontrolled data collection through mobile sensors with continuous coverage at specific locations through static sensors: for example, FC data could be updated using GPS data and static sensors in key areas targeted for more complete monitoring. Future work will focus on the collection of larger datasets to allow better comparisons between the various data sources. The next step in this research will investigate the methods for the combination of the data sources.

**Table 1. Comparison of the traffic data sources.**

| | Cost | Spatial coverage | Temporal coverage | Data types | Challenges |
|---|---|---|---|---|---|
| FC | high | moderate | small | speed, travel times | small sample size |
| GPS | low | large | large | speed, travel times | matching GPS positions to the network, varying sample size, no control |
| BT | low | small | continuous | travel times | ongoing development with some detection issues |
| Video | moderate | small | continuous | speed, counts, density, behavior | ongoing development, varying accuracy |

## ACKNOWLEDGMENTS

The authors wish to acknowledge the continuous contributions and support of the MTQ, namely the team of Pierre Tremblay, both in terms of financial support and data provision. They also wish to acknowledge the contributions of Communauto inc that provided access to datasets of GPS traces (Benoît Robert and Marco Viviani).

The authors wish to acknowledge the work of several students from École Polytechnique who assisted in the processing of various datasets: Ali El Husseini (video data), Vincent Grasset (GPS traces), Pierre Loustau (FC data) and Nakul Reddy Nalla (BT data).

## REFERENCES

Fournier, P. (2009) Expérimentation des relevés de temps de parcours à l'aide de capteurs Bluetooth Ministère des transports du Québec.

Haghani, A., Hamedi, M. and Sadabadi, K. F. (2009) I-95 Corridor Coalition Vehicle Probe Project Validation of INRIX Data July-September 2008. I-95 Corridor Coalition. Final Report.

Ismail, K.; Sayed, T. and Saunier, N. (2010) Camera Calibration for Urban Traffic Scenes: Practical Issues and a Robust Approach, Presented at the 89th Annual Meeting of the Transportation Research Boarch (CD compendium of papers), Washington D.C.

Litman, T. (2008). Well Measured. Developing Indicators for Comprehensive and Sustainable Transport Planning, Report, Victoria Transport Policy Institute, 47 pages (http://www.vtpi.org/wellmeas.pdf, checked in May 2008).

Loustau, P., Morency, C., Trépanier, M. and L. Gourvil. (2010a) Travel Time Reliability on a Highway Network: Some Estimations using Floating Cars Data, Transportation Letters, Volume 2, Issue 1 / January 2010, J.Ross Publishing, pp.27-37.

Loustau, P., Morency, C. and Trépanier, M. (2010b) Measuring, Describing and Modeling Travel Time Reliability, Presented at the 89th Annual Meeting of the Transportation Research Boarch (CD compendium of papers), Washington D.C.

Nesamania, K.S., Chub, L., McNallyc, M.G., Jayakrishnan, R. (2005). Estimation of vehicular emissions by capturing traffic variations, Center for Activity Systems Analysis. Paper UCI-ITS-AS-WP-05-2. http://repositories.cdlib.org/itsirvine/casa/UCI-ITS-AS-WP-05-2

Saunier, N. and Sayed, T. (2006) A feature-based tracking algorithm for vehicles in intersections, Third Canadian Conference on Computer and Robot Vision, IEEE.

Saunier, N. and Sayed, T. (2008) A Probabilistic Framework for Automated Analysis of Exposure to Road Collisions, Transportation Research Record, 2083, pp. 96-104.

Saunier, N.; Sayed, T. and Ismail, K. (2010) Large Scale Automated Analysis of Vehicle Interactions and Collisions, Transportation Research Record, 2147:42-50, 2010.

Schmöcker, J.-D., Fonzone, A, Quddus, M., Bell, M.G.H. (2006). Changes in the frequency of shopping trips in response to a congestion charge, Transport Policy, Volume 13, pp. 217–228.

Schrank, D., Lomax, T. (2003). The 2003 Urban Mobility Report, Texas Transportation Institute, http://mobility.tamu.edu, May 2003.

Schrank, D., Lomax, T. (2005). The 2005 Urban Mobility Report, Texas Transportation Institute, http://mobility.tamu.edu, May 2005.

Schrank, D., Lomax, T. (2007). The 2007 Urban Mobility Report, Texas Transportation Institute, checked in May 2008), September 2007. http://tti.tamu.edu/documents/mobility_report_2007_wappx.pdf

Szekeres, D., Heckman, M. (2005). Statewide Congestion Measurement Study – Final Report, Pennsylvania Department of Transportation, Contract N° 030419, Juillet 2005.

# Simulation and Analysis of Emergency Routing Planning Based on Vulnerability Identification of Urban Transportation Network

AN Shi[1], CUI Na[2*], WANG Jian[1], YU Hang[2]

[1] School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin, 150090, China

[2] School of Management, Harbin Institute of Technology, Harbin 150001, China

## ABSTRACT

In this study, link vulnerability index was derived based on a given set of traffic condition and availability of alternate paths when a road link was damaged. The identification results were considered as the basis of traffic assignment. With the purpose of decreasing the system disruption risks, a modified System Optimal (SO) traffic assignment model was presented. The evacuation situations under the conditions of classical and modified SO traffic assignment modes were simulated and compared. The results showed that although the modified traffic assignment mode could lead to an increase in the evacuation network clearance time and average travel time, it could make the system disruption risk decrease about 20%, and the results also provided a useful reference for the urban emergency planning department when making emergency routing strategies.

**KEY WORDS:** Urban emergency routing planning, network vulnerability, vulnerability identification, traffic assignment, Starlogo simulation

## INTRODUCTION

For at least a decade, the vulnerability and reliability of urban critical infrastructures have attracted a lot of attention (Bell, 2000; Chen, 2006; Jenelius, 2009; Sullivan, 2010). However, road transportation vulnerability has not obtained much focus as an explicit subject so far. Sometimes the vulnerability in the context of the road transport network has been confused with the concept of reliability. For example, Berdica (2002) proposed that "vulnerability in the road transport networks can be considered as the complement of reliability". However, reliability was focused on the probability of the road network performing its purpose adequately under the operating conditions. It can be found that reliability is generally concerned with probabilities, which actually often is the case. Taylor and D`Este (2004) chose to separate vulnerability completely from reliability due to the mainly probabilistic base of reliability, which is: "the concept of vulnerability is mainly related to the consequences of link failure, irrespective of the probability of failure". Jenelius et al. (2006) gave a detailed description on this point, and it's also the emphasis of this paper. According to several studies, the concept of vulnerability does not yet have a commonly accepted definition (Berdica, 2002; Einarsson, 2006; Taylor and D' Este, 2004). Many studies define the vulnerability of road network from the point of operability, serviceability and accessibility of a link/route/road network, but generally speaking, the vulnerability can be seen in the same way as risk (Sarawitz, 2003), and vulnerability should be dissociated into two components: probability and consequence (Jenelius, 2006). Reducing vulnerability can hence be regarded as reducing the risks

involved in various incidents, which can be achieved in two ways: (1) reduce the probability of failure, or (2) reduce the resulting consequences when the failure occurs. In fact minimizing consequence is a very important aspect of vulnerability studies, since it is not always suitable to predict the probabilities of certain incidents, e.g. natural disaster or terrorist actions. It also accords with the need of this paper in reducing the consequences of road network failure and keeping the normal operation conditions of entire evacuation network when making emergency routing plans.

## VULNERABILITY EVALUATION MODEL

In this section, the vulnerability evaluation model of the road network element is presented. The model is to measure the importance of a link to the connectivity of an origin-destination pair under the context of emergency evacuation, and it's also a way of measuring link failure risk. Previous studies used such measures as guidance to routine road administrations or network maintenance; however, in emergency situations it seems to be more significant. It could provide useful assistance for emergency managers to know the operating situation of current road network, and make corresponding evacuation strategies to realize the purpose of reducing the risk of road network operation and improving the efficiency of whole evacuation.

The link vulnerability explicitly accounts for flow on the disrupted link and the availability of alternate paths, such as travel time and excess capacity. Link is a component of path, and a link may belong to many travel paths, therefore the flow on the link often from many O-D pairs. When the link is disrupted, it will need other alternative paths to accommodate the affected flow, and how to allocate the excess

capacity of alternative path to many O-D pairs is a key issue. Previous study, e.g. Murray-Tuite (2003), proposed that "no prioritization among the O-D pairs is permitted". It means that all roads are equally significant, regardless how often they are utilized. This point reflects the public characteristic of roads, and people should have equal opportunities everywhere. However, in the process of actual evacuation, with the considerations of emergency evacuation requirement, the travels on some O-D pairs, such as the evacuation travel between the bus distributing center and the refuges, are regarded to be more important. These O-D pairs are often main emergency passages or accommodate more traffic, in order to ensure the evacuation efficiency emergency managers often allocate these O-D pairs a higher traffic diversion priority. In the process of real evacuation, it shows as providing signal priority for these roads or implementing behavior adjustment management. If neglecting the phenomenon of traffic diversion priority, it may result in that the vulnerability of some links on the main emergency passage is too high, e.g. the main evacuation passages or rescue passages, or too low, e.g. the links on the routine travel paths. In this way it couldn't provide objective references for the emergency managers, so this paper attends to put emphasis on the vulnerability evaluation issues with traffic diversion priority. The mathematical formulation of the vulnerability index can find its predecessor from the seminal work of Murray-Tuite (2003). The model proposed in the following advances the state of the art by incorporating the generalized diversion priority weights that intuitively devise the mechanism for reassigning the affected traffic flow on disrupted link equally.

According to the interdependent relationship between path and link, for a given link $i$, when it is damaged, the affected traffic flow is the sum of the flow on the paths containing this link.

$$X_i = \sum_{r,s} \sum_{j} \Psi_{ij} f_j^{r,s} \tag{1}$$

$X_i$ is the amount of flow that would have to be accommodated by excess capacity on alternate paths, $\Psi_{ij}$ is the arc-path incidence matrix, and $f_j^{r,s}$ is the flow on path $j$ from origin $r$ to destination $s$. Obviously these alternate paths cannot contain link $i$. Let $b_j$ be the bottleneck link of path $j$, where the bottleneck is defined as the link with the minimum excess capacity, so the excess capacity $C_j$ of alternative path $j$ is the excess capacity of its bottleneck link. It is calculated as the difference of the link maximum service flow rate $\rho_{b_j}$ and the current flow $x_{b_j}$. Let $L_j$ be the set of links in alternative path $j$.

$$C_j = \min_{l \in L_j} c_l = c_{b_j} = \rho_{b_j} - x_{b_j} \tag{2}$$

With the considerations of allocating the path excess capacity to each O-D pair and providing traffic diversion priority for some O-D pairs, it is known that when the flow on the disrupted link $i$ contains the traffic flow of O-D pair $(r,s)$, the excess capacity $C_j^{r,s}$ of alternative path $j$ available to the O-D pair $(r,s)$ is given in equation (3).

$$C_j^{r,s} = \min_{l \in L_j} c_l \cdot \frac{X_i^{r,s} \cdot w_{r,s}}{\sum_{r,s} X_i^{r,s} \cdot w_{r,s}} = (\rho_{b_j} - x_{b_j}) \cdot \frac{X_i^{r,s} \cdot w_{r,s}}{\sum_{r,s} X_i^{r,s} \cdot w_{r,s}} \tag{3}$$

Where $X_i^{r,s}$ is the flow on the disrupted link $i$ from origin $r$ to destination $s$, and $w_{r,s}$ is the diversion weight of O-D pair $(r,s)$. Here in order to reflect relative importance of different O-D pairs, a weight $w$ is allocated for each O-D pair.

Actually $w$ not only depends on the evacuation demand that the O-D pair accommodates, but also relates to the particularity of the travel demand, e.g., some O-D pairs are main evacuation passages or rescue passages, therefore the diversion weight of each O-D pair can be shown as follows:

$$W_{r,s} = \frac{D_{r,s}}{\sum_{r,s} D_{r,s}} \cdot \alpha^a \cdot \beta^b \tag{4}$$

$$w_{r,s} = \frac{W_{r,s}}{\sum_{r,s} W_{r,s}} \tag{5}$$

$$\sum_{r,s} w_{r,s} = 1 \tag{6}$$

Where $W_{r,s}$ is the initial diversion weight of O-D pair $(r,s)$; $D_{r,s}$ is the travel demand between O-D pair $(r,s)$; $\alpha$ and $\beta$ are importance indices of special evacuation passages or rescue passages relative to the routine O-D pairs, e.g. $\alpha = 2$ means when the O-D pair $(r,s)$ is the special evacuation passage, its importance is two times of the routine O-D pair; $a$ and $b$ are both binary indicators, when the O-D pair is a special evacuation passage, $a = 1$; and when the O-D pair is a rescue passage, $b = 1$, and 0 otherwise. The diversion weight $W_{r,s}$ of each O-D pair should be normalization processed so as to ensure the sum of the weight is equal to 1.

In the actual travel, the number of available alternate paths between O-D pair is limited, thus when the link $i$ between O-D pair is disrupted, the maximum number of available alternate paths is assumed to $N_i^{r,s}$. In general, the number of alternate paths considered depends on the affected traffic flow $X_i^{r,s}$ and the available excess path capacity of each alternative path $C_j^{r,s}$. Obviously, when considering the traffic diversion priority of O-D pairs, the traffic assignment of the affected flow $X_i^{r,s}$

among alternative paths is following the thought of system optimal traffic assignment. Alternate paths are considered in order of marginal path travel time, obviously the alternative path with lowest marginal travel time will be taken first.

If the traffic assignment result shows $X_i^{r,s}$ can be accommodated by alternate paths, that is $n_i^{r,s} \le N_i^{r,s}$, where $n_i^{r,s}$ is the real number of available alternate paths, the relative usefulness of alternative path *j* for the given O-D pair $(r,s)$ can be treated as two parts: "How many traffic flow it could accommodate for me" and "How about the current situation of the alternative path". In other words, the utility of the alternate path is a function of the excess capacity of alternative path *j* and its current path travel time. As the excess capacity of alternative path is mainly related to its bottleneck link, the utility $U_j^{r,s}$ is shown in equation (7).

$$U_j^{r,s} = \frac{C_{b_j}^{r,s}}{C_{b_j}} \cdot \frac{T_{b_j}^f}{t_{b_j}} \tag{7}$$

As mentioned above, the denominator and numerator of the first term of the right hand side are respectively the excess capacity of bottleneck link $b_j$, $C_{b_j}$, and available excess capacity for the O-D pair $(r,s)$ $C_{b_j}^{r,s}$. $C_{b_j}^{r,s} \le C_{b_j}$, so the value scope of the first term is [0, 1]. The first tem indicates that as the excess capacity allocated to O-D pair $(r,s)$ increase, the feasibility of alternate path for the O-D pair $(r,s)$ increases. If available excess capacity for the O-D pair $(r,s)$ equals to zero, it means the alternative path is not a viable alternative. The denominator and numerator of the second term of the right hand side are the current travel time $t_{b_j}$ and free flow travel time $T_{b_j}^f$ of bottleneck link $b_j$, and on this basis to describe the current state and

baseline characteristics of path $j$. $T_{b_j}^f$ couldn't equal to zero, so the value scope of the second term is (0, 1]. Thus the utility of alternate path $j$ is bounded by 0 and 1, inclusive.

Correspondingly, the vulnerability index $V_i^{r,s}$ of link $i$ can be expressed as:

$$V_i^{r,s} = 1 - \sum_{j=1}^{n_i^{r,s}} U_j^{r,s} \cdot \frac{X_{i,j}^{r,s}}{X_i^{r,s}} \tag{8}$$

Where $X_{i,j}^{r,s} / X_i^{r,s}$ is the ratio of the flow allocated to alternative path $j$ and the part belonging to O-D pair $(r,s)$ on the disrupted link.

On the contrary, if the traffic assignment result shows that $X_i^{r,s}$ can't be accommodated by other alternative paths, that is $n_i^{r,s} > N_i^{r,s}$, or there is no available alternate paths; it indicates that link $i$ is critical to the connectivity or normal operation of O-D pair $(r,s)$, and its vulnerability $V_i^{r,s}$ takes the maximum value of 1.0, so the vulnerability index $V_i^{r,s}$ of disrupted link $i$ can be given in equation (9).

$$V_i^{r,s} = \begin{cases} 1 - \sum_{j=1}^{n_i^{r,s}} U_j^{r,s} \cdot \frac{X_{i,j}^{r,s}}{X_i^{r,s}} & n_i^{r,s} \le N_i^{r,s} \\ 1.0 & \text{otherwise} \end{cases} \tag{9}$$

Incorporating the vulnerability index of each O-D pair, the vulnerability of link $i$ over the whole road network $V_i$ is defined in equation (10).

$$V_i = \sum_{r,s} V_i^{r,s} \cdot \frac{X_i^{r,s}}{X_i} \tag{10}$$

It could be found that vulnerability not only describes the importance of link to the connectivity of road network, but also presents the link failure risk. A value of 0 indicates that under the current evacuation conditions damaging the given link would have no impact on the traffic. A value at the upper bund indicates that the failure risk

of this link is top high, once it is disrupted or damaged, there is no alternative path to accommodate this flow, and it will result in the inability of the network to transmit the origin-destination demand.

## EVACUATION ROUTING MODELING

Based on the identification of road link vulnerability in current evacuation network, how to make a reasonable and effective evacuation routing plan become a core problem. The traditional traffic assignment model based the system optimal thought takes the travel time minimization as the object function. However, under the emergency conditions, minimizing the travel time is only one of evaluation indices that the users care about. "Whether the link could maintain the normal transportation function" and "how about the link failure risk" are both core issues for the emergency managers and the evacuees. As mentioned above, link failure risk assessment also takes into account the travel time of bottleneck link, thus in this paper the classical SO traffic assignment model is modified by integrating the link vulnerability index, the model is explained as follows:

$$\min z = \sum_i x_i \cdot V_i \tag{11}$$

$$s.t. \quad \sum_j f_j^{r,s} = D_{r,s} \qquad \forall r,s \tag{12}$$

$$f_j^{r,s} \geq 0 \qquad \forall j,r,s \tag{13}$$

$$x_i = \Psi_{ij} \cdot f_j \tag{14}$$

Compare to previous function form with the objective of minimizing the travel time of whole network, equation (11) minimizes the travel risk of all users in the same network, $x_i$ is the flow on link $i$, and $V_i$ is the vulnerability index of link $i$. Constraint

(12) ensures that the sum of the flows $f_j^{r,s}$ on the paths connecting $r$ and $s$ meet the demand $D_{r,s}$ for the origin-destination pair $(r,s)$. Equation (13) is the non-negativity constraint. The final equation (14) relates the link and path flows through an arc-path incidence matrix $\Psi_{ij}$; the values of the entire matrix are 0 if link $i$ does not lie on path $j$ and 1 if link $i$ does lie on path $j$.

According to the design of objective function, the traffic assignment will be realized following the principle of reducing the travel risk of whole network. In order to compare the evacuation situations of the classical and the modified traffic assignment modes, a simulation experiment is given in the following.

## NUMERICAL STUDIES

### Simulation Network

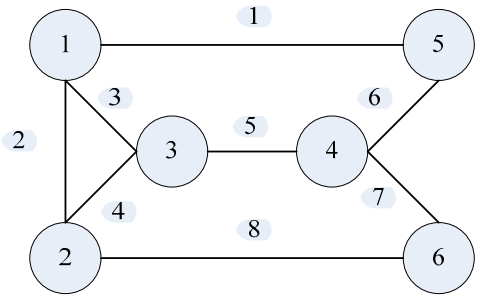



**Figure 1 Simulation network**

**Table 1 Link characteristics for the real network**

| Link | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Number of Lanes | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 2 |
| Free Flow Travel Time (min) | 10 | 2 | 2 | 1.5 | 6 | 2 | 3 | 11 |
| Max Service Rate (vehicle/h) | 200 | 170 | 160 | 170 | 180 | 180 | 180 | 200 |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

The paper chooses the road network of Murray-Tuite (2003) as the simulation reference object. As illustrated in Figure 1, the test network consists of 6 nodes, 20 links, and 4 OD pairs: (1, 5), (1, 6), (2, 5), (2, 6). Table 1 shows link characteristics.

**Table 2 Link characteristics for the simulation network**

| Link | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Length (grid) | 47 | 33 | 14 | 14 | 15 | 14 | 14 | 47 |
| Free Flow Travel Time ($s'$) | 16 | 11 | 5 | 5 | 5 | 5 | 5 | 16 |
| Max Volume (cars) | 110 | 82 | 30 | 30 | 35 | 30 | 30 | 110 |

Note: since the timing mode of software is different from the routine, e.g. second, it is defined as $s'$ in this paper.

The traffic simulation tool is Starlogo, developed by MIT. The length and volume of each simulation link could be given as the design of experiment road network, and the simulation value of each parameter is shown in table 2.

**Simulation Designing.** Based on the capacity limitation of the simulation network, it is assumed that the volume of the simulation interface is 200. Taking the evacuation demand of 1000 for example, it is presented in the way of continuously loading. For illustration purposes, three different loading levels are examined: under-saturation, saturation, and over-saturation. Volume of 200 is taken as the reference with the fluctuation of 20%, thus the volume of simulation network includes three levels: 160, 200, and 240. The proportion of travel demand among each O-D pair is assumed to follow the way of $D_{1,2} : D_{1,6} : D_{5,2} : D_{5,6} = 6 : 2 : 2 : 5$. During the process of simulation experiment, there are some basic steps as follows.

Step 1: Initialization. Take the link free flow travel time $t_f$ as the reference, and apply the classical system optimal traffic assignment model and Matlab software to calculate the link initial flow $x_i^0$;

Step 2: Based on the traffic assignment results $x_i^0$ in step 1, calculate the link travel time $t_i^0$ by using the BPR function $t_i = t_f + 0.15(\frac{x_i}{C_i})^4$; and then calculate the link vulnerability index $V_i^0$ following the equations (1) ~ (10); O-D pair (1, 5) is assumed to be the main evacuation passage, so $\alpha = 2$, and O-D pair (2, 5) is the main rescue passage, $\beta = 3$;

Step 3: Apply the model (11) ~ (14) to calculate the new link flow $x_i^1$ in Matlab;

Step 4: Return to step 2 and step 3 again, let $n = 1$, if $\sum_i \left| \frac{x_i^{n+1} - x_i^n}{x_i^n} \right| < \xi$, $\xi > 0$, the iteration stops, obtain the optimal solution and turn to step 5; otherwise, let $n = n + 1$, return to step 2 and come to the next iteration;

Step 5: Apply the traffic assignment result in step 4 and the initial traffic assignment result in step 1 to make the simulation experiment and comparative analysis, and the sequence of intermediate nodes in the trip chain is followed by the "cars" in Starlogo.

**Simulation Results Analysis.** During the simulation process, we will record corresponding information, such as average travel time, average travel distance, and network clearance duration, etc. Table 3 presents the simulation results, where the unit of numerical value is the system unit of the simulation software.

**Table3 Simulation results analysis of the classical system optimal traffic assignment model**

| Simulation Network Volume (cars) | Average Travel Time ($s'$) | Average Travel Distance (grid) | Network Clearance Duration ($s'$) | The Sum of Link Failure Risk |
|---|---|---|---|---|
| 160 | 22.5 | 49.8 | 156.4 | 241.9 |
| 200 | 23.9 | 51.4 | 161.8 | 302.4 |
| 240 | 25.3 | 50.5 | 177.3 | 392.7 |

**Table 4 Simulation results analysis of the modified system optimal traffic assignment model**

| Simulation Network Volume (cars) | Average Travel Time ($s'$) | Average Travel Distance (grid) | Network Clearance Duration ($s'$) | The Sum of Link Failure Risk |
|---|---|---|---|---|
| 160 | 23.3 | 50.2 | 173.2 | 187.2 |
| 200 | 25.1 | 50.7 | 179.1 | 249.6 |
| 240 | 27.9 | 51.1 | 195.6 | 311.8 |

According to the simulation results of table 3 and table 4, it could be found that both in the classical and modified traffic assignment experiments, when the capacity of the road network was over-saturated (V = 240 cars), the average travel time of simulation vehicles and the network clearance duration would increase sharply. It led to the conclusion that in order to ensure the evacuation operate in a high efficiency, it was important to keep the whole road network in a saturated or under-saturated state. In this way, the road resources could be utilized to the great extent, and what's more, the efficiency, stability and safety of the evacuation could be guaranteed, which was also the important theoretical basis for implementation of staged or regional evacuation management strategies. In addition, due to all the evacuation activities were carrying out under a unified command of emergency management department, so the average travel distance presented an approximate numerical result in various traffic conditions, which was consistent with the actual evacuation.

On the other hand, in a certain network capacity, when comparing the average travel time, average travel distance, and network clearance duration in the classical and modified simulation scenarios, the latter always showed a relative high value. For instance, in the saturated case, the average travel time of the latter was 5% more than the former. While in the over saturated conditions, the increase of the latter was even up to 10.3%. The network clearance duration of the whole network also changed a little, the increase proportions in three volumes were all about 10%. This was mainly because the objective of the modified traffic assignment mode was to minimize the system travel risk, while the links with a smaller vulnerability index were not exactly the links having the shortest travel time. However, when calculating the sum of link failure risk in the two simulation scenarios, it could be found that the failure risk of the latter under three network volumes were all presenting a trend of significantly decrease, the decreasing range reached 22.61%, 17.46% and 20.60% respectively. In the actual evacuation process, even though a 10% decrease may bring a very positive consequence, so it is indispensable to take into account the failure risk of the evacuation road network in emergency evacuation routing planning.

## CONCLUSIONS

Under the conditions of emergency evacuation, assessing the vulnerability of the road network and identifying the link failure risk are important part of the emergency evacuation routing planning. In order to highlight the diversion priority of some origin-destination pairs in the evacuation network, a vulnerability evaluation model for the road network element was improved based on the previous studies by

introducing the concept of traffic diversion priority. Different O-D pairs are assigned with different diversion priorities. The practical significance of the concept is to provide priority for the main emergency evacuation or rescue passages, so as to make up the defect of a too high vulnerability value of the links on the special evacuation passages or a lower evaluation value of the links on the routine paths in the previous evaluation. In this way it could be helpful to realize the target of accurately describing the failure risk of the emergency evacuation road network.

On this basis, a modified evacuation traffic assignment model with the objective of minimizing the system travel risk of the whole network was built. The purpose of modeling was to assign the evacuation demand from the angle of maintaining the normal transport function of network system. That was taking the safety and stability of the evacuation as the first considerable factor. For comparison purposes, the evacuation situations of the classical and modified traffic assignment modes were solved by Matlab, and simulated in the Starlogo developed by MIT. The simulation experiments under various traffic conditions were carried out with the presupposed evacuation demand, road network capacity and evacuation environment. The simulation results demonstrated that the modified system optimal traffic assignment model would lead to a slight increase in the average travel time and the network clearance duration, however, the sum of link failure risk in this scenario would sharply decrease about 20%, comparing to the classical model. If the evacuation routing planning did not consider the link vulnerability and overlook the link failure

risk, it might result in great losses. The achievements of the research provide a useful reference for the further study of evacuation routing planning under emergency conditions.

## ACKNOWLEDGEMENTS

This research was supported by National Natural Science Foundation of China (Project No. 70973032) and National Soft Science Research Project (Project No. 2009GXS5D130)

## REFERENCES

Bell, M. G. H. (2000). "A game theory approach to measuring the performance reliability of transport networks." *Transportation Research Part B*, 34(6): 533-545

Berdica, K. (2002). "An introduction to road vulnerability: what has been done, is done and should be done." *Transport Policy*, 2002, 9(2): 117-127

Chen, A., Yang, C., Kongsomsaksakul S., Lee, M. (2006). "Network-based accessibility measures for vulnerability analysis of degradable transportation networks." *Networks and Spatial Economics*, 7(3):241-256

Einarsson, S., Rausand, M. (2006). "An approach to vulnerability analysis of complex industrial systems." *Risk Analysis*, 18(5), 535-546

Jenelius, E. (2009). "Network structure and travel patterns: explaining the geographical disparities of road network vulnerability." *Journal of Transport Geography*, 17(3): 234-244

Jenelius, E., Petersen, T., Mattsson, L. G. (2006). "Importance and exposure in road network vulnerability analysis". *Transportation Research Part A*, 40(7): 537-560

Murray-Tuite, P. M. (2003). "Identification of vulnerable transportation infrastructure and household decision making under emergency evacuation conditions." *Doctoral Degree Thesis,* The University of Texas at Austin

Sarawitz, D., Pielke, J. R., Keykhah, M. (2003). "Vulnerability and risk: some thoughts from a political and policy perspective." *Risk Analysis*, 23(4): 805-810

Sullivan, J. L., Novak, D. C., Aultman-Hall, L., Scott, D. M. (2010). "Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: a link-based capacity-reduction approach." *Transportation Research Part A*, 44(5): 323-336

Taylor, M. A. P., D'Este, G. M., Nicholson, A. (2004). "Critical infrastructure and transport network vulnerability: developing a method for diagnosis and assessment." *Proceedings of the Second International Symposium on Transportation Network Reliability*, Christchurch, New Zealand, July, 96-102

# Enhancing the value of an incidents database with an interactive visualization tool

Catherine Morency[1], Martin Trépanier[2], Nicolas Saunier[3] and Jean-Philippe Akélaguélo[4]

[1]Assistant professor, Department of civil, geological and mining engineering (cmorency@polymtl.ca), [2]Associate professor, Department of mathematics and industrial engineering (mtrepanier@polymtl.ca), [3]Assistant professor, Department of civil, geological and mining engineering (nicolas.saunier@polymtl.ca), [4]Civil engineering student, Ecole Polytechnique of Montreal, P.O. Box 6079, Station Centre-Ville, Montreal, PQ, CANADA, H3C 3A7

## ABSTRACT

The growing capabilities of computational tools and the advent of multiple data providing systems have increased the needs and ability to process large sets of microscopic data. Since the data are not directly suited for analysis or modeling, they are often stocked without further question, while they could shine a new light on critical traffic phenomena such as congestion.

This research is a spin-off of a project on the modelling of travel time reliability on highway corridors. The project relies on floating cars data but has also led to the exploration of historical incidents datasets that required various pre-processing steps. A visualization tool was developed using Microsoft Excel in an innovative way. This paper reports on the development of this tool that provides some insight into the analytical potentialities of the incident databases. Four years of data are examined and converted into usable information through cross-examination and interactivity.

## INTRODUCTION

The increasing capacity of technological tools, as well as the advent of multiple data providing systems, has multiplied the needs and ability to process large sets of microscopic data. At the same time, expectations towards infrastructure managers are increasing both from the public and decision makers. On the one hand, travelers are seeking more detailed and up to date information on the various transportation networks while decision makers ask for continuous assessments of the quality and reliability of the infrastructures and services they are responsible for. Data are being gathered, in various formats, at different scales, in multiple locations and at varying time frames. Depending on their origins they can be ready for analysis and modeling or simply unusable in their current form. Often, multiple sets of data are stocked without further analysis, while they could shine a new light on critical traffic indicators such as congestion or travel time reliability.

In this context, tools to enhance the value of datasets, of various forms, are welcome contributions both for analysts and decision makers. The tool that is described in this paper is a spin-off of a project on the modeling of travel time

reliability on the main highway corridors of the Greater Montreal Area. While the main focus of the project was on floating cars data (probe vehicles), other peripheral datasets were made available since they were thought to increase understanding of the congestion phenomena. Amongst these were historical incidents datasets. Since these datasets had not been examined previously, pre-processing was required (geo-codification, classification of incidents). Also, a visualization tool was developed using Microsoft Excel spreadsheet software in an innovative way. This paper reports on the development of the interactive visualization tool that provides some insight into the analytical potentialities of the incident databases. The paper first proposes some background info on the travel time reliability project as well as on the role and relevance of visualization tools to assist analysis and dissemination of results. It then describes the structure of the tool as well as the development process. Some insights into the potentialities of incidents datasets to contribute to the modeling of congestion in urban areas are then proposed.

## BACKGROUND

### Project context

In 2007, our research group was mandated by the Quebec Ministry of Transportation (MTQ) to assess the challenges and potentials of travel time's data to monitor the reliability of the Montréal highway network. At the core of the project was the processing of 29,229 measurements from 6 years of floating car data, covering some 800 km of freeways. This was aimed to model the frequency distribution of travel times on road segments and develop reliability indicators (see Loustau et al., 2010 a, b and c). Clustering techniques were used to classify segments according to the similarity of the frequency distributions of travel times. This process led to the identification of 16 clusters (8 for each peak period) and distributions were modeled using a three lognormal distribution approach.

### Modeling context

During the modeling process, various data requests were formulated. While the current models on the frequency distribution of travel times provide good results through simulation, we are looking for explanatory variables to explain the distribution patterns. We also want to apply the models where travel time data are not available as well as for forecasting purposes. Our ultimate goal is to integrate three types of information to enhance the explanatory power of the model:

- Travel times and speed data from various types of technology: floating cars, GPS traces Bluetooth devices, video data, etc.
- Incident data recorded by the MTQ, such as accidents, vehicle breakdowns and other sources of traffic interruptions
- Urban features data describing the geometry of the road segments, the features of the neighborhood as well as the surrounding travel conditions

The research team first got access to 5 years of incident data, files that had hardly been used. To borrow Olsen (1999) data use cycle, data was in a stage of

chaos (raw, unorganized data). Wongsuphasawat et al. (2009) make similar observations about many state departments of transportation. Unification and codification (second stage) was required to convert these data into data usable for analysis and modeling (stage 3).

### The role of visualization in the analysis and decision making process

According to Buliung and Morency (2009), visualization can be compared to an anamorphosis interpreter wherein the act of visualization makes use of specialized devices (e.g., computer programs, statistical tools, GIS, interactive spreadsheets), or compels the viewer to occupy a specific perspective (spatial, temporal, or social feature), with a view to reconstituting the "original" for the purpose of developing a clearer understanding of "process." Actually, visualization is relevant at all stages of the transportation planning process, from data acquisition to data dissemination (Manore et al., 2008). Early interest in visualization in transportation revolved around "creating 3D geometry, images and animations to better communicate how proposed projects will "look"" (Visualization in Transportation Committee web site). Hughes (2004) insists on the need to engage the public and other stakeholders and get feedback throughout the design process, beyond the presentation of engineering facts. The boundary between geographical information systems (GIS) and other visualization tools is quickly disappearing as "visualization is migrating from conceptual illustration into geospatially accurate decision aids and tools" (Hughes 2005). Multimedia dissemination of origin-destination survey data has been initiated by Chapleau et al. (1997). Several web-based visualization tools have been developed at the Center for Advanced Transportation Technology Laboratory (CATT LAB) of the University of Maryland, among which some allow the visualization of incidents (Wongsuphasawat et al. 2009) and try to identify correlations between incidents and congestion (Lund et al. 2010). In a more general framework, MacEachren's cube (MacEachren, 1995) details the multiple roles of visualization according to the audience (public vs. private/technical), the objective (presenting known facts vs. discovering patterns), and the level of interactivity (high vs. low). Depending on how these three aspects are combined, the use of data visualization tools can vary from communication to scientific analysis. In our project, the visualization tool aims at assisting the descriptive analysis of the raw dataset and at obtaining a better understanding of its content.

## METHODOLOGY

Data processing leading to the construction of the Excel spreadsheet is shown at Figure 1. First, the incidents from the raw datasets are geocoded by using the incident description, road number and nearest intersection. The GIS layer comes from the highway network. The same network has been used to estimate travel times from floating cars data. Therefore, the incidents are matched with the road segments that were probed. Then, incident data is prepared: incidents are classified by type, and most of the indicators presented in the next section are calculated. Finally, data is stored in hidden Excel sheets within a regular Excel spreadsheet. Visual Basic for Applications (VBA) macros and cell formulas are used to format and display data in

several interactive sheets that are presented hereafter. Key figures regarding the dataset that was made available to us are summarized in the Table 1.

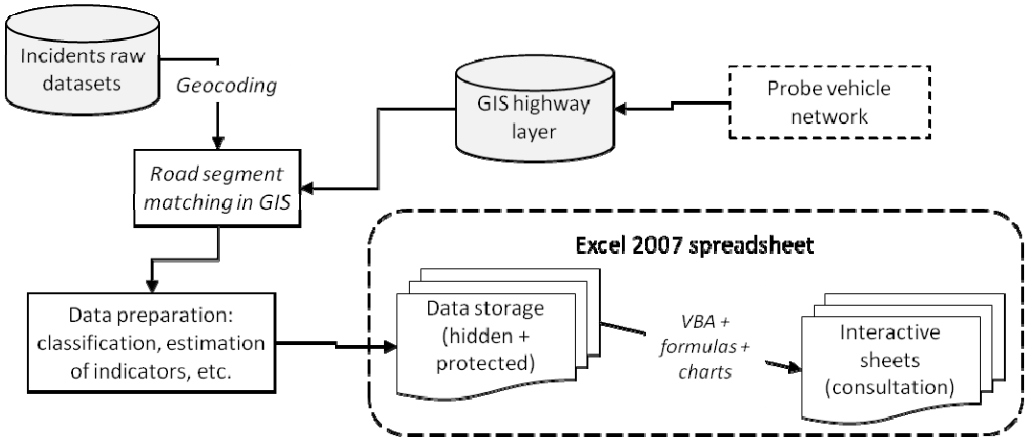



**Figure 1. Data processing and spreadsheet construction**

**Table 1. Key figures of the available dataset**

| | |
|---|---|
| Temporal coverage | July 23rd, 2001 to December 31st, 2004 |
| | 1,249 different days with incidents |
| Number of spatial locations | 359 different spatial locations (x-y coordinates) |
| Total number of incidents | 107,286 |
| Types of incidents | 60 different types |
| ***5 most frequent types of incidents and proportion of the total number of incidents*** | |
| *Failure* | 52.98% |
| *Accident* | 15.97% |
| *Debris* | 14.53% |
| *Car that left without assistance* | 3.12% |
| *Pedestrian* | 2.13% |

## KEY FEATURES OF THE TOOL

### Structure

The visualization tool was developed using straightforward functions of Excel and is composed of seven worksheets. Table 2 lists these worksheets, identify what features are used in it and briefly explain their contents and role. Each of the sheets is described hereafter.

### Study of incidents frequency and duration

The first two interactive sheets look similar but display two different types of information. On the first sheet (Figure 2), the user can make a selection of roads, and then use comboboxes to choose the type of incident, the year and the period. The map displays the frequency of incidents for the selection, showing a different color for each road and direction. The legend is automatically adjusted to the observations. The lower right part of the screen displays numerous statistics about the selection. With

this tool, it is easy to isolate the part of the network and the circumstances that we want to examine.

**Table 2. Spreadsheet structure and features**

| Worksheet | Excel features | Contents / Role |
|---|---|---|
| Main menu | Shape controls (buttons) | Navigation tool to other sheets. |
| Incident frequency | Shape controls, background bitmap, bubblechart, checkboxes | Presents the incident frequency with respect to many dimensions: location (spatially and by road), type of incident, time period during the day, year, week period. |
| Incident duration | Shape controls, background bitmap, bubblechart, checkboxes | Presents the incident duration with respect to many dimensions: location (spatially and by road), type of incident, time period during the day, year, week period. |
| Consultation of incidents by highway | Shape controls, line and column charts, listboxes | Presents the incident distribution by road stretch, by interchange crossing or section), type of incident, time period during the day, year, week period. |
| Multidimensional comparison tool | Shape controls, comboboxes, pie charts, spreadsheet tables | Provides a tool for comparison of two road sections, for two different time periods. In addition, user can select a specific type of incident and a specific period of the day. Results are displayed in tables and pie charts. |
| Animated evolution of a single highway | Shape controls, bubblechart, listboxes, comboboxes | Displays, for a single road, the distribution of incidents (or incident rate per km) according to the type of incident, the duration class and the day of week. |
| Animated evolution of multiple highways | Shape controls, line chart, checkboxes, comboboxes | Displays, for a selection of roads, the distribution of incidents (or incidents rate per km) according to the type of incidents and the duration class. |

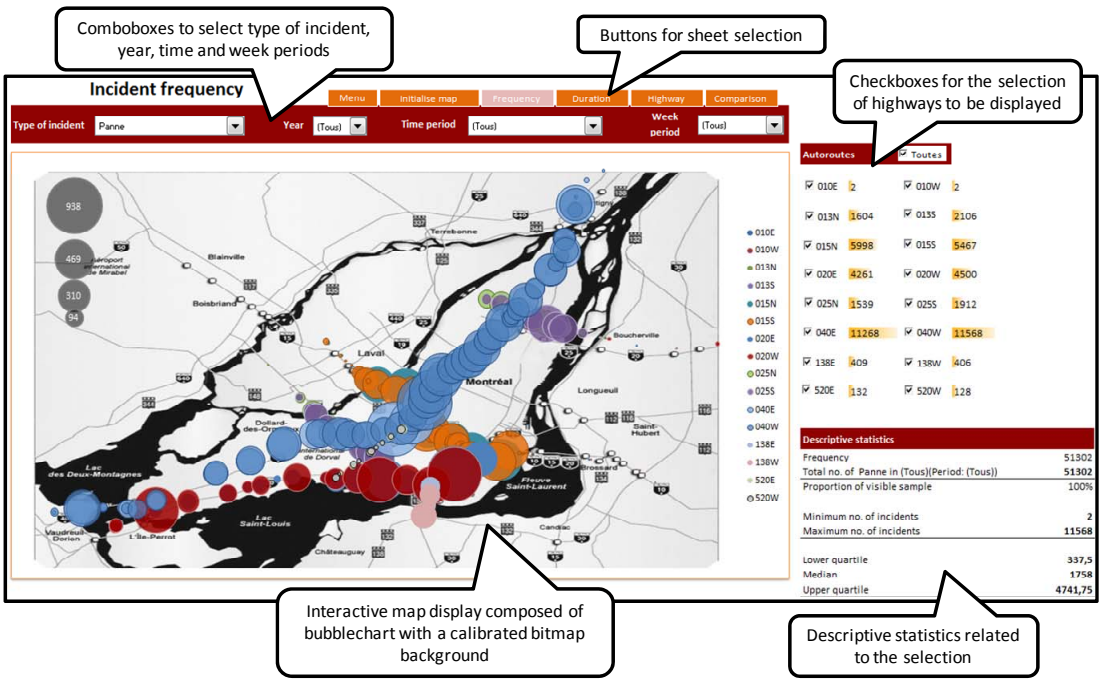



**Figure 2. The "frequency" worksheet**

The second sheet displays the average incident durations in the same manner. In Figure 3, we zoomed in on the bubblechart to show only two major freeway axes

of the region. We can see that the average duration of incidents is quite variable over the region. The most affected part is the "Turcot" interchange, linking several freeways near the downtown area. The bridges at the extremities of the Montreal Island are also experiencing higher incident durations. This figure is for vehicle failures only. Therefore, the analysts can browse through all types of incidents, like accidents, animal or trash removal, ice and snow removal, regular maintenance, power outage, etc.

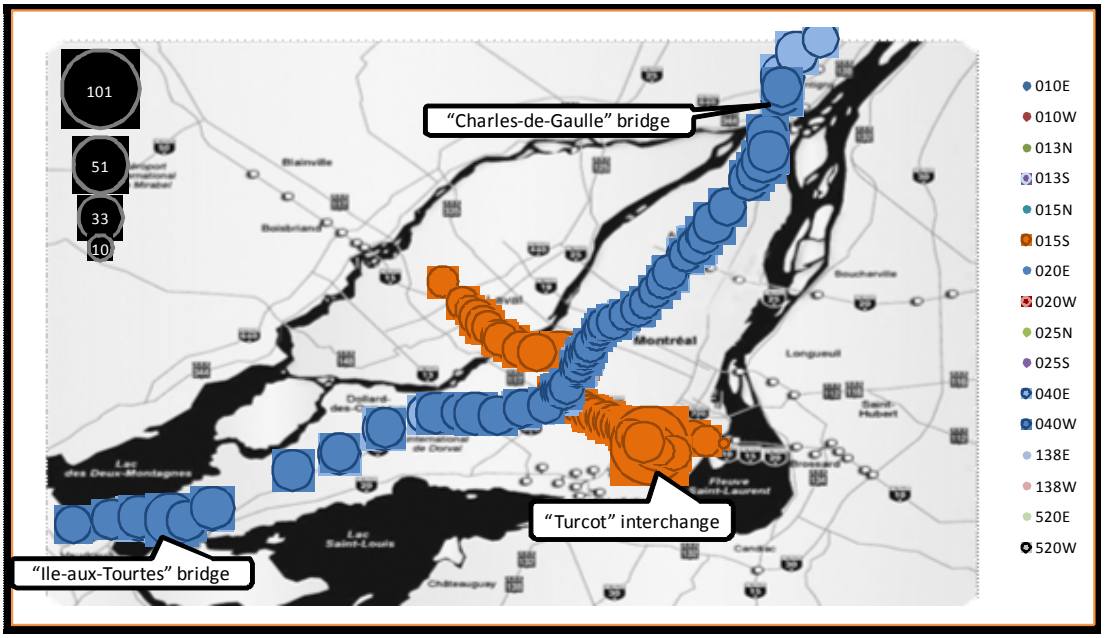



**Figure 3. Average duration of failure incidents over the whole period for two major freeway axes (in minutes)**

## Study of incidents by highway

The next sheet is aimed to display the number and the duration of incidents for a given highway, in both directions (Figure 4). The tool permits to filter this information according to the type of incident, the time period of the day (AM, PM peaks, midday and evening off-peak) and the week period (weekday of week-end). This very detailed examination of data can help to identify problematic sectors that would be hard to see on a map. It also helps to identify missing or erroneous data.

## Comparison tool

Figure 5 presents the comparison tool worksheet. Here, the user can select alternatively two highways with their own incident type, and two daily periods filtered by year. The results are displayed for each combination of these dimensions, in the form of charts and tables. With this tool, it is easier to see the differences between two different sub-sets of data, may it be in the same circumstances or not.

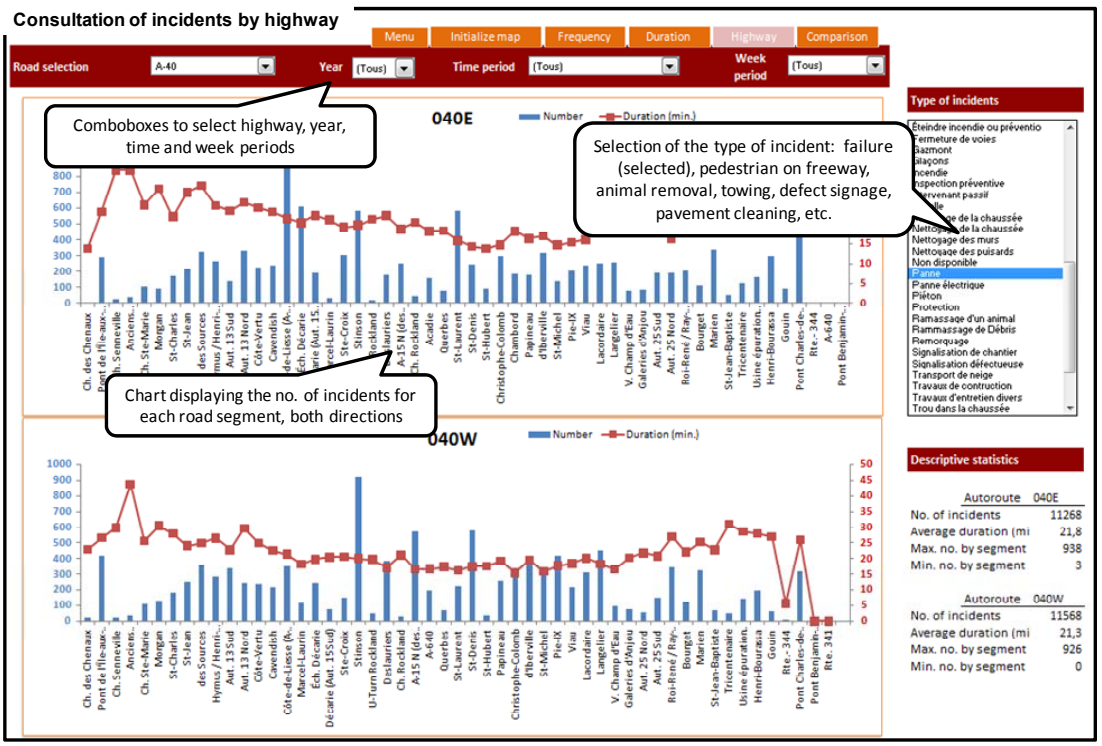



Figure 4. Worksheet for the consultation of incidents by highway

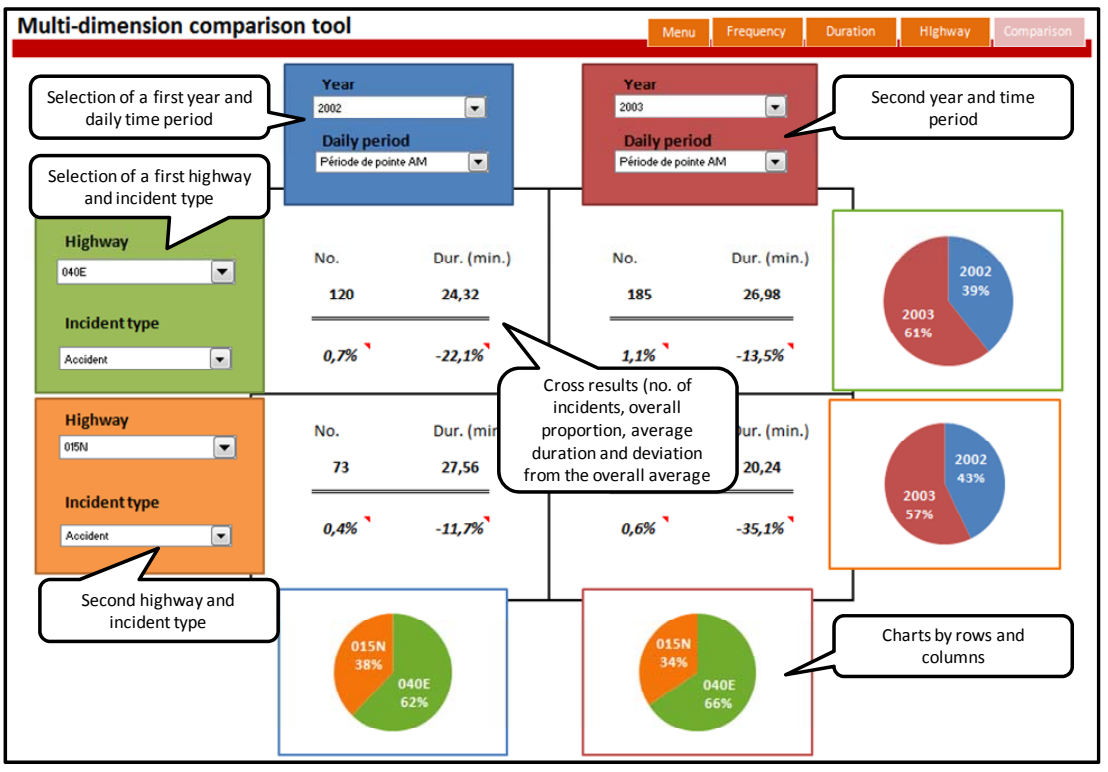



Figure 5: Multi-dimension comparison tool worksheet

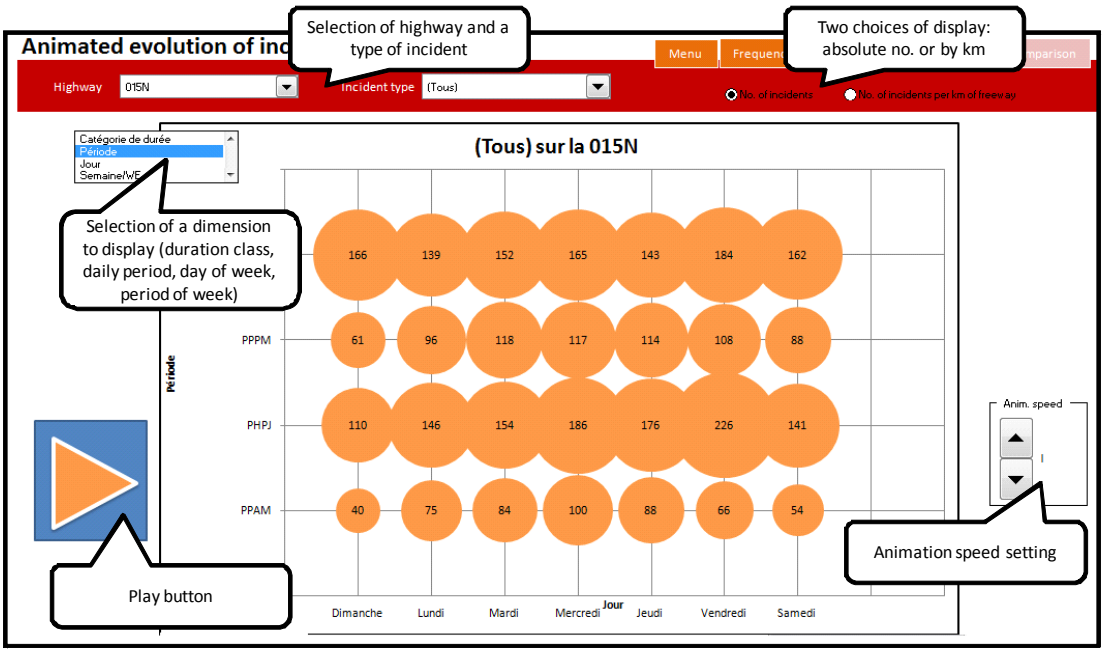



Figure 6. Worksheet for the animated evolution of incidents on a single highway

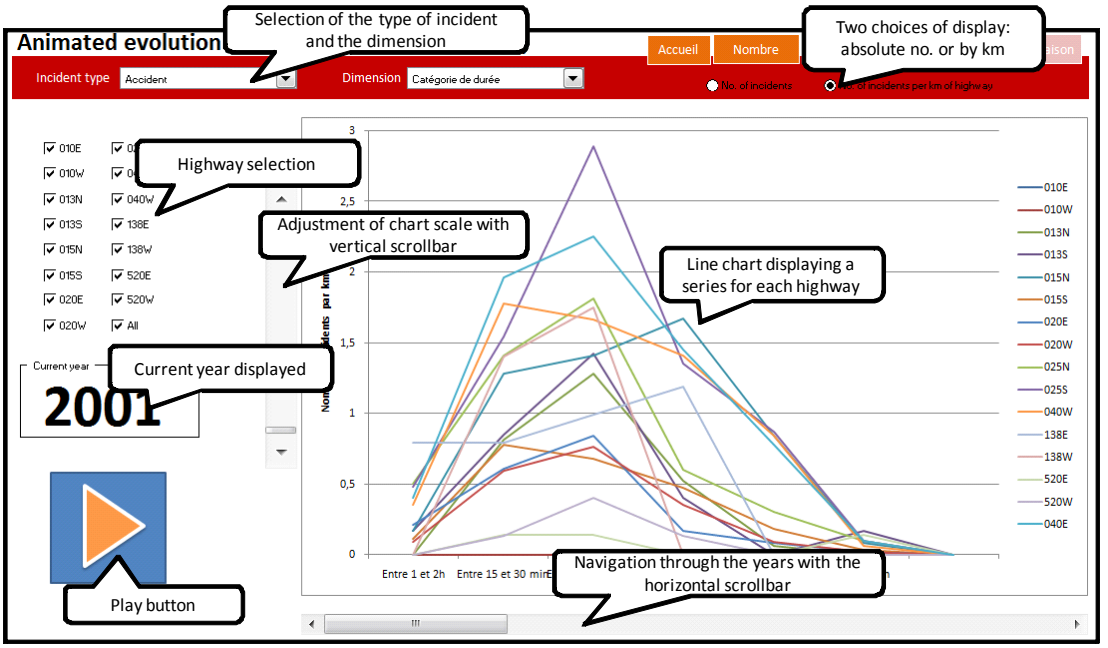



Figure 7. Worksheet for the animated evolution of incidents on multiple highways

## Animated evolution of incidents

While the characteristics of the incidents can be compared between two different highways, one may be interested to see the evolution of such characteristics over the years. The two last worksheets provide a way of showing this evolution with the help of chart animation. On the first screen (Figure 6), the user will select a specific highway and a type of incident. Then, after the choice of a time period (period, day of week, etc.), the bubblechart is animated to show the results from the

first to the last year of observation. This helps to identify the elements that could have changed differently through time. The second screen (Figure 7) provides a similar feature but for a selection of highways. The results are displayed on a line chart with multiple series, one for each highway. The user can interactively change the Y-scale of the chart with the help of the vertical scrollbar.

## CONCLUSION AND PERSPECTIVES

In this paper, we have presented an interactive tool, implemented in Microsoft Excel that was developed to enhance the visualization of data on incidents in the Greater Montreal Area. Data have been embedded in the spreadsheet software in an innovative way. Seven worksheets acts as the application software interface by providing multiple choices of parameters to the user. Therefore, results are displayed using tables and charts in the most interactive and intuitive way. The tool benefits from the conviviality of the spreadsheet software interface, the calculation power of its formulas and the interactivity of the Visual Basic for Application underlying programming language.

Let us remind that the tool is part of a larger project on the modeling of travel time on urban highways in the area. The use of this tool by transportation planners is essential to the thorough understanding of the incidents phenomenon in the area, hence to conduct a deeper analysis of the cause-to-effect relation between different types of incidents and congestion on highways. In a future research, we look forward to integrate these results in a data mining analysis combining travel time, incidents and road geometry databases to determine the underlying factors of congestion in urban areas.

## ACKNOWLEDGMENTS

The authors wish to acknowledge the support of the Quebec Ministry of Transportation for this project, namely the team of Pierre Tremblay.

## REFERENCES

Buliung, R., Morency, C. (2009). “Seeing is Believing": Exploring Opportunities for the Visualization of Spatiotemporal Dimensions of Activity-Travel and Land Use Processes, in Paez, A., Le Gallo, J., Buliung, R., Dall’Erba, S. Progress in Spatial Analysis: Theory and Computation, and Thematic Applications. pp.119-148.

Chapleau R., Trépanier M., Lavigueur P., Allard B., (1997). Origin-Destination Survey Data Dissemination in a Metropolitan Context: A Multimedia Experience. Transportation Research Record, no. 1551, pp. 26-36.

Chapleau, R., Morency, C., Bourgeois, M. (2008). Simple and Interactive Spatial Mobility Analysis Tool for Data Visualization, International Conference on Survey Methods in Transport: Harmonisation and data comparability, Annecy, France.

Hughes, R. (2004) Visualization in Transportation: Current Practice and Future Directions Transportation Research Record, no. 1899, pp. 167-174.

Hughes, R. (2005) Research Agenda for the Application of Visualization to Transportation Systems, Transportation Research Record, no. 1937, pp. 145-151.

Loustau, P., Grasset, V., Morency, C., Trépanier, M. (2010) Comparing floating car data and carsharing GPS data for travel time assessment, to be presented at the 12th World Conference on Transport Research, Lisbonne, Portugal.

Loustau, P., Morency, C., Trépanier, M. (2010). Measuring, Describing and Modeling Travel Time Reliability, presented at the 89th Annual Meeting of the Transportation Research Board, Washington D.C.

Loustau, P., Morency, C., Trépanier, M., Gourvil, L. (2010). Travel time reliability on a highway network: estimations using floating car data, Transportation Letters, Vol. 2(1), pp. 27-37, 2010

Lund, A., Pack, M. (2010) Dynamic Wide-Area Congestion and Incident Monitoring Using Probe Data, Transportation Research Record, in press.

MacEachren AM (1995) How maps work: representation, visualization, and design. The Guilford 828, Press, New York.

Manore, M., Krygsman, S. (2008), Workshop B5 – Data visualization techniques, Workshop Summary Report, International Conference on Survey Methods in Transport: Harmonisation and data comparability, Annecy, France.

Morency, C., Trépanier, M., Piché, D., Chapleau, R. (2008) Bridging the gap between complex data and decision makers: an example of innovative interactive tool, accepted for publication in the Journal of Transportation Planning and Technology.

Olsen, Dan. R. Jr. (1999). Interacting in Chaos, Interactions Journals, September + October 1999.

Visualization in Transportation Committee web site http://www.trbvis.org/MAIN/TRBVIS_HOME.html

Wongsuphasawat, K.; Pack, M.; Filippova, D.; VanDaniker, M. & Olea (2009) A. Visual Analytics for Transportation Incident Data Sets Transportation Research Record, no. 2138, pp. 135-145

# An Exploratory Study of Speed Limit Compliance in Missouri Work Zones

Mojtaba A. Mohammadi[1], Ghulam H. Bham[2], Hadi Khazraee[3]
1. Graduate Research Assistant. Civil, Architectural, and Environmental Engineering, Missouri University of Science and Technology, ma998@mail.mst.edu
2. PhD. Civil, Architectural, and Environmental Engineering, Missouri University of Science and Technology, ghbham@mst.edu
3. Graduate Research Assistant. Civil, Architectural, and Environmental Engineering, Missouri University of Science and Technology, sk387@mail.mst.edu

## ABSTRACT

This study evaluated the speed compliance with the posted speed limit both quantitatively and qualitatively using objective and subjective data analysis. Free flow speed data for three work zones on I-44 in Missouri were used for the objective evaluation. Speeds analyzed for cars and trucks were distributed normally. Construction activity had a significant effect in reducing the speed of vehicles. Results indicate that the speeds of vehicles were statistically higher than the speed limit regardless of the construction activity, when the lane width was not reduced. The subjective evaluation investigated the drivers' perception of driving through the work zone by using a survey questionnaire. About 40% of drivers indicated that the speed of other drivers were less than the speed limit which contradicted the results of the objective analysis. More than 70% of drivers experienced delay and more than 90% agreed that construction activity reduced their speed which was ascertained from the results of the objective analysis.

## INTRODUCTION

Many states have enacted temporary speed reduction regulations. Different speed limits are in place, however, regulatory and advisory speed limit signs alone do little to reduce traffic speeds in work zones (Graham-Migletz Enterprises, 1996; Migletz et al., 1999), and drivers do not follow speed limits that they consider unreasonable (Outcult, 2009; Agent et al. 2005). A 10-mph reduction below the normal speed limit is desirable in work zones when work takes place on or near the open lane (Migletz et al., 1998, Outkalt, 2009). Migletz et al. (1999) found that compliance with work zone speed limits was generally highest where the speed limit was not reduced and decreased where the speed limit was reduced by more than 10 mph. The most commonly used speed reductions are 0 mph, 5 mph, 10 mph, and 20 mph. These differences have made enforcement difficult, as motorists see different speed limits in work zones.

Traffic speeds adjacent to a work zone are an important safety concern. High speed provides drivers' shorter time to react, requires longer distance to come to a stop and provides less time to workers to get out of the way in case of work zone intrusion. A high speed vehicle is also more likely to penetrate a barrier and threaten workers. Ensuring safety in work zones while maintaining highway capacity has therefore become one of the most interesting challenges confronted by traffic

engineers and researchers. Experience has shown that the effectiveness of signs to reduce speed depend on the posted speed limit, the geometry of the work zone, and driver's sight distance (Outkalt, 2009). Speed reduction signs posted for work zones with little to no evidence of any activity are not only ineffective, but can also make drivers skeptical of the validity of signs posted at work zones. Research has shown that drivers reduce speeds in work zones, particularly when workers are present, regardless of whether or not speed limit reductions are posted (Migletz et al., 2005).

Agent et al. (2005) showed that travel speeds on most types of highways are substantially above the posted speed limit. In addition, they found that car speeds are slightly above those of trucks. Data taken before and after speed limit changes indicate that operating speeds change much less than the change in speed limit.

Research have been carried out on determining the appropriate speed limits for work zones (Migletz et al., 1999; Graham-Migletz Enterprises, 1996; Knodler et al., 2008) and on evaluation of speed control methods (Richards et al., 1985; Sandberg et al., 2006). In these studies, the speed parameters that were used as measures of effectiveness included mean and standard deviation of speeds, 85th percentile speed, speed limit compliance, speed variance, and percentage of drivers exceeding the posted speed limit by more than 5 mph, 10 mph and 15 mph. Many researchers used the FFSs as the data for their studies. Examples include Knodler et al. (2008) on effective speed management and setting rational speed limits, studies by Benekohal and Shu (1992), and Wang et al. (2003) on changeable message signs and speed reduction strategies.

The objectives of this study were to investigate traffic speeds and driver compliance with posted speed limits in work zone to asses objectively and subjectively the impact of static speed limits on work zones. This study also evaluated the impact of work zone activity on free-flow speed (FFS) of vehicles. Three work zone sites were selected, traffic was videotaped, and traffic flow characteristics were extracted. FFSs within the work zone activity area were compared with the speed limit. The effect of construction activity and lane width reduction on vehicles' speeds was also studied. For subjective evaluation, drivers were surveyed about safety of traveling through the work zone, driver's preference for work zone signage such as "work zone ahead", "left/right lane closed" and "speed limit" signs and their effect on lane change behavior, drivers' actual speed versus posted speed limit, driver's opinion about the speed limit compliance, perceived delay and the effects of construction activity on their driving behavior. The responses were then analyzed and evaluated in regard to the objective analysis.

## METHODOLOGY

Traffic data extracted were used for studying speed of cars and trucks. Various statistical analyses were performed; speeds were examined for normality using the Anderson-Darling test, and t-test was used to determine the significance of difference in the speeds of trucks and passenger cars and to evaluate the effects of construction activity on speeds. Speed compliance was analyzed based on the difference between the mean speed and the posted speed limit. A 90% confidence

level was used for all tests. This analysis identified statistically significant differences based on vehicle types and presence of construction activity.

The FFS data distributions were tested for normality. The null hypothesis, $H_0$, assumed that FFS distribution was normal and the alternative hypothesis, $H_A$, stated that FFS distribution was not normal. Two-sample t-test was performed to determine whether there was any significant difference between the mean value of the FFS of cars and trucks. The null hypothesis was $\mu_{FC} \leq \mu_{FT}$ where $\mu_{FC}$ represents the mean FFS of cars and $\mu_{FT}$ the mean FFS of trucks. The alternative hypothesis was $\mu_{FC} > \mu_{FT}$.

For work zones with construction activity, two-sample t-test was performed to evaluate the effect of construction activity on the speeds of cars and trucks. The null hypothesis was the difference between mean speed during periods of no activity and mean speed during periods of activity was less than or equal to zero. The alternative hypothesis, $H_A$, was that this difference was greater than zero.

To evaluate speed compliance, t-test was used to compare the mean FFS of cars and trucks with the posted speed limit of 60 mph. The null hypothesis for both cars and trucks was that the mean value of the speed distribution was 60 mph. The alternative hypothesis was that this value was greater than 60.

As a primary step for the subjective analysis, a survey questionnaire was distributed to drivers in at three different times and the results for these three times were combined. Drivers were asked if they had passed through the work zone; those who answered positively received a questionnaire. Some of the questions were multiple choice based on a Likert Scale from totally disagree to totally agree, and other types of multiple choice questions.

## DATA COLLECTION

The work zone locations were selected in consultation with MoDOT. All sites were located on I-44 in Missouri. They included two right-lane-closed work zones in Waynesville and Rolla, MO and one left-lane-closed work zone near Cuba, MO. The data collection locations were near work activity within the work zone. The regulatory speed limit of the work zones was 60 mph and the highway speed limit was 70 mph. The FFSs of passenger cars and heavy vehicles traveling through the work zone were evaluated to determine driver compliance with the speed limit.

For the work zone near Rolla, the open lane width was reduced because the work zone tubular markers were placed about a foot inside the open lane. All the videos were taken during daytime. Table 1 presents the different attributes of each work zone.

From the video data, free flowing vehicles were identified based on a 5-second headway (Benekohal and Shu, 1992; Bella, 2005; Wang et al., 2003) and extracted using the Autoscope system. When the extracted data included periods during which the activity of construction was atypical, the data were classified as pertaining to periods of either construction activity or no activity. Specifically, the mean, median and variance of the speeds were calculated. Thus the mean speeds

could be compared with the speed limit corresponding to vehicle type and construction activity to evaluate driver compliance.

Table 1. Work zone sites in this study

| Mile Post | Location | Short or long term | Activity | Date | Duration |
|---|---|---|---|---|---|
| 152.8 | Waynesville | Long | Pavement rehabilitation | Aug 13, 2009 | 17:05 to 20:05 |
| 186.0 | Rolla | Long | Pavement rehabilitation | Oct 2, 2009 | 12:00 to 19:00 |
| 202.6 | Cuba | Short | Pavement rumble striping | Nov 6, 2009 | 11:30 to 16:30 |

## RESULTS

Vehicles' speeds were analyzed to find speed characteristics. Statistical tests were performed to find the significance of difference between speeds of passenger cars, heavy vehicles, and the posted speed limit. The effects of construction activity and lane width reduction on vehicles' speeds were studied. The results are presented in the following sections: speed characteristics, statistical analysis and speed compliance.

### Speed Characteristics

Vehicle speeds of less than 45 mph (mostly stop and go conditions) were deemed to be too slow to represent free flowing vehicles and thus excluded from extracted data. The FFSs were then classified as either car or truck speeds. The minimum, mean, and maximum speed, standard deviation, frequency distribution, and percentage of vehicles exceeding a given speed level were then determined. Tables 1 summarize the speed characteristics of each group for the work zone sites.

FFSs from the Waynesville and Rolla work zone sites were classified based on two criteria, vehicle type and construction activity. Out of all the vehicles counted for Waynesville site, 12.49% were free flowing, including 6.22% passenger cars and 6.27% trucks. Of all the vehicle speeds measured at Cuba site, 4.13% represented free flowing cars and 1.14% represented free flowing trucks. These figures demonstrate that traffic was heavy for Cuba site on the date of data collection, as a queue was observed during various time periods.

For the Waynesville site, the results indicate that the mean speed for all conditions was higher than the speed limit. For the Rolla site, on the contrary, it was lower than the speed limit; however, only statistical tests can determine whether the difference was significant. Furthermore, for both work zone sites, the mean speed of cars was higher than that of trucks for both levels of construction activity. From Tables 1 it can be observed that when there was no construction activity in the work zones, both the number of free flowing vehicles and their speeds increased. For example, at Waynesville site, when there was no significant construction, the

proportion of cars and trucks traveling at least 10 mph above the posted speed limit, was greater than 10%, while during periods of construction activity, this figure was lower. Speed data statistics gathered for Cuba site indicate that the mean speed of all vehicles was higher than the posted speed limit of 60 mph.

Table 1. Speed Characteristics of Vehicles: Waynesville Work Zone

| Speed Characteristics | Waynesville | | | | Rolla | | | | Cuba | |
|---|---|---|---|---|---|---|---|---|---|---|
| | No Construction (mph) | | Construction (mph) | | No Construction (mph) | | Construction (mph) | | No Construction (mph) | |
| | C[1] | T[2] | C[3] | T[4] | C[1] | T[2] | C[3] | T[4] | C[1] | T[2] |
| Mean | 66.6 | 63.6 | 62.6 | 61.4 | 56 | 51.9 | 51.5 | 48.9 | 62.8 | 60.6 |
| Median | 66 | 64 | 63 | 61.5 | 56 | 52 | 52 | 50 | 63 | 60 |
| Standard Deviation | 7.01 | 7.49 | 5.83 | 6.30 | 7.07 | 6.4 | 6.11 | 6.39 | 4.72 | 4.93 |
| Minimum | 48 | 47 | 49 | 46 | 37 | 38 | 38 | 36 | 46 | 48 |
| Maximum | 82 | 83 | 72 | 75 | 70 | 69 | 65 | 61 | 76 | 72 |
| Count | 107 | 89 | 23 | 42 | 162 | 64 | 110 | 47 | 221 | 133 |
| % More than Speed Limit | 80.4 | 67.4 | 65.2 | 59.5 | 25.3 | 7.8 | 5.4 | 2.1 | 157 | 61 |
| % 5 mph more than Speed Limit | 61.7 | 43.8 | 30.4 | 23.8 | 10.5 | 4.7 | 0.0 | 0.0 | 62 | 23 |
| % 10 mph more than Speed Limit | 29.0 | 15.7 | 8.7 | 7.1 | 0.0 | 0.0 | 0.0 | 0.0 | 9 | 3 |

1 Passenger cars for the period with no construction

2 Trucks for the period with no construction

3 Passenger cars for the period with construction

4 Trucks for the period with construction

## Statistical Analysis

Table 2 presents the results of the Anderson-Darling normality test. None of the p-values were significant in 90 % level of confidence. The results fail to reject the null hypothesis and demonstrate that all of the FFS distributions for cars and trucks were normal. Table 3 presents the results of two-sample t-tests comparing the speeds of cars and trucks, and Table 4 provides the results of t-test comparing speeds during construction activity and those during periods of no construction.

Table 2. Results of Normality for FFSs

| Sites | Test parameters | No Construction | | Construction | |
|---|---|---|---|---|---|
| | | Car | Truck | Car | Truck |
| Waynesville | AD value | 0.169 | 0.323 | 0.274 | 0.41 |
| Rolla | AD value | 0.377 | 0.552 | 0.402 | 0.588 |
| Cuba | AD value | 0.702 | 0.641 | - | |

Two-sample t-tests were performed to determine whether the difference between the mean values of the FFSs of cars and trucks was statistically significant.

Table 3 presents the test results. The results of the t-tests indicate that the speeds of cars were significantly greater than the speeds of trucks. For the Waynesville site when there was no construction activity, the difference between the speed of cars and trucks was statistically significant. This difference, however, was insignificant with no construction activity in the work zone. This result suggests that cars were influenced more by construction activity than were trucks. This difference at Rolla site was always significant and that may suggest the higher influence of construction activity on heavy vehicles speed when the lane width was reduced in work zone.

Table 3. Results of t-test: Comparing the Mean Speeds of Cars and Trucks

| Sites | Test parameters | Car vs. Truck speed (mph) | |
|---|---|---|---|
| | | No Construction | Construction |
| Waynesville | Difference | 3.03 | 1.14 |
| | T statistic | 2.92*** | 0.71 |
| Rolla | Difference | 4.162 | 2.55 |
| | T statistic | 4.27*** | 2.32** |
| Cuba | Difference | 2.20 | - |
| | T statistic | 4.18*** | |

*** Significant at 99% level of confidence; ** Significant at 95% level of confidence;

The speed data indicate that the mean speeds of both cars and trucks were lower during periods of construction than during periods of no construction. As indicated in Table 4, the difference was significant for the speeds of passenger cars and trucks at Waynesville and Rolla sites.

Table 4. Results of t-test: Comparison of Mean Speeds during Construction Activity and No Construction

| Sites | Test parameters | No Construction vs. Construction | |
|---|---|---|---|
| | | Car (mph) | Truck (mph) |
| Waynesville | Difference | 4.03 | 2.14 |
| | T statistic | 2.57*** | 1.61* |
| Rolla | Difference | -4.53 | -2.92 |
| | T statistic | -5.62*** | -2.38** |

*** Significant at 99% level of confidence; ** Significant at 95% level of confidence; * Significant at 90% level of confidence

**Speed Compliance Analysis**

To evaluate driver compliance with work zone speed limit, FFSs were statistically compared with the posted speed limit, and the significance of the difference was determined. For this purpose, the speed data were subjected to t-tests with a confidence level of 90%, and the results were analyzed based on the p-values obtained. Table 5 presents the results of the t-tests.

The results of the tests for Waynesville site demonstrate that during both periods of construction and no construction, the speeds of cars and trucks were statistically greater than the posted speed limit. Based on the p-value, with a 90% confidence level, truck speeds were statistically higher than the posted speed limit.

These results demonstrate that construction was not heavy enough and did not prompt vehicles to travel below the speed limit. As noted above, construction did prompt car drivers to slow significantly, but not below the speed limit of 60 mph.

The results of the tests for Rolla site demonstrate that the speeds of cars and trucks were statistically lower than the posted speed limit. These results compared with the Waynesville site demonstrate that the reduced lane width had a greater effect in reducing vehicle speeds. From the results it can be observed that all speeds, whether when there was construction or not, were significantly less than the speed limit. And the results for Cuba site indicate that vehicles traveled significantly faster than the posted speed limit which is similar to the results of the Waynesville site.

Table 5. Results of t-test: comparison of mean speeds with posted speed limit

| Work zone | Test parameters | No Const Car | No Const Truck | Const Car | Const Truck |
|---|---|---|---|---|---|
| Waynesville | Difference | 6.6 | 3.6 | 2.6 | 1.4 |
| | T statistic | 9.74*** | 4.5*** | 2.11** | 1.47* |
| Rolla | Difference | -4.0 | -8.1 | -8.5 | -11.0 |
| | T statistic | -7.13*** | -10.16*** | -14.56*** | -11.85*** |
| Cuba | Difference | 2.8 | 0.61 | - | |
| | T statistic | 8.85*** | 1.43* | | |

*** Significant at 99% level of confidence; ** Significant at 95% level of confidence; * Significant at 90% level of confidence

**SUBJECTIVE ANALYSIS**

A survey questionnaire was prepared and filled by drivers; surveyed drivers on demographics and questions on types of vehicle driven, driver's attention to the posted speed limit and work zone signage such as "work zone ahead", "left/right lane closed" and "speed limit", and their effect on driver behavior, vehicle speed versus posted speed limit, driver's opinion regarding posted speed limit, perceived delay and the effects of construction activity on their driving behavior. A total of 61 drivers participated in the survey.

A survey questionnaire was distributed to drivers at different sites during different times and the results were combined. Drivers were asked if they had passed through the work zone; those who answered positively received a questionnaire. Some of the questions were multiple choice based on a Likert Scale from totally disagree to totally agree, and some were other types of multiple choice questions. The responses were analyzed to determine driver perceptions of reduced work zone speed limits, their compliance , and any hazards present within a work zone.

## Demographics

Participants in this survey ranged from young, 18 to older drivers, above 65 years old. The survey included 74% male and 26% female respondents. They had different driving experiences ranging from 1 year to more than 20 years. Responses also included a variety of vehicle types such as passenger cars, RVs, single and

double trailer trucks. Sixty percent of the participants drove through the highway regularly or at least once a week.

## Survey Analysis

Drivers were asked if operating their vehicle among various types of vehicles posed any safety risks. Eighty percent of drivers disagreed. This statistic shows that most drivers generally have little safety concerns about driving among other types of vehicles. Although this question was not directly related to the speed limit in work zone, answers to this question can be regarded as a factor influencing driver behavior. Drivers whose answers were positive may have reduced their speed because of reasons other than the reduced speed limit. Since most of the drivers disagreed, therefore, speed reduction can be assumed independent of vehicle composition in the traffic stream.

When asked about their preferred distance of “work zone ahead” sign from the start of the work zone, none of the participants felt that they should be alerted within 1 mile of the work zone. These results show that drivers prefer to be well informed more than a mile from the work zone where workers are present and may be hazardous. Only 21% of respondents preferred to see the "closed lane ahead" sign 0.5-1 mile before the work zone. And about 2% of participants preferred to see the “closed lane ahead” sign less than 0.5 mile from the work zone; these were likely the most aggressive drivers in the sample. This figure is in accordance with the driver behavior preference expected regarding the "work zone ahead" sign. Very few participant indicated a desire to see the reduced speed limit sign less than 0.5 mile before the work zone, and more than 73% preferred to see it more than a mile before the work zone.

When asked about their lane changing behavior, more than 80% of participants stated they move out of the closed lane before they can actually see the work zone, of which, 61% preferred to take action immediately after the "lane closed" sign indicating their conservative behavior. From the remaining participants, 15% and 5% preferred to move out of the closed lane only after they saw the work zone, and immediately before the transition taper, respectively, which once again shows the small number of aggressive drivers.

When asked about drivers’ opinion on the speed limit compliance of other drivers, about 49% of participants agreed that other drivers followed the speed limit in the work zone. This should be interpreted in the light of comparison between responses to the question asked about vehicle speed and the posted speed limit which is presented in the next paragraph.

When drivers were surveyed regarding their own speed, all participants indicated that they drove at less than 55 mph within the work zone. When this response is compared with the response to the previous question, that asked participants’ opinion on the speed limit compliance of other drivers, – this indicates that drivers perceive themselves to be obeying the law and others to be disobeying it. All participants indicated traveling below the posted speed limit of 60 mph, but around 50% indicated that other drivers had followed the speed limit in the work zone, which indicates a discrepancy of about 50%.

When surveyed about the appropriateness of the posted speed limit, almost all of the participants suggested a speed limit between 55 and 65 mph as appropriate for work zones. About 78% agreed that driving at the work zone speed limit was safe and this was in agreement with responses to the previous question that asked the participants what speed limit they would suggest.

When asked about the construction activity and if that would make them reduce their speed, 92% agreed completely. One can attribute the low percentage of disagreement with this issue to the aggressive behavior of a small portion of drivers as suggested by responses to previous questions. Those drivers also indicated that construction activities had no effect on their speed, and they preferred to change their lanes only just before the transition taper.

## CONCLUSIONS

This study measured the effects of construction activity and lane width on speed reduction of passenger cars and heavy vehicles in work zones. In addition, it evaluated compliance with posted speed limits in work zones.

The FFS of cars and trucks were found to be normally distributed. With no work zone construction activity, the difference between the speed of cars and that of trucks were statistically significant, but this difference was not significant during periods of construction. Construction activity had a significant effect on speed reduction of passenger cars and heavy vehicles at 90% level of confidence. With both construction and no construction activity, cars and trucks traveled at speeds statistically higher than the posted speed limit. Reduced lane width also had a significant effect on reduction of vehicle speeds. With lane width reduced, construction activity had a greater influence on speeds of heavy vehicles in a work zone.

All participants in the subjective analysis survey indicated that they drove at less than 55 mph within the work zone and about 49% of participants agreed that other drivers followed the speed limit in the work zone. This contradicted the results of the objective analysis and indicates that drivers perceive themselves to be obeying the law and others to be disobeying it.

More than 70 percent of drivers experienced delay and more than 90 percent of drivers agreed that construction activity reduced their speeds which were in agreement with the results of the objective analysis.

## REFERENCES

Agent, K. R., J. G. Pigman, and J. M. Weber (2005). “Evaluation of Speed Limits in Kentucky”. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1640, Paper No. 98-1542, Transportation Research Board of the National Academies, Washington, D.C.

Bella, F. (2005). ‘‘Driving simulator validation for work zone design’’. *CD ROM Proceedings of the 84th Annual Transportation Research Board Meeting*, January 9-13, Washington D.C.

Benekohal, R. F., and J. Shu (1992). "Speed Reduction Effects of Changeable Message Signs in a Construction Zone". *Publication FHWA/IL/UI-239*. FHWA, U.S. Department of Transportation.

Graham-Migletz Enterprises, Inc (1996). "Procedure for Determining Work Zone Speed Limits*"*. *NCHRP Research Results Digest*, No. 192, Transportation Research Board, National Research Council.

Knodler Jr, M. A., D. S. Hurwitz, and H. Rothenberg (2008). "An evaluation of rationally implemented speed limits on collector roadways". *CD ROM Proceedings of the 87th Annual Transportation Research Board Meeting*, January 13-17, Washington D.C.

Migletz, J., J. Graham, B. Hess, I. Anderson, D. Harwood, and K. Bauer (1998). "Effectiveness and Implementability of Procedures for Setting Work Zone Speed Limits". *Publication NCHRP Project 3-41(2)*. Independence: MO: Graham–Migletz Enterprises.

Migletz, J., J. L. Graham, I. B. Anderson, D. W. Harwood, and K. M. Bauer (1999). "Work Zone Speed Limit Procedure". *Transportation Research Record: Journal of the Transportation Research Board*, No. 1657, Paper No. 99-0920, Transportation Research Board of the National Academies, Washington, D.C..

Outkalt, W. (2009). "Work Zone Speed Control". *Publication Colorado Department of Transportation*, CDOT-2009-3.

Richards, S. H., R. C. Wunderlich, and C. L. Dudek (1985). "Field Evaluation of Work Zone Speed Control Techniques". *Transportation Research Record: Journal of the Transportation Research Board*, No. 1035, Transportation Research Board of the National Academies, Washington, D.C., 66–78.

Sandberg, W., T. Schoenecker, K. Sebastian, and D. Soler. (2006). "Long-Term Effectiveness of Dynamic Speed Monitoring Displays (DSMD) for Speed Management at Speed Limit Transitions". *CD ROM ITE Annual Meeting and Exhibit Compendium of Technical Papers.* August 6 - 9, Milwaukee.

Wang, C., K. K. Dixon, and D. Jared (2003). "Evaluating Speed-Reduction Strategies for Highway Work Zones''. *Transportation Research Record: Journal of the Transportation Research Board*, Paper No. 03-4099, Transportation Research Board of the National Academies, Washington, D.C.

# Speed Photo Enforcement Effects on Headways in Work Zones

Ali Hajbabaie1, Hani Ramezani1, and Rahim F. Benekohal1

## Abstract

Determining the effects of Speed Photo-radar Enforcement (SPE) on headways of vehicles traveling in a work zone is important since maintaining a larger headway provides drivers with a longer time to avoid a potential rear-end collision. In addition, when the headway between two vehicles is less than the follower's reaction time, a rear end collision is likely to happen if the leader suddenly decreases its speed. In this paper, the effects of the SPE on the headway of vehicles traveling along a work zone are studied. For this purpose, we collected data from two locations that were around 1.5 miles apart in a work zone on interstate highway 55 at. The SPE van was present at the upstream location. The headway of all vehicles within a period of 45 minutes was measured at the upstream location as well as the downstream location. The results of this study showed that the presence of the SPE did not reduce the headway of the traffic stream at the location that it was present, as well as at the downstream location. This was confirmed by comparing the headway of followers (headway less than four seconds) when the SPE was deployed in the work zone to the headway of them when the SPE was not present, for both lanes combined, and for the shoulder and median lanes separately. The same result was observed for different pairs of leaders and followers. In addition, our analysis showed that the SPE prevented a reduction in the headway of car-car pairs traveling on the median lane at the downstream location.

## Introduction

The severity and frequency of freeway traffic crashes have always been a major concern, particularly in work zones. Understanding the effects of Speed Photo Enforcement (SPE) in work zones on headway of vehicles, safe following distance of vehicles in platoon, and platoon size can help in better management of traffic in construction zones.

A study by Benekohal et. al (2010) showed that Speed Photo-Radar Enforcement has significantly reduced the speed of vehicles in work zones. However, the effects of the SPE on traffic headway in work zones remain to be unknown. There have been some concerns that using the SPE van in a work zone may result in a reduction in traffic headway while vehicles reduce their speed to comply with the work zone speed limit. We have designed this study to determine if the SPE reduces traffic headway in work zones or not. In addition, we will determine SPE's effects on traffic headway on shoulder lane and on median lane separately, and for four different pairs of leaders and followers based on their vehicle type (passenger car, truck).

1 The University of Illinois –Urbana Champaign

Keywords: Traffic Headway, Speed Photo-Radar Enforcement, Work Zone Safety

Not only the effects of the SPE at where it was implemented (Treatment location) will be studied, its effects at a location about 1.5 miles down the road (Downstream location) will be investigated as well. It was shown by Benekohal et. al (2010) that SPE still has some effects on average speed of vehicles at the Downstream location. These effects are less prominent than those at the Treatment location. In other words, vehicles speed up as they are getting away from the Treatment location. The question is that "does this speeding up at the Downstream location cause a reduction in traffic headway?" Hence, headway data at the Treatment location is compared with headway data at the Downstream location. However any difference between headway data at the Treatment location versus the Downstream location in the SPE case could be either due to different effects of SPE at these locations or due to factors such as geometric condition, influencing recurrent traffic pattern, regardless of whether SPE is in the site or not.

For this purpose, we collected data on interstate highway 55 at two locations when SPE van was in the work zone as well as when it was not. Collected data was processed, and analyzed. The rest of this paper discusses our methodology to analyze the data, the data collection set up, findings, and finally our concluding remarks.

## Methodology

To determine if the SPE had any significant effect on headways at the Treatment location, the headway of vehicles at this location should be measured for two cases when a) the SPE van was present in the work zone (the SPE case) and, b) when the SPE van was not present (the Base case). Then using appropriate student t-test, the average headway of followers for these two cases is compared. Followers are those vehicles with headway of less than four seconds. The rest of the vehicles with a headway of more than or equal to four sec. are considered as free flowing vehicles. This study focuses on the headway of followers, since they have short or medium headways and hence they are more likely to be involved in a rear-end collision than free flowing vehicles.

At the first level, the average headway of all followers (cars and trucks combined) in both lanes (shoulder and median together) for the SPE case is compared to that for the base case. This test reveals if overall, SPE has any effect on the headway of followers in the work zone.

Generally, speeds of vehicles travelling on the median lane are higher than those on the shoulder lane; in addition, there was a static sign in the site, asking truck drivers to take the median lane. Since there are some differences in the traffic conditions between shoulder lane and median lane, it is needed to analyze SPE effects on these lanes separately. Moreover, it will be statistically shown that headway data for the shoulder lane should not be combined to the headway data for the median lane. Thus the headway of followers in the shoulder lane for the SPE case is compared to that for the base case. Similar comparison is done for the median lane.

The effect of the lead vehicle type on the headway of the followers should be considered in the headway analysis. The following four types of leaders and followers are defined:

1. A Car, follows another Car: C→C,
2. A Truck, follows a Car: T→C,

3. A Car, follows a Truck: C→T, and
4. A Truck, follows another Truck: T→T

The Headway of a follower vehicle is the time elapsed since the front bumper of the lead vehicle crosses the marker until the front bumper of the follower crosses the same marker. Based on this definition the length of the leader influences the value of the follower's headway. As a result, for a certain travel speed, if for example two cars maintain the same (distance) gap with their leaders, the headway for a C→T pair is longer than the headway of a C→C pair. Thus, the effects of SPE on headways, has to be studied on above-mentioned four different types of leaders and followers and as already mentioned all these analyses should be done for shoulder lane and median lane separately.

All of the above-mentioned analysis reveals any potential effect of the SPE on headway at the location that the SPE is implemented. To identify if the SPE has any potential effect on the headway 1.5 miles downstream of where it is parked, we collected data for base and SPE cases at a location that was 1.5 miles downstream of where the SPE is parked and conducted the same analysis.

It should be mentioned that for all tests we will report the sample size, average headway and the p-value corresponding to the test that is done. Data collection and reduction are explained in details in the next section.

## Data Collection/Reduction

Data was collected in a work zone in I-55 freeway near Chicago during off-peak hours in the afternoon. The work zone started around milepost 255 and was about 7 miles long. The SPE was placed at milepost 259.

The posted speed limit in the work zone was 55 mph and two lanes were open to through traffic. The construction zone was for lane addition and bridge deck repair that was taking place at the time of the data collection. Data was collected at two locations: several hundred feet downstream of the SPE hereafter is called Treatment Location, and 1.5 miles downstream of that, hereafter is called Downstream Location. At the locations that the data was collected, there were two traveled lanes open to the traffic and a normal-width-right-hand-side shoulder was available.

Treatment location was several hundred feet downstream from where the treatment was actually located. This distance was far enough to provide sufficient time for drivers to react to the SPE. SPE was visible before the drivers reached this point.

At both Treatment and Downstream Locations data was collected using camcorders and two markers that were about 200 feet apart. All of the data collection equipments were placed outside of the shoulder with no interference with the traffic stream. Data was collected for the Base case in which normal traffic control devices were in the work zone and for the SPE case in which the SPE van was parked on the right shoulder in the work zone.

For data reduction, the frame number of images was used that provided an accuracy of 0.033 second (one frame) in determining headway. The following information was recorded for all vehicles:

- Time at the both markers

- Vehicle type (passenger car or heavy vehicle)
- Vehicle lane (shoulder lane or median lane)
- Whether the vehicle is free flowing or a follower

During a period of about 45 minutes, headway, vehicle type, and travel lane data for all vehicles in the traffic stream were collected. Table 1, presents vehicle count, percentage of following vehicles, and percentage of trucks in both locations for the base and SPE cases.

**Table 1: Vehicle count, Following Vehicle and Truck Percentage at both locations for the base and SPE cases.**

| Treatment | Location | Vehicle Count | Percentage of | |
|---|---|---|---|---|
| | | | Following Vehicle (%) | Truck (%) |
| SPE | Treatment | 1722 | 78.6 | 20.6 |
| | Downstream | 1676 | 79.8 | 20.3 |
| Base | Treatment | 1787 | 81.1 | 18.7 |
| | Downstream | 1782 | 82.4 | 19.3 |

The headway between two consecutive vehicles was determined using the times at the second marker (because it was closer to the camcorder).

## Findings

To determine if SPE has an overall effect on the headway of vehicles at the location it was implemented, the average headway of vehicles at the Treatment location for the SPE case was compared to that for the Base case using student t-test. As shown in Table 2, the average headway in the base case was 1.76 s. while it was 1.79 s. when the SPE was present in the work zone. However, the difference was not statistically significant as indicated by the P-value of 0.3907. This means that the presence of the SPE van in the work zone did not result in a reduction in the average headway of all following vehicles in both lanes.

**Table 2: Headway in shoulder, median, and both lanes, at the treatment at downstream location for the base and SPE cases.**

| Location | Lane | Treatment | Sample size | Mean Headway (sec) | P-value |
|---|---|---|---|---|---|
| Treatment location | Both lanes | Base | 1449 | 1.76 | 0.3907 |
| | | SPE | 1354 | 1.79 | |
| | Shoulder | Base | 746 | 1.76 | 0.3793 |
| | | SPE | 657 | 1.79 | |
| | Median | Base | 703 | 1.77 | 0.7439 |
| | | SPE | 697 | 1.79 | |
| Downstream location | Both lanes | Base | 1469 | 1.66 | 0.5614 |
| | | SPE | 1337 | 1.68 | |
| | Shoulder | Base | 755 | 1.62 | 0.8029 |
| | | SPE | 676 | 1.63 | |
| | Median | Base | 714 | 1.71 | 0.5921 |
| | | SPE | 661 | 1.73 | |

SPE effect should be studied separately on the shoulder and median lanes. Average headway of followers travelling on the shoulder lane was compared to that on the median lane for a given location and treatment. Based on the results displayed in Table 3, average headway of the followers on the median lane and on the shoulder lane at the Downstream location were different for both the SPE case as well as the Base case. Thus, headway data for shoulder lane and median lane at the Downstream location should be studied separately. It was also decided to conduct the study for the shoulder lane separately from the median lane at the treatment location because, traffic condition on the median lane was different than that on the shoulder lane; Overall, speed of vehicles on the median lane was higher than that on the shoulder lane and also as already mentioned, truck drivers were asked to travel on the median lane by a static sign.

As shown in Table 2, the average headway of all followers on the shoulder lane in the SPE case was not statistically less than that in the Base case. Similar result was obtained for the Median lane.

SPE did not reduce the headway of all followers at its location as well as 1.5 miles downstream of that as shown in Table 2. Similar to the treatment location, at downstream location the SPE did not reduce the average headway of followers on the shoulder, median, and both lanes and in fact, it numerically increased the average headway however, the difference was not statistically significant.

At this point, it was clear that the SPE did not significantly reduce the overall headways on any of the lanes when cars were combined with trucks. However, it was possible that SPE had some effects on the headway of some leader-follower pair but not on the headway of all pairs. Thus, the headway of the four different leader-follower pairs at both locations for the SPE case should be compared to that for the base case but as already mentioned, the traffic condition on the shoulder lane was different than that on the median lane. Thus the analysis should be done separately on the shoulder lane and one the median.

**Table 3: Comparing the average headway in median lane to that in shoulder lane for both locations and cases**

| Treatment | Location | Lane | Sample size | Mean Headway (sec) | P-value |
|---|---|---|---|---|---|
| SPE | Treatment | Median | 697 | 1.79 | 0.8704 |
| | | Shoulder | 657 | 1.79 | |
| | Downstream | Median | 661 | 1.73 | 0.0318 |
| | | Shoulder | 676 | 1.63 | |
| Base | Treatment | Median | 703 | 1.77 | 0.7008 |
| | | Shoulder | 746 | 1.76 | |
| | Downstream | Median | 714 | 1.71 | 0.0526 |
| | | Shoulder | 755 | 1.62 | |

As shown in Table 4, for all four different leader-follower pairs at the Treatment location, the headway in the SPE case was not significantly different than that in the Base case. This indicated that the SPE did not reduce the headway of any pairs of leaders and followers at the location it was present. This observation was desired and showed that at the Treatment location SPE did not have any negative effect on traffic headway. Similar results were found at the downstream location that was expected. It should be noted that the sample size was not large enough to

conduct the test for trucks, following trucks at the Treatment location as well as at the Downstream location.

Table 4: Average Headway for different leader-follower pairs

| Location | Lane | Pair-type | Treatment | Sample size | Mean Headway (sec) | P-value |
|---|---|---|---|---|---|---|
| Treatment location | Shoulder | C→C | Base | 624 | 1.70 | 0.4083 |
| | | | SPE | 568 | 1.74 | |
| | | C→T | Base | 63 | 2.13 | 0.9112 |
| | | | SPE | 47 | 2.15 | |
| | | T→C | Base | 54 | 1.96 | 0.4897 |
| | | | SPE | 36 | 2.10 | |
| | | T→T | Base | 5 | 2.21 | 0.2709 |
| | | | SPE | 6 | 2.83 | |
| | Median | C→C | Base | 429 | 1.52 | 0.2496 |
| | | | SPE | 355 | 1.46 | |
| | | C→T | Base | 116 | 1.85 | 0.2902 |
| | | | SPE | 145 | 1.93 | |
| | | T→C | Base | 71 | 2.33 | 0.2540 |
| | | | SPE | 101 | 2.17 | |
| | | T→T | Base | 87 | 2.46 | 0.4925 |
| | | | SPE | 96 | 2.39 | |
| Downstream location | Shoulder | C→C | Base | 661 | 1.57 | 0.7977 |
| | | | SPE | 580 | 1.58 | |
| | | C→T | Base | 46 | 2.00 | 0.4686 |
| | | | SPE | 50 | 1.90 | |
| | | T→C | Base | 43 | 1.97 | 0.9777 |
| | | | SPE | 41 | 1.97 | |
| | | T→T | Base | 5 | 2.11 | 0.8982 |
| | | | SPE | 5 | 2.17 | |
| | Median | C→C | Base | 409 | 1.36 | 0.2396 |
| | | | SPE | 352 | 1.42 | |
| | | C→T | Base | 114 | 1.85 | 0.7424 |
| | | | SPE | 127 | 1.88 | |
| | | T→C | Base | 76 | 2.31 | 0.1817 |
| | | | SPE | 83 | 2.11 | |
| | | T→T | Base | 115 | 2.40 | 0.3720 |
| | | | SPE | 99 | 2.31 | |

Up to this point, the effects of the presence of the SPE on the average headway were determined at both the Treatment and Downstream locations. As mentioned earlier and shown by Benekohal et. al (2010), at the Downstream location, the speed of vehicles in the SPE case was significantly lower than the speed of vehicles in the Base case. In addition, when vehicles travelled from the treatment location to the downstream location they speeded up. During this interval of speeding up, these vehicles might reduce their headway which was not desired. As a result, we conducted the following tests to check if the traffic headway changed from the

Treatment location to the Downstream location or not. For this purpose in the SPE case, the average headway of traffic stream (excluding free flowing vehicles) at the Treatment location was compared to that at the Downstream location. As shown in Table 5, the average headway of the traffic at the Treatment location in both lanes (1.79 s) was significantly longer than that at the Downstream location (1.68 s). The same trend was observed in the shoulder lane but in the median lane, the headway at the Treatment location was not significantly different from that at the Downstream location. At the first glance, these observations were not desirable since one could conclude that the headway was reduced when vehicles were speeding up. However, this reduction was not due to the presence of the SPE in the work zone as in the Base case the same trend to the same extent was observed (headway reduced from 1.76 s at the Treatment location to 1.66 s at Downstream location for both lanes combined). This observation indicated that the reduction in the headway from the Treatment location to the Downstream location did not happen due to the presence of the SPE. The reason for that might be due to some other factors such as the different roadway geometric condition, but not the presence of the SPE.

**Table 5: Headway at the treatment location versus headway at downstream location.**

| Treatment | Lane | Location | Sample size | Mean Headway (sec) | P-value |
|---|---|---|---|---|---|
| SPE | Both lanes | Treatment | 1354 | 1.79 | 0.0005 |
| | | Downstream | 1337 | 1.68 | |
| | Shoulder | Treatment | 657 | 1.79 | 0.0003 |
| | | Downstream | 676 | 1.63 | |
| | Median | Treatment | 697 | 1.79 | 0.1927 |
| | | Downstream | 661 | 1.73 | |
| Base | Both lanes | Treatment | 1449 | 1.76 | 0.001 |
| | | Downstream | 1469 | 1.66 | |
| | Shoulder | Treatment | 746 | 1.76 | 0.0016 |
| | | Downstream | 755 | 1.62 | |
| | Median | Treatment | 703 | 1.77 | 0.1414 |
| | | Downstream | 714 | 1.71 | |

We also compared the average headway of different pairs of leaders and followers at the Treatment location to that in the Downstream location when the SPE was present in the work zone. The results are shown in Table 6. Except for the C→C pairs in the shoulder lane, for all other pairs in both shoulder and median lanes, when the SPE was present in the work zone, the headway at the Treatment location was not significantly different from the headway at the Downstream location. This meant that among different pairs of leaders and followers in the shoulder lane the C→C pair was the most significant contributing pair to the difference between the headway at the Treatment and Downstream locations.

For the base case, except for the C→C pair in both shoulder and median lane, for all leader-follower pairs, the average headway at the Treatment locations was similar to that at the Downstream location. Comparing the SPE case to the Base case in the median lane for the C→C pairs showed that when the SPE was not present in the work zone C→C pairs reduced their headway when they traveled from the Treatment location to the Downstream location. However, when the SPE was present in the work zone, these C→C pairs did not significantly reduce their headway in travelling from the Treatment location to the Downstream location. This was a

desirable finding as it indicated that the SPE resulted in no reduction in the headway of C→C pairs in the median lane when traveled from the Treatment location to the Downstream location.

**Table 6: Headway at the treatment location versus headway at downstream location for different leader-follower pairs**

| Treatment | Lane | Pair-type | Location | Sample size | Mean (sec) | P-value |
|---|---|---|---|---|---|---|
| SPE | Shoulder | C→C | Treatment | 568 | 1.74 | 0.0013 |
| | | | Downstream | 580 | 1.58 | |
| | | C→T | Treatment | 47 | 2.15 | 0.1105 |
| | | | Downstream | 50 | 1.90 | |
| | | T→C | Treatment | 36 | 2.10 | 0.5478 |
| | | | Downstream | 41 | 1.97 | |
| | | T→T | Treatment | 6 | 2.83 | 0.0953 |
| | | | Downstream | 5 | 2.17 | |
| | Median | C→C | Treatment | 355 | 1.46 | 0.5278 |
| | | | Downstream | 352 | 1.42 | |
| | | C→T | Treatment | 145 | 1.93 | 0.4870 |
| | | | Downstream | 127 | 1.88 | |
| | | T→C | Treatment | 101 | 2.17 | 0.6325 |
| | | | Downstream | 83 | 2.11 | |
| | | T→T | Treatment | 96 | 2.39 | 0.4563 |
| | | | Downstream | 99 | 2.31 | |
| Base | Shoulder | C→C | Treatment | 624 | 1.70 | 0.0047 |
| | | | Downstream | 661 | 1.57 | |
| | | C→T | Treatment | 63 | 2.13 | 0.3797 |
| | | | Downstream | 46 | 2.00 | |
| | | T→C | Treatment | 54 | 1.96 | 0.9619 |
| | | | Downstream | 43 | 1.97 | |
| | | T→T | Treatment | 5 | 2.21 | 0.8782 |
| | | | Downstream | 5 | 2.11 | |
| | Median | C→C | Treatment | 429 | 1.52 | 0.0022 |
| | | | Downstream | 409 | 1.36 | |
| | | C→T | Treatment | 116 | 1.85 | 0.9680 |
| | | | Downstream | 114 | 1.85 | |
| | | T→C | Treatment | 71 | 2.33 | 0.8711 |
| | | | Downstream | 76 | 2.31 | |
| | | T→T | Treatment | 87 | 2.46 | 0.5386 |
| | | | Downstream | 115 | 2.40 | |

## Conclusions

Our study showed that the presence of the SPE van in work zones does not reduce traffic headway at the location of the SPE.

The average headway of all traffic stream (excluding free flowing vehicle) when the SPE was present in the work zone was not greater than the headway of all traffic stream when the SPE was not present in the work zone. The similar results were obtained when data were separated by

lane and effects of the SPE on the traffic travelling on the median lane were studied separately from the shoulder lane.

Analysis of the four different pairs of leaders and followers revealed that SPE did not reduce the headway of any of these pairs. Similar results were observed at the downstream location. In summary, we found out that the SPE did not have any negative effect on traffic headway neither at the location it was parked nor around 1.5 miles downstream of that location.

Comparing the headway at the Treatment location to that at the Downstream location showed that whether the SPE was present in the work zone or not, the headway of traffic stream significantly reduced at the Downstream location. This fact was observed for the headway of traffic stream at both lanes combined, as well as in the shoulder lane separately but, not in the median lane. This observation indicated that the drivers reduced their headway as they travel to the Downstream location. However, analyzing the headway of different pairs of leaders and followers showed that when the SPE was present in the work zone, the headway of cars following cars (C→C pairs) on the median lane was not reduced. This means that the SPE resulted in a longer headway for C→C pair in the median lane. In summary, we found that not only the SPE did not reduce the headway of traffic stream in the work zone, for C→C pairs, it increased the headway at the Downstream location.

In our analysis we compared the average headway of all following vehicles in the traffic stream for all cases. However, it was possible that some of these cars that were followers at one of the locations became the platoon leader and a free flowing vehicle at the other location. As a result a study that tracks the vehicles from the treatment location to the downstream location is needed and is underway.

In addition, the behavior of those drivers that maintain extremely short headways in work zones needs to be studied. This could be done by matching the vehicles at both location to determine if these vehicles maintained the same short headway at both locations, or only at one,

or they left the platoon and become a free flowing vehicle at the downstream location, and so on. A study to address these questions is underway.

## Reference

Benekohal, R. F., Hajbabaie, A., Medina, J. C., Wang, M., & Chitturi, M. V. (2010). *Speed Photo-Radar Enforcement Evaluation in Illinois Work Zones.* Illinois Center for Transportation.

# Analysis of Single-Lane Roundabout Slip Lanes Using SIDRA

Majed Al-Ghandour P.E.[1]; William Rasdorf Ph.D., P.E.[2]; Billy Williams Ph.D., P.E.[3]; and Bastian Schroeder Ph.D.[4]

[1]Assistant Branch Manager, Program Development Branch, North Carolina Department of Transportation, Raleigh, NC 27699-1534, PH (919) 395-8859; Fax: (919) 733-3585; email: malghandour@ncdot.gov
[2]Professor, Department of Civil, Construction, and Environmental Engineering, North Carolina State University, Raleigh, NC 27606-7908, PH (919) 515-7637 Fax: (919) 515-7908; email: rasdorf@.ncsu.edu
[3]Associate Professor, Department of Civil, Construction, and Environmental Engineering, North Carolina State University, Raleigh, NC 27606-7908, PH (919) 515-7813 Fax: (919) 515-7908; email: billy_williams@ncsu.edu
[4]Senior Research Associate, Institute for Transportation Research & Education, North Carolina State University, Centennial Campus Box 8601, Raleigh, NC 27695-8601, PH (919) 515-8565 Fax: (919) 515-8898; email: bastian_schroeder@ncsu.edu

## ABSTRACT

Roundabout intersections, increasingly used in the U.S., sometimes incorporate slip lanes to facilitate right-turning traffic flow and reduce delay, thereby increasing capacity and safety. Performance of a single-lane roundabout with an adjacent slip lane is modeled with the SIDRA Intersection analysis tool for three types of slip lane: free-flow, yield, and stop, and is compared to having no slip lane. The gap acceptance-based assessment considers four experimental traffic percentage distribution matrices representing flow scenarios. SIDRA results confirm that average delay and circulating conflict volumes in a roundabout with a slip lane are related exponentially to slip lane volumes. A free-flow slip lane exit type helps to reduce total average delay in the roundabout and the slip lane approach. Both yield and stop slip lane exit types also reduce roundabout total average delay but to a lesser degree than a free-flow slip lane. Finally, theoretical capacity threshold values for slip lane volumes are estimated to range from 150 to 350 vehicles per hour for traffic volume distribution scenarios.

**Author keywords**: Roundabout**,** slip lane, gap acceptance, SIDRA, traffic delay, exit type, experimental traffic percentage distribution matrices.

## INTRODUCTION

As modern roundabouts gain popularity in the U.S., consideration of the safety and performance effects of slip lanes also gains importance. A slip lane, a separate lane that facilitates right-turning traffic flow, reduces approach delay by allowing right-turning movements to bypass the roundabout, thereby reducing vehicle conflicts. Though roundabouts are an increasingly common form of intersection control in the U.S., research has yet to quantify slip lane contributions to operational and safety improvements when slip lanes are installed.

NCHRP Report 572 (NCHRP 2007) defines the types of slip lane: a non-yield slip lane, merging with the roundabout exit leg and forming a new acceleration (free-flow) lane adjacent to exiting traffic; and a yield slip lane, terminating at a sharp angle with the roundabout exit approach so that right-turning traffic is yielding.

Operational performance of roundabouts, measured as roundabout capacity, typically is based on one of three capacity methods: gap acceptance; empirical regression; or a hybrid of gap and empirical methods. TRB (2000) and NCHRP Report 572 (NCHRP 2007) provided roundabout capacity models as a function of the circulating flow in the roundabout, follow-up headway, and critical gap. They estimated the capacity of a roundabout's approach (entry lanes) via input parameters such as circulating conflicting traffic volume, follow-up time, and critical gap. U.K. and German linear (empirical) regression methods used roundabout geometry parameters without consideration of driver behaviors (NCHRP 2007). In Australia, Akcelik (2007) derived the gap acceptance capacity model as an exponential relationship between capacity and opposing flow rate in a roundabout.

Neither capacity nor safety evaluations were found in the literature for roundabouts where slip lanes were installed. Nothing was found specifically focused on roundabout slip lanes.

SIDRA software (Signalized and Unsignalized Design and Research Aid, SIDRA 2007) is commonly used to analyze traffic operations at roundabouts. Akcelik and Besley (2004) also observed that both the AUSTROADS Roundabout Guide and the SIDRA software use gap acceptance techniques for roundabout capacity and performance analysis based on empirical models to estimate gap-acceptance parameters.

SIDRA was used in this study to explore experimental traffic flows in a single-lane roundabout with a slip lane. The work reported in this paper used SIDRA to test two hypotheses:

1. All slip lane types should reduce average delay in a single-lane roundabout, thereby improving its operational performance (its capacity as indicated by average delay), compared to a base single-lane roundabout without a slip lane. A free-flow slip lane exit type is expected to reduce delay more than a yield exit type and a yield exit type more than a stop exit type.

2. Slip lane theoretical threshold capacity volumes increase with increased roundabout right-turn traffic volumes. A threshold value is the limit in volumes where the roundabout operates with a capacity as transition to level of service F (traffic congestion is the result of more traffic flow demand than capacity).

## METHODOLOGY

Gap acceptance theory is used to determine the capacity of each approach and of the entire roundabout (including slip lanes). SIDRA software quantifies parameter values for available operations and geometric data (SIDRA 2007). SIDRA automatically sets gap acceptance parameters for the roundabout and slip lane as a function of geometry, circulating flow rate, entry flow rate, and other factors; it limits other parameters such as critical gap headway ($t_c$) to range from 2 to 8 seconds, and follow-up headway ($t_f$) to range from 1 to 5 seconds. Inputs to SIDRA include vehicle volumes and movement paths, yield and stop slip lane exit

configurations, gap acceptance and follow-up headway attributes for selected movements, and roundabout geometry attributes (inscribed diameter, number of entry lanes, and average entry lane width). Average roundabout delay, the average vehicle delay in seconds for all vehicles entering the roundabout, is the Measure of Effectiveness (MOE).

In this study, for a single-lane roundabout with a slip lane, four experimental traffic percentage turning volume distributions (scenarios) were assumed for balanced scenarios (traffic flow into and out of every roundabout approach is the same). The scenarios (S1, S2, S3, and S4) were initialized, analyzed, and then controlled through several iterations. A slip lane was assumed to be placed at the northbound (NB) entry to the roundabout as shown in Figure 1.

Several variables were tested across the four traffic percentage distribution scenarios as follows:

1. Slip lane exit type (free flow lane, yield sign, and stop sign) compared to having no slip lane (base case).
2. Slip lane right-turning traffic volume as the dominant turn (in increments of 50 vehicles per hour and ranging from 50 vehicles per hour to 500 vehicles per hour—representing low, moderate, and high volumes).
3. Approach entry volume.
4. Traffic percentage distribution flow patterns.

A range of traffic conditions intended to imitate real-world traffic flow patterns was represented in these traffic percentage distribution scenarios. In the four flow scenarios, shown in Figure 1, roundabout entry and exit flows for each approach are the same, although the dominant right-turning traffic percentages in the slip lane are different (33%, 45%, 60%, and 75%).

Volume distributions for the roundabout were developed from the traffic percentage distribution matrices and are summarized as shown in a sample of three volumes: 50 vehicles per hour, 250 vehicles per hour, and 500 vehicles per hour (Table 1). Volumes for each roundabout approach ($V_a$) are the same as for exit approach ($V_{exit}$) volumes, based on the assumption of experimental balanced roundabout scenarios. For example, scenario S1, dominant right turn flow percent (33%), sustains more traffic volumes on both approach entry ($V_a$) and exit approaches ($V_{exit}$) and circulating flow ($V_c$), compared to S2 (45%), S3 (60%), and S4 (75%) dominant turn levels. At a slip lane (right turn) volume $V_{sl}$=500 vehicles per hour for S4, a 75% dominant level (high right-turn flow) shows a volume at each approach ($V_a$) of 667 vehicles per hour, (highlighted in yellow, Table 1), smaller values than S1 (1515 vehicles per hour), S2 (1110 vehicles per hour), and S3 (834 vehicles per hour) dominant levels. The conflicting circulating volumes ($V_c$) for S4 (467 vehicles per hour) are also smaller values than S1 (1515 vehicles per hour), S2 (1110 vehicles per hour), and S3 (709 vehicles per hour). All traffic volume distributions scenarios (S1, S2, S3, and S4) were coded into SIDRA for different scenarios to evaluate the performance of a slip lane in terms of average delay and SIDRA sensitivity.

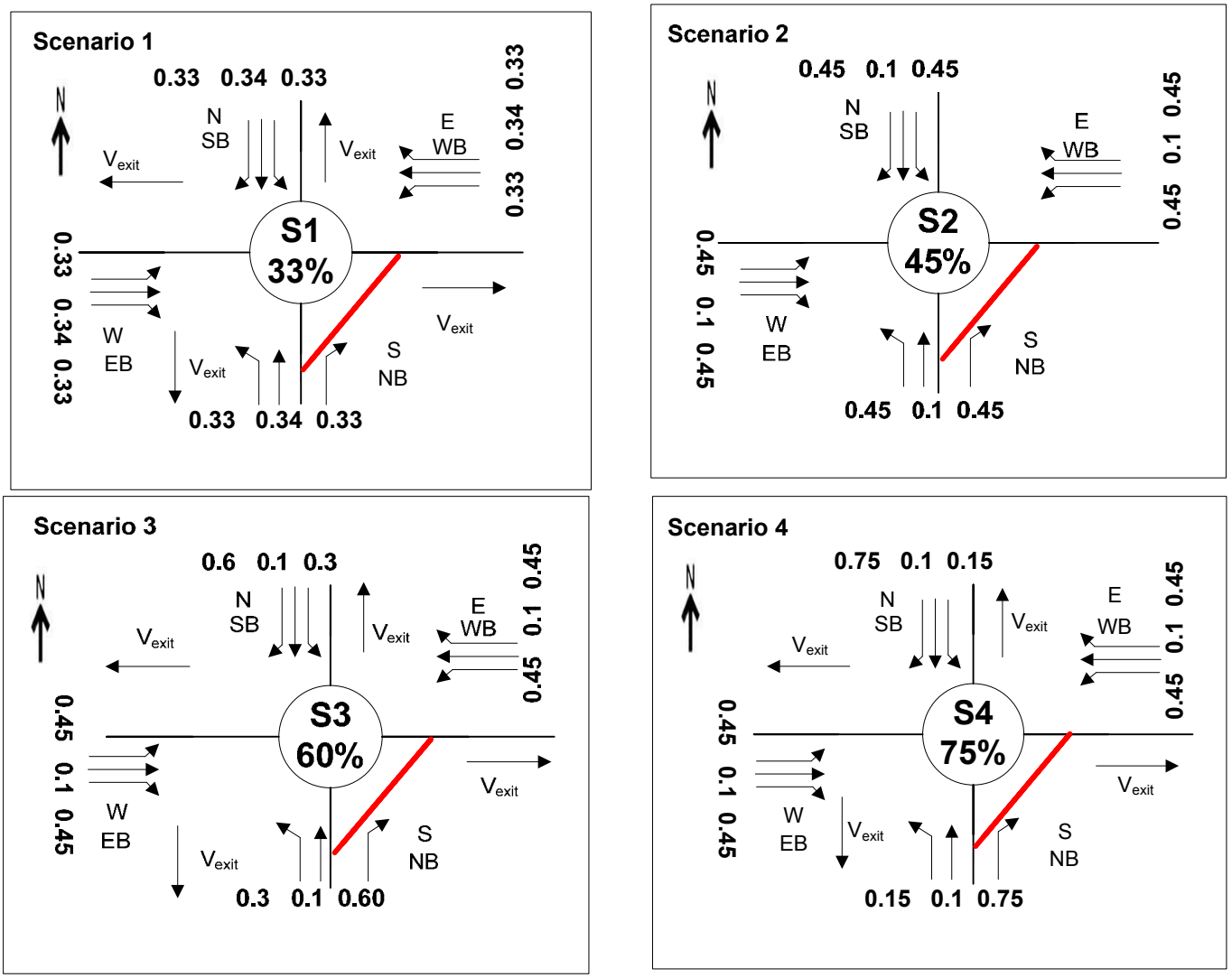



**Figure 1. Traffic Percentage Distribution Flow Pattern Scenarios (S1-S4).**

**Table 1. Total Approach and Conflicting Volumes for Scenarios S1-S4.**

| Slip Lane Volume as Dominant Right Turn, (vehicles per hour), $V_{sl}$ | | Scenarios | | | |
|---|---|---|---|---|---|
| | | S1 (33%) | S2 (45%) | S3 (60%) | S4 (75%) |
| $V_{sl}$ =50 (Low) | $V_a$<br>$V_c$ | 150<br>(150) | 111<br>(111) | 83<br>(71) | 66<br>(46) |
| $V_{sl}$ =250 (Moderate) | $V_a$<br>$V_c$ | 757<br>(757) | 555<br>(555) | 416<br>(354) | 333<br>(233) |
| $V_{SL}$ =500 (High) | $V_a$<br>$V_c$ | 1515<br>(1515) | 1110<br>(1110) | 834<br>(709) | 667<br>(467) |

$V_a$: Approach volumes per approach, vehicles per hour. $V_c$: Conflicting volumes for northbound entry (NB), vehicles per hour. $V_{sl}$: Slip lane volumes as dominant right turn, vehicles per hour.

## ANALYSIS AND RESULTS

A conflict point is any point "where a vehicle path crosses, or merges with another vehicle path" (FHWA 2000). The most likely conflict point in a single-lane roundabout is merging, based dynamically on vehicle traffic events in a specific time and space. To illustrate conflict volumes, a sample is shown in Figure 2 showing the northbound details for Scenario S1 (33%). Slip lane right turns volumes ($V_{sl}$), approach volumes ($V_a$), exit approach volumes ($V_{exit}$), entry volumes ($V_e$), conflicting circulating volumes for the roundabout at northbound ($V_c$), and conflicting volumes off slip lane ($V_m$) are assumed and calculated as follows:

$V_a$= volumes at (4) + volumes at (5) + volumes at (6) = 0.33+0.34+0.33=1.0.
$V_e$= volumes at (4) + volumes at (5) = 0.33+0.34 =0.67.
$V_c$= volumes at (1) + volumes at (2) + volumes at (3) = 0.33+0.33+0.34=1.0.
$V_{sl}$= volumes at (6) = 0.33.
$V_m$= volumes at (1) + volumes at (3) = 0.33+0.34=0.67.
$V_{exit}$= (volumes at (1) + volumes at (3)) + volumes at (6) =$V_m$ +$V_{sl}$ = 0.67+0.33=1.0.

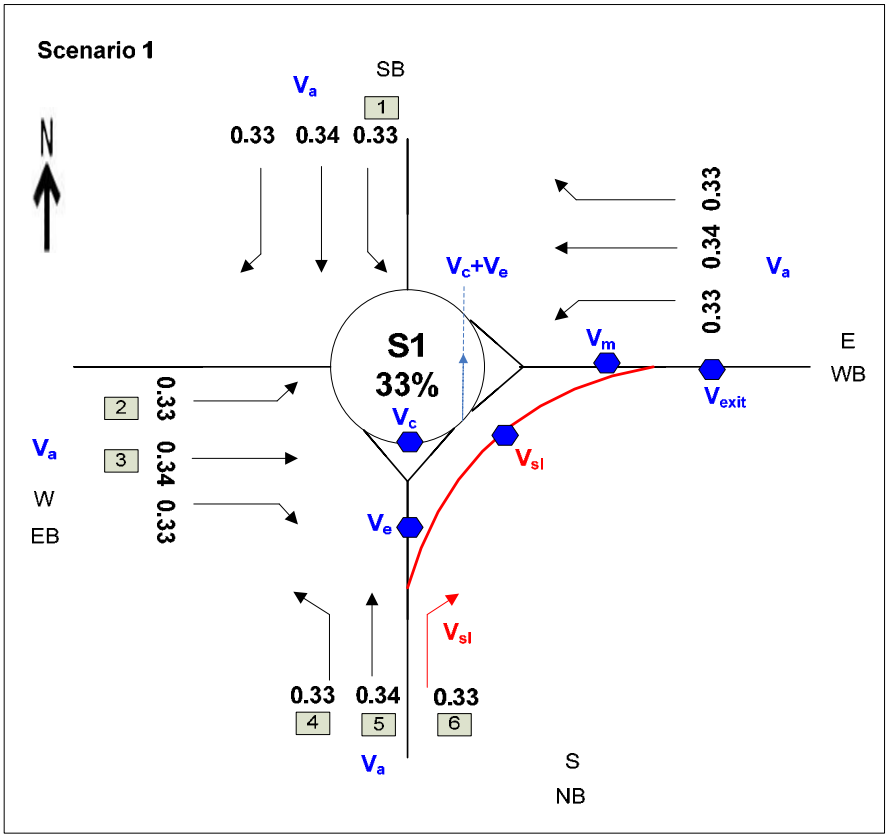



**Figure 2. Sample of Traffic Percentage Distribution Flow Pattern Scenario S1.**

$V_a$: Approach volumes per approach, vehicles per hour. $V_{exit}$: Exit volumes per approach, vehicles per hour. $V_e$: Entry volumes for the specific northbound entry (NB), vehicles per hour. $V_c$: Conflicting volumes for northbound entry (NB), vehicles per hour. Vm: Conflicting volumes off slip lane, vehicles per hour. $V_{sl}$: Slip lane volumes as dominant right turn, vehicles per hour.

As, more traffic is diverted outside the roundabout using the slip lane (right-turn movement), slip lane will have the greatest impact on reducing the roundabout conflicting circulating volumes ($V_c$) and the conflicting off slip lane approach volumes ($V_m$).

### Conflicting Volumes

Conflicting volume ($V_c$) flow pattern ratios for all scenarios are summarized (Table 2). Compared to scenarios S1-S3, S4 shows the lowest slip lane ($V_c$) ratio (0.7, highlighted in yellow). Thus, the highest right-turn slip lane volumes ($V_{sl}$) reduce the conflicting volumes ($V_m$) off slip lane (0.25). A higher dominant right-turning traffic percentage in the slip lane also causes a lower conflicting circulating volume ($V_c$) for the northbound (NB) entry approach and lower entry volumes ($V_e$) into the roundabout.

**Table 2. Conflicting Volume Ratios for Scenarios S1-S4.**

| Type of Volumes | Scenarios | | | |
|---|---|---|---|---|
| | S1 (33%) Low Right Turn Flow | S2 (45%) Moderate Right Turn Flow | S3 (60%) Moderate Right Turn Flow | S4 (75%) High Right Turn Flow |
| $V_a$ | 1.0 | 1.0 | 1.0 | 1.0 |
| $V_e$ | 0.67 | 0.55 | 0.40 | 0.25 |
| $V_c$ | 1.0 | 1.0 | 0.85 | 0.7 |
| $V_{sl}$ | 0.33 | 0.45 | 0.60 | 0.75 |
| $V_m$ | 0.67 | 0.55 | 0.40 | 0.25 |
| $V_{exit}$ | 1.0 | 1.0 | 1.0 | 1.0 |

$V_a$: Approach volumes per approach, vehicles per hour. $V_{exit}$: Exit volumes per approach, vehicles per hour. $V_e$: Entry volumes for the specific northbound entry (NB), vehicles per hour. $V_c$: Conflicting volumes for northbound entry (NB), vehicles per hour. $V_m$: Conflicting volumes off slip lane, vehicles per hour. $V_{sl}$: Slip lane volumes as dominant right turn, vehicles per hour.

### Average Roundabout Delay

The SIDRA results for the four scenarios are compared (Table 3). A free-flow slip lane exit type with high traffic volumes (500 vehicles per hour) shows significant reduction (operational improvement) in roundabout average delay, from 42.7 sec/vehicle (no slip lane) to 31.8 sec/vehicle: a 25.5% reduction in S4, as highlighted in yellow. Within a yield slip lane exit type, the reduction of roundabout average delay is 25.48%. Within a stop slip lane exit type, reduction is 24.8%. As expected, the results from SIDRA for the four scenarios show that delay is significantly reduced in a single-lane roundabout with any type of slip lane, before oversaturation occurs, with only marginal differences in delay reduction between the slip lane types (yield and stop). The northbound (NB) approach, the slip lane approach, shows significant average delay

reduction—91.68%—when slip lanes are used, regardless of exit types, compared to having no slip lane at high traffic volumes (500 vehicles per hour). The northbound average delay reduction percentage is calculated as (for example, using a yield sign exit type) 91.68% = ((3.8-45.7)/45.7). The slip lane average delay within a stop slip lane exit type (32.1 sec/vehicle in S4, 500 vehicles per hour) is slightly higher than the average delay of a yield exit type(31.8 sec/vehicle). Because a free-flow right-turning slip lane exit type has no opposing exiting flow from the roundabout—and therefore a high capacity—a free-flow slip lane exit type has no delay (zero). Reduction of delay via the use of free-flow slip lanes is shown to be greater than in stop or yield sign slip lanes. Thus, a free-flow slip lane exit type provides significant reductions in total roundabout average delay and to the slip lane approach.

Samples from the SIDRA results are shown in Figures 3 and 4 for all scenarios using a yield slip lane exit type and having no slip lane as the base. Corresponding oversaturated conditions (volume/capacity>1.0) that are expected to occur as roundabout approach volumes and right-turning traffic volumes are increased are also shown. Lower right-turning volumes ($V_{sl}$) showed fast-occurring roundabout oversaturation. For example, Scenario S1 (33%), presenting the greatest delay, oversaturated at a right-turning volume of 200 vehicles per hour, S2 at 300 vehicles per hour, and S3 at 400 vehicles per hour; these oversaturation points are marked by dashed lines in Figure 3. In contrast, as expected, average delay is reduced within a roundabout with a slip lane, within a higher percentage of slip lane volumes as shown by S1 (from 715.7 sec/vehicle with no slip lane to 672.29 sec/vehicle with a yield slip lane) to S4 (from 42.68 sec/vehicle with no slip lane to 31.80 sec/vehicle with a yield slip lane).

**Table 3. Summary of SIDRA Average Delays.**

| Slip Lane Exit Type | $V_{SL}$: Slip Lane Volume, Right-Turn Volume (vehicles per hour) at NB Approach | Average Delay (sec/vehicle) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Scenarios | | | | | | | | | | | |
| | | S1 (33%) | | | S2 (45%) | | | S3 (60%) | | | S4 (75%) | | |
| | | SL | North App. (NB) | RBT | SL | North App. (NB) | RBT | SL | North App. (NB) | RBT | SL | North App. (NB) | RBT |
| Yield | 50 (Low) | 0.6 | 1.0 | 1.0 | 0.4 | 0.7 | 0.7 | 0.2 | 0.5 | 0.5 | 0.1 | 0.3 | 0.3 |
| | 250 (Med) | 3.1 | 26.7 | 151.1 | 2.2 | 7.1 | 22.1 | 1.2 | 2.7 | 3.9 | 0.6 | 1.6 | 1.8 |
| | 500 (High) | 6.6 | 384.1 | 672.3 | 3.8 | 83.8 | 428.1 | 2.6 | 7.9 | 141.7 | 1.6 | 3.8 | 31.8 |
| Stop | 50 (Low) | 0.6 | 1.0 | 1.0 | 0.4 | 0.7 | 0.7 | 0.2 | 0.5 | 0.5 | 0.1 | 0.3 | 0.3 |
| | 250 (Med) | 3.1 | 26.7 | 151.1 | 2.2 | 7.1 | 22.1 | 1.2 | 2.7 | 3.9 | 0.6 | 1.6 | 1.8 |
| | 500 (High) | 6.0 | 384.1 | 672.2 | 3.5 | 83.3 | 428.0 | 2.4 | 7.9 | 141.7 | 1.6 | 3.8 | 32.1 |
| Free-Flow | 50 (Low) | 0.0 | 1.0 | 1.0 | 0.0 | 0.7 | 0.7 | 0.0 | 0.5 | 0.4 | 0.0 | 0.3 | 0.2 |
| | 250 (Med) | 0.0 | 26.6 | 150.9 | 0.0 | 7.1 | 21.8 | 0.0 | 2.7 | 3.8 | 0.0 | 1.6 | 1.7 |
| | 500 (High) | 0.1 | 384.1 | 671.8 | 0.1 | 83.5 | 427.6 | 0.1 | 7.9 | 141.3 | 0.0 | 3.8 | 31.8 |
| No Slip | 50 (Low) | 1.1 | 1.1 | 1.1 | 0.8 | 0.8 | 0.8 | 0.5 | 0.5 | 0.5 | 0.3 | 0.3 | 0.3 |
| | 250 (Med) | 162.8 | 162.8 | 162.8 | 28.5 | 28.5 | 28.0 | 4.7 | 4.7 | 4.7 | 2.1 | 2.1 | 2.1 |
| | 500 (High) | - | 715.7 | 715.7 | 480.6 | 480.6 | 480.6 | - | 177.2 | 170.3 | - | 45.7 | 42.7 |

$V_{SL}$: Slip lane volumes as dominant right turn, vehicles per hour. RBT: roundabout. NB: Northbound approach, slip lane approach. S1 to S4: Scenarios. SL: Slip lane. Yield: Slip lane with a yield exit type. Stop: Slip lane with a stop exit type. Free-Flow: Free slip lane exit type. No Slip: No slip lane (base).

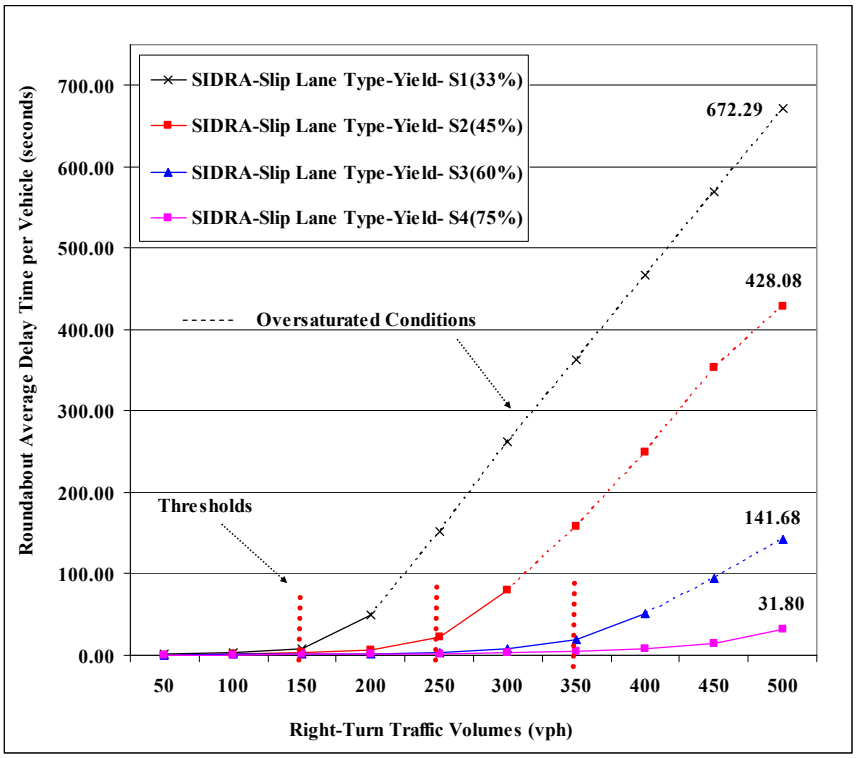



**Figure 3. SIDRA Roundabout Average Delay for Slip Lane Yield Exit Type.**

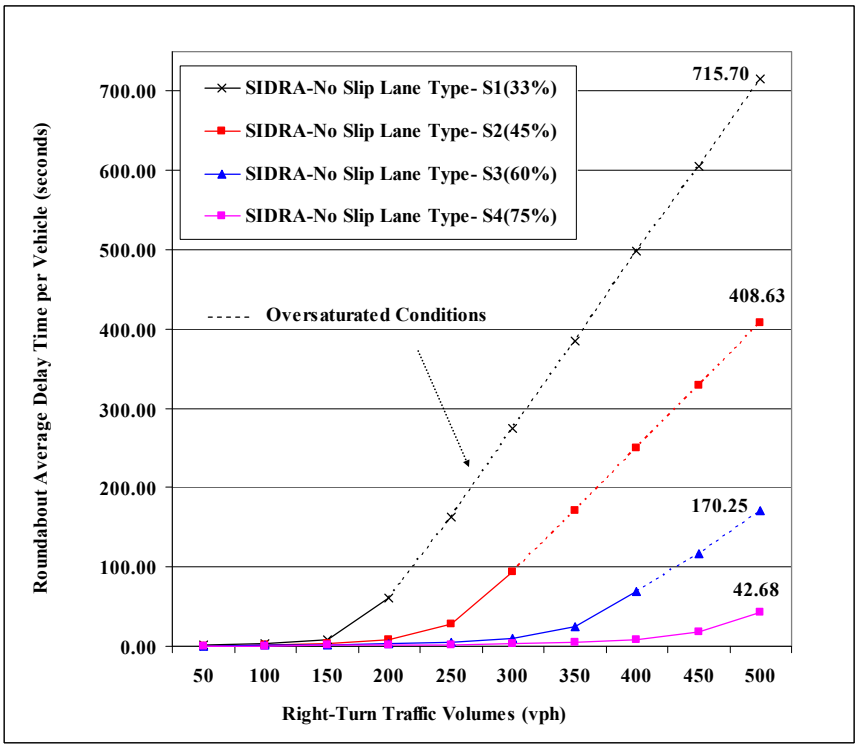



**Figure 4. SIDRA Roundabout Average Delay Having No Slip Lane (base).**

As slip lane (right-turning) traffic volume ($V_{sl}$) increases, the conflicting circulating volumes ($V_c$), decrease and the average delay also decreases, in a non-linear, or exponential, relationship. The highest roundabout average delay, observed in S1, was a result of the combined highest approach volumes ($V_a$), highest total roundabout volumes, and highest conflicting circulating flow ($V_c$). The lowest roundabout average delay, observed in S4, was a result of the combined lowest approach volumes ($V_a$), lowest total roundabout volumes, and lowest conflicting circulating flow ($V_c$). Therefore, under different scenarios, slip lane performance is most effective under a higher right-turning traffic pattern distribution. Slight changes between entry lane flow ($V_e$) and circulating flow ($V_c$) cause less impact on the total average delay within a roundabout with a slip lane

### Threshold Capacity Values

Slip lane theoretical threshold volumes were determined when the roundabout operates with a volume/capacity ratio greater than 1.00 as transition to Level of Service F (traffic congestion is the result of more traffic flow demand than capacity). The threshold capacity volumes are based on SIDRA average delay results (Figure 3: threshold limits are shown by vertical dotted lines). At oversaturated conditions, all delay values greater than 50 seconds also are shown by dashed lines, representing the proposed LOS F threshold for U.S. roundabouts (NCHRP 2007). As roundabout approach volumes ($V_a$) and right-turning volume ($V_{sl}$) increase, threshold capacity occurs at higher volumes: whereas S1 shows a threshold of around 150 vehicles per hour to be oversaturated, the S2 threshold occurs at 250 vehicles per hour, and the S3 threshold occurs at 350 vehicles per hour, for a slip lane yield exit type. An interesting point is that for all slip lane types in these scenarios, threshold values are the same per same scenario. SIDRA analysis shows therefore that the single-lane roundabout operates with a volume/capacity ratio greater than 1.00 regardless of slip lane type. Hence, the slip lane theoretical capacity operational thresholds volumes increase with greater right-turn volumes ($V_{sl}$).

## CONCLUSIONS

SIDRA results confirm that average delay and circulating conflict volumes in a roundabout with a slip lane are related exponentially to slip lane volumes. Higher right-turn traffic volumes (percentages) in the slip lane reduce roundabout conflicting circulating volume ($V_c$) and conflicting volumes off slip lane ($V_m$). Results also showed overall roundabout delay was reduced 25.51% with use of a free-flow slip lane; 25.48% with use of a yield slip lane; and 24.80% with use of a stop slip lane. Hence the most effective roundabout delay performance generally is obtained from a free-flow slip lane.

Average delay is more effectively reduced within a roundabout with a slip lane than in one without. The most striking example from this study is a 91.68% delay in reduction from use of a slip lane in the northbound (NB) slip lane approach (scenario S4 at 500 vehicles per hour) regardless of the slip lane exit types.

When roundabout traffic becomes oversaturated, any reduction of delay obtained is similar for both yield or stop slip lane exit types. Theoretical capacity thresholds values (limits) for right-turning slip lane volumes ($V_{sl}$) are estimated to be within a range of 150 to 350 vehicles per hour for traffic distribution volumes.

## RECOMMENDATIONS

SIDRA can be used to analyze a slip lane's contribution to improved roundabout capacity and safety before traffic flow becomes oversaturated. Determining theoretical threshold value ranges can be helpful to practitioners who are considering the use of a slip lane in a roundabout design. Fully testing multiple traffic volume distribution matrices with other percentages of dominant right-turning traffic provides insight to different roundabout pattern flows. Therefore, understanding the effect of traffic demand and distribution patterns of traffic on roundabout delay will help in assessing a slip lane's impact on improving operational performance. Additional analysis should be conducted for other variables: different unbalanced flow scenarios (traffic flow into and out of different roundabout approach is different); number of lanes within a slip lane (one or two); slip lane widths (10 ft or 12 ft) and lengths; distance of slip lane exit/merge from the roundabout; number of pedestrian crosswalks (one or two); and other geometric configurations. To validate SIDRA sensitivity, future analysis should vary its default values for gap acceptance parameters (critical gap and follow-up headway per lane) and compare results to field data.

## REFERENCES

Akcelik, R. (2007). "A review of gap-acceptance capacity models." Paper presented at the 29th Conference of Australian Institutes of Transport Research (CAITR 2007), Adelaide, December 2007.

Akcelik, R. A., and Besley, M. (2004). "Differences between the AUSTROADS roundabout guide and aaSIDRA roundabout analysis methods." Presented at 26th Conference of Australian Institutes of Transport Research (CAITR 2004), Clayton, Melbourne, 8-10 December 2004.

Federal Highway Administration (FHWA) (2000). "Roundabouts: An Informational Guide." Publication No. FHWA-RD-00-067, Washington, D.C.

National Cooperative Highway Research Program (NCHRP) (2007). "Roundabouts in the United State." National Research Council, Transportation Research Board, National Cooperative Highway Research Program; NCHRP Report 572, Washington, D.C.

SIDRA (2007). SIDRA User's Manual. P.O. Box 1075G, Greythorn, Vic 3104, Australia.

Transportation Research Board (TRB) (2000). *Highway Capacity Manual.* 4th edition. National Research Council, Transportation Research Board, Washington, D.C.

# An Evaluation of New Jersey Jug-handle Intersection (NJJI) with and without Pre-Signals

Md. Shoaib Chowdhury, PhD, PE, MASCE [1]

[1]Adjunct Associate Professor, Department of Civil Engineering, The City College of New York, 160 Convent Avenue, New York, NY 10031; PH (212)-650-7829; email: mchowdhury@ccny.cuny.edu

## ABSTRACT

Conventional treatment of left-turn at major arterial intersections is often found to be the major cause of poor arterial operations, especially under a heavy volume condition. New Jersey Jug-handle Intersection (NJJI), an unconventional intersection scheme, eliminates direct left turn and has been in place on NJDOT heavy-volume arterial roadways for quite a long time. The NJJI scheme separates left turn vehicles prior to or after crossing the intersection via a right side jug-handle (forward or reverse) ramp and then allow them to make turn onto cross-street via ramp terminus separated from the main intersection. By introducing jug-handle ramps at all four approaches of a four-legged intersection, four identical ramp termini can be constructed. These ramp termini can be signalized along with the main intersection and each of them can be operated on a two-phase signal. CORSIM, a micro-simulation model, has been utilized to evaluate and compare the operational performance of a forward ramp NJJI scheme (with and without signals at the ramp termini) with a conventional intersection (CI) scheme. The simulation study reveals that the NJJI scheme with pre-signals significantly improves traffic operations while increasing the intersection capacity. Other benefits of the NJJI scheme include reduction of vehicle stop time, fuel consumption and emission level.

## INTRODUCTION

Intersections in general and signalized intersections in particular are the major determinant of arterial capacities (*AASHTO, 2004*). Therefore, the key to improving arterial performance is to improve the operations of signalized intersections along its way. Conventional treatment of left-turn is generally the major cause of poor intersection operations, especially under a heavy turning volume condition. Therefore, traditional measures to improve intersection operations such as retiming signal, adding left turn phase and/or left/right turn bays, increasing storage lengths of left/ right turn bays etc. does not adequately increase intersection capacity due to the inherent limitations of conventional intersection. As a result, transportation Engineers have been considering various alternative measures including unconventional means to eliminate direct left turns at signalized intersections (*Reid, 2000; Reid and Hummer, 2000; Tabernero and Sayed 2005)*. NJ Jug-handle intersection (NJJI), an unconventional intersection scheme, has been in place on NJDOT arterial streets for quite a sometime. It has also been introduced in other states including Canada *( FHWA, 2004; Furtado, Tencha and Devos 2003)*.

NJ Jug-handle (NJJ) ramp separates turning vehicles at the right side of a roadway either in advance of the intersection (in case of a forward ramp), or after passing the intersection (in case of a reverse ramp) as a means of eliminating turning

maneuvers at the intersection. Generally, left and right turn vehicles at the ramp terminals are controlled by stop and yield signs, respectively (*Reid and Hummer, 2000*). However, under a heavy turning volume condition, it may require signalizing those ramp terminals to improve traffic operations and safety. A previous study has indicated that the NJJI scheme has potential to reduce the number of left turn collisions as it eliminates direct left turn at the intersection while reducing the number of potential conflicts as compared to a conventional intersection (*FHWA, 2004)*.

In suburban areas, where adequate space is available and arterials are congested, the NJJI scheme could improve arterial operations. In this paper, a forward ramp NJJI scheme (See Figure 1) with and without signals at the ramp termini (e.g., with and without pre-signals) is evaluated as an isolated intersection and its performance is compared with a conventional scheme by using CORSIM *(FHWA, 2006)* micro-simulation model.

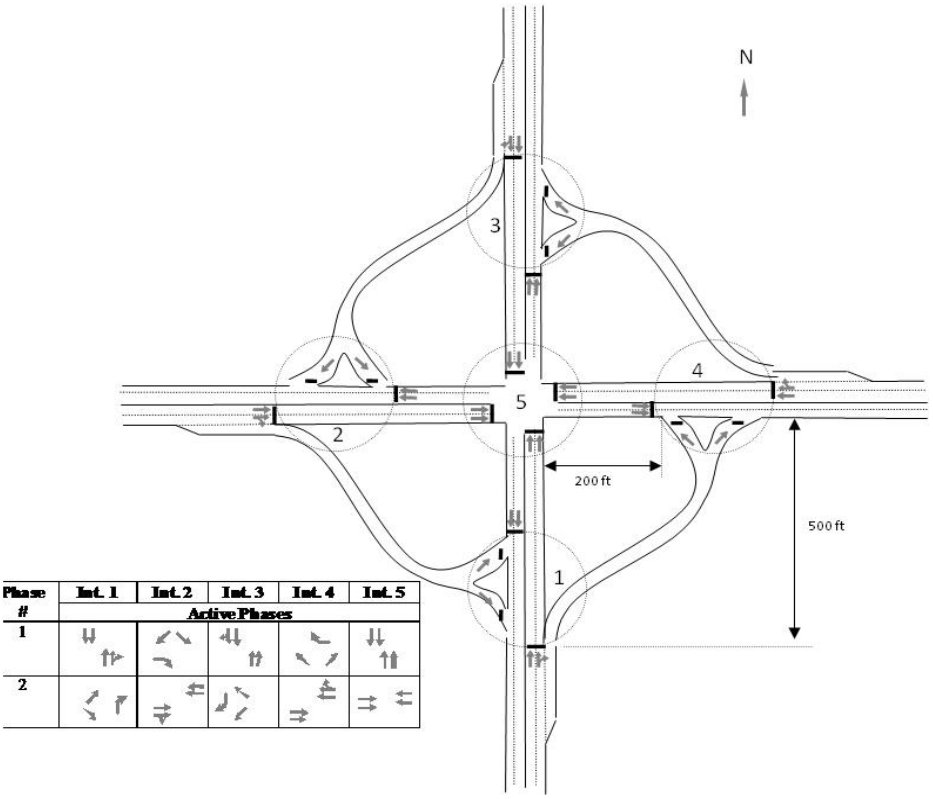



**FIGURE 1 NJJI with forward ramps**

## DESCRIPTION OF NJJI

NJDOT has defined the jug-handle as an "at-grade ramp that could be placed at or between intersections to permit the motorists to make indirect left and/or U-turns. Around the block designs that use interconnecting local street patterns to accomplish indirect left turns or U turns are not considered jughandles". NJDOT also described three types of NJJI namely -- forward ramp, reverse ramp and u-turn ramp and published their design standards *(NJDOT, 2010).* A forward ramp NJJI scheme, as illustrated in Figure 1, eliminates direct left and right turns at the intersection by separating them at the right side

upstream of the intersection via a ramp called "forward jug-handle ramp (FJR)". The left and right turn vehicles are then allowed to turn left/right onto cross street at ramp terminus separated from the primary intersection (Int. # 5 in Figure 1). It requires that storage length between the primary and secondary (Int. # 1, 2, 3 & 4 in Figure 1) intersections (ramp terminus) should be adequate to accommodate left turn vehicles that are generated during a signal cycle before processing them through the primary intersection. At an intersection between a major and minor street, FJR may only be required on the major street approaches depending on turning volumes. However, at an intersection of two major streets, FJR could be required on all four approaches.

## MODELLING AND DESIGN CONSIDERATIONS

CORSIM *(FHWA, 2006),* a traffic micro-simulation model developed by the Federal Highway Administration (*FHWA*), has been utilized to conduct simulation experiments, and evaluate and compare traffic operations between the NJJI and CI schemes. CORSIM is a time step based stochastic micro-simulation model and capable of simulating signalized intersections. Over the years, CORSIM has been successfully applied in the operations evaluations of both conventional and unconventional intersections *(Reid and Hummer, 2000; Hummer and Boone, 1995; Chowdhury, 2006a)*.

CORSIM consists of two micro-simulation models namely NETSIM and FRESIM. NETSIM model simulates surface streets and signalized intersections. Thus, NETSIM model capabilities are utilized in this study. The default calibration parameters of NETSIM including driver behavior (e.g., queue discharge headway and start-up lost time, gap acceptance at stop signs, gap acceptance for left and right turns etc) and vehicle performance parameters (e.g., speed and acceleration characteristics) are used in absence of field data. The default model parameters typically represent field condition and are reasonable for exploratory simulations and planning level alternative evaluations such as the one has been performed in this study. The following section describes geometric, traffic and operational characteristics of the intersection modeled.

### Geometric Considerations

As shown in Figure 1, the proposed NJJI scheme is a two-lane two-way four-legged arterial-to-arterial at-grade intersection consisting of FJRs at all four approaches. Thus, the simulated NJJI scheme consists of four identical ramp termini. The space between the primary intersection and the beginning of forward ramp is 500 ft, while the space between the primary intersection and the end of forward ramp is 200 ft. Each approach of the equivalent conventional intersection consists of two through lanes and one left and one right turn bays with storage lengths of 200 ft each. Each of the four approach links extends 1000 ft upstream of the intersection under both the NJJI and CI schemes to ensure that vehicle queue lengths do not exceed simulated network. It is to note here that no attempt has been made to optimize storage lengths between the primary and secondary intersections of the NJJI scheme and left/right turn bay of the conventional intersection at this planning level evaluation study. However, such lengths could be optimized given the turning volumes of respective movements.

**Traffic Considerations**

The traffic stream consists of cars and trucks. Two car types including high performance (with 75% of fleet component) and low performance (with 25% of fleet component) and a single unit truck (with 100% of fleet component) are considered. Truck traffic represents 5 percent of the traffic stream.

In order to evaluate the traffic operations under a wide range of demand and turning volume conditions, series of cases (e.g. nine volume levels and two turn percentages as discussion will follow) are designed. For simplifying the simulation experiments and systematically comparing the results, it is assumed that vehicular demands on all of the four intersection approach links are identical for a given case. Vehicles are released onto the network through the entry links at a uniform rate. To examine how different levels of demands may affect the intersection operations under each of the three schemes (e.g., conventional, NJJI without pre-signals and NJJI with pre-signals), each scheme is simulated repeatedly by gradually increasing the entry volume from a starting low volume of 500 vph and with an increment of 250 vph until the intersection reaches oversaturation condition. As it would be shown later, the over saturations of conventional and NJJI schemes with and without pre-signals reach at different volume levels. The range of input volumes varies between 500 vph and 2500 vph to ensure that the oversaturation conditions reach under all three schemes. Similarly, to examine how turning volumes may affect the intersection operations, two turn percentages including a low of 5 and a high of 15 are considered. Percentages of left and right turn volumes for a particular case are identical. Free flow speed on an approach link is 35 mph, while it is 25 mph on a jug-handle ramp.

**Signal Considerations**

It is assumed that the conventional intersection is controlled by a four-phase signal with leading left turn phases. The NJJI scheme with pre-signals is controlled by five two-phase-signals, one at the primary and the remaining four at four secondary intersections. The NJJI scheme without pre-signals is controlled by a two-phase-signal at the primary intersection, while left and right turn movements at each of the four ramp termini are controlled by stop and yield signs, respectively. Right turn on red is permitted where applicable.

As shown in Figure 1, phase one of the NJJI with pre-signals allows north-south bound movements at the primary (Int. # 5) and two secondary intersections (Ints. # 2 & 4) and phase two allows east-west bound movements at the primary (Int. # 5) and two other secondary intersections (Ints. # 1 & 3). During the two-phase signal cycle, left turn traffic experiences both the phases, while through and right turn traffic generally experience only one phase. For instance, in the first phase, north-south bound left turn traffic is loaded onto the east-west arterial between the primary and secondary intersections and in the subsequent phase, they are unloaded.

Since each simulation run under a given case considers identical entry volumes at all four approaches, therefore green time for each of the two phases is also identical. Several other reasonable and simplified assumptions are made in regards to cycle lengths for all the schemes and left turn phase for the CI scheme. CORSIM is a simulation tool and has no capability to optimize cycle length. Researchers have shown that the cycle

length that may be found optimal under a macroscopic model may not be the optimal under a micro-simulation environment, such as in CORSIM, due to the inherent differences between the micro and macro simulation models *(Stevanovic and Martin, 2006)*. In addition, for a given traffic scenario, a rigorous model calibration is also required to achieve consistent MOE's from both the micro and macro models. Therefore, rather than applying a macro-simulation based optimization tool, the cycle length is arbitrarily selected to 90 seconds for all cases under the conventional and NJJI schemes, which is a widely used cycle length in the real world under a similar roadway and traffic environments. It is to note here that under a low demand condition, cycle length would generally be shorter to avoid excessive delays, while under a high demand condition, it could be longer to minimize the lost time between phases. Amber and all-red intervals are assumed to be 3 and 2 seconds respectively. Under the conventional scheme, green split for the leading left turn phase has been achieved after performing some experimental trial and error CORSIM runs with variable green splits. Considering the approach volume and intersection geometry, it is found that 7 sec and 10 sec green times are reasonable for the leading left turns with 5 and 15 % turning volumes, respectively.

Considering nine different volume levels, two different turn percentages and three different schemes (e.g., NJJI with pre-signals, NJJI without pre-signals, and CI), a total of 54 simulation runs each with a duration of one hour has been performed to generate the results, where each simulation run captures intersection operations for 40 consecutive signal cycles.

## ANALYSIS OF RESULTS

After performing the simulation experiments for all the cases designed, intersection operations, energy consumption and emission related Measures of Effectiveness (MOEs) including approach capacity (throughput), control delay, speed, vehicle stop time, fuel consumption, and HC, CO and NO emissions are extracted from the model outputs and presented through a series of figures. Since all intersection approaches are identical from traffic, geometric and operations prospective for a particular case is concerned, therefore, the approach MOE of a particular case presented here is the average value of the MOEs of four approaches.

Figure 2 (a and b) shows the relationship between entry volumes and average intersection approach capacities (throughputs) considering two different turn percentages (e.g., 5 and 15%). It shows that as the input volume increases, the throughput also increases nearly identically for a certain range of volumes and thereafter, depending on the percentage of turning volumes, further increase of entry volumes does not increase throughput indicating that intersection has reached at its capacity. The approach capacities of the conventional intersection with 5 and 15 percent turning volumes are nearly identical and it is somewhere between 1550-1600 vph. Note that considering the stochastic nature of capacity variations especially before and after a flow breakdown, a difference of 50 vph is not considered a meaningful difference (*Chowdhury, 2006b)*. However, approach capacities of NJJI scheme with and without pre-signals under 5 and 15 percent turning volumes are found to be different. The approach capacity of NJJI scheme without pre-signal decreases roughly from 1900 vph to 1550 vph (approximately 18%) as turn percentage increases from 5 to 15. Contrary to NJJI scheme without pre-signals, the approach capacity of NJJI scheme with pre-signals increases roughly from

2100 vph to 2300 vph (approximately 10%) as turn percentage increases from 5 to 15. This is a very important finding and it can be explained as the demand including left turn volume increases under the NJJI scheme without pre-signals, left turn vehicles find it difficult to get acceptable gaps to turn them onto cross street at the ramp termini. At the same time the spaces between the primary and secondary intersections are mostly occupied by through traffic as their movements are not controlled and coordinated by signals at ramp termini while vehicles are discharged uniformly onto the network through entry links. This finding could have been different, if the subject intersection could be coordinated with other neighboring signals to ensure excellent progression of flow. Therefore, two important conclusions are that (1) isolated NJJI scheme without pre-signals may not perform satisfactorily under a heavy overall and turn volume conditions although it may perform efficiently under a low overall and turn volume conditions, (2) isolated NJJI scheme with pre-signals increases intersection capacity the most (roughly increases the intersection capacity by 45% over the conventional scheme) and thus performs the best under a heavy overall and turn volume conditions as compared to the other two schemes. Although safety analysis is not considered exclusively in this study, however, it is expected that the NJJI scheme with pre-signals will not only improve intersection operations and thus arterial operations significantly, but also probably will improve intersection safety by providing protected right-of-ways to both the vehicular and pedestrian traffic.

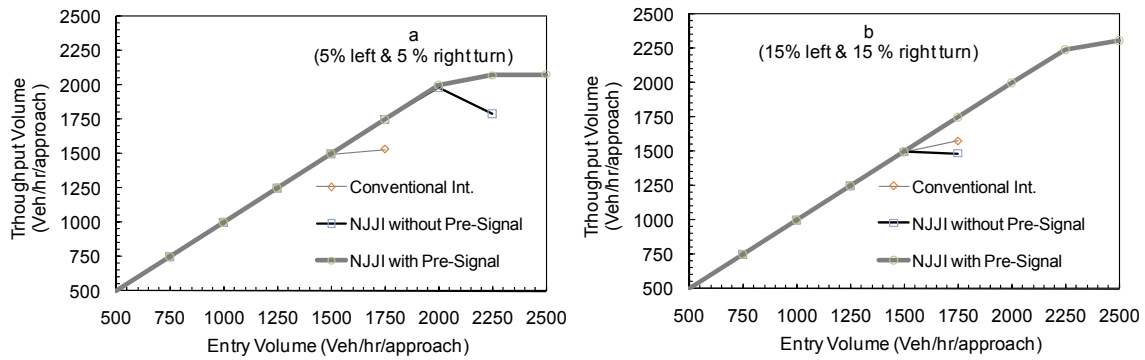



**FIGURE 2 Throughput Volume vs. Entry Volume.**

Figure 3 (a and b) shows the relationship between control delay and entry volume. Control delay is defined in the HCM 2000 *(Highway Capacity Manual, 2000)* as "the portion of the total delay for the vehicle approaching and entering a signalized intersection that is attributable to traffic signal operations. Control delay includes the delays of initial deceleration, move-up time in the queue, stops, and re-acceleration." It is found that control delays under both the NJJI schemes and the conventional scheme are relatively low and compatible for entry volumes up to 1500 vph. As the entry volume further increases (depending on percentage of turning volumes), control delay under the conventional scheme begins to increase sharply. As expected, control delay under the NJJI scheme without pre-signals is relatively low for entry volume up to nearly 1900 vph (which is also the approach capacity for this scheme) with a 5 percent turning volume. The control delay under the NJJI scheme with pre-signals is low even at or beyond the entry volume of 2000 vph depending on percentage of turning volume due to its higher

approach capacity. Therefore, it can also be concluded that the NJJI scheme with pre-signals has the potential to reduce delays and congestions on arterials that are suffered from extreme congestions while conventional intersections are in place.

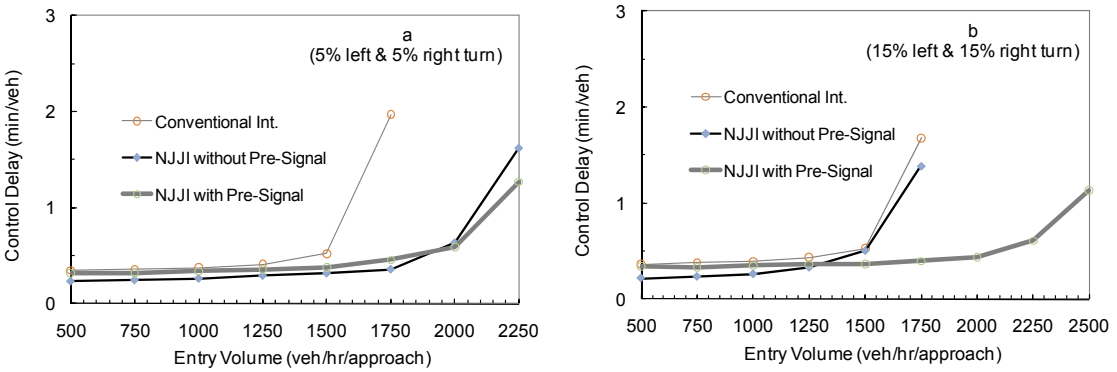



**FIGURE 3 Average Control Delay vs. Entry volume.**

This finding is further supported by average speed vs. entry volume relationship presented in Figure 4 (a and b). The figure shows that the NJJI scheme with pre-signals has the potential to maintain higher speed under a heavy overall and turning volume conditions, although the NJJI scheme without pre-signals maintains higher speed at low overall volume conditions (approximately at or below 2000 vph with 5 percent turning volumes and at or below 1250 vph with 15 percent turning volumes).

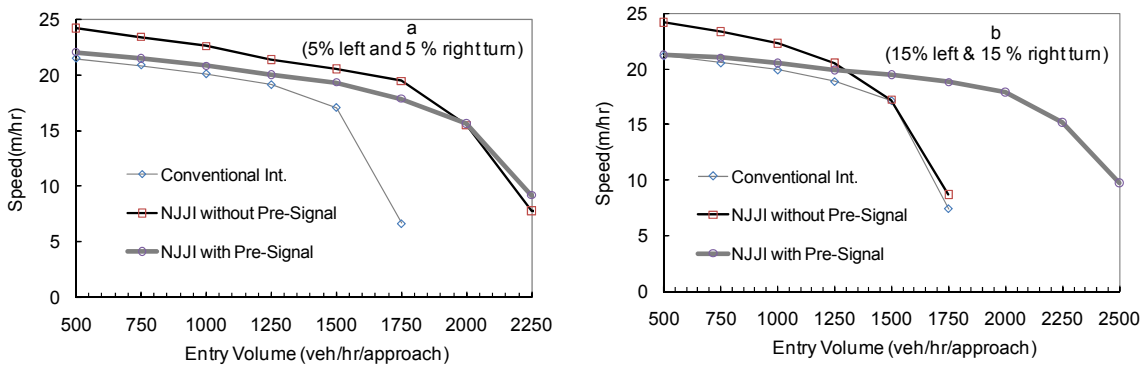



**FIGURE 4 Average Speed vs. Entry Volume.**

Figure 5 (a and b) shows that the capacity failure and speed reduction also increases average vehicle stop times. As can be seen, average vehicle stop time increases significantly under the conventional scheme with 5 percent turning volumes as compared to both the NJJI schemes as entry volume exceeds its capacity of 1550 vph. Under a 15 percent turning volume, both the conventional scheme and NJJI scheme without pre-signals perform poorly and compatibly as compared to the NJJI scheme with pre-signals. As explained before, this happens as turning vehicles suffer significant delays at ramp terminus and such delays propagate onto arterial streets.

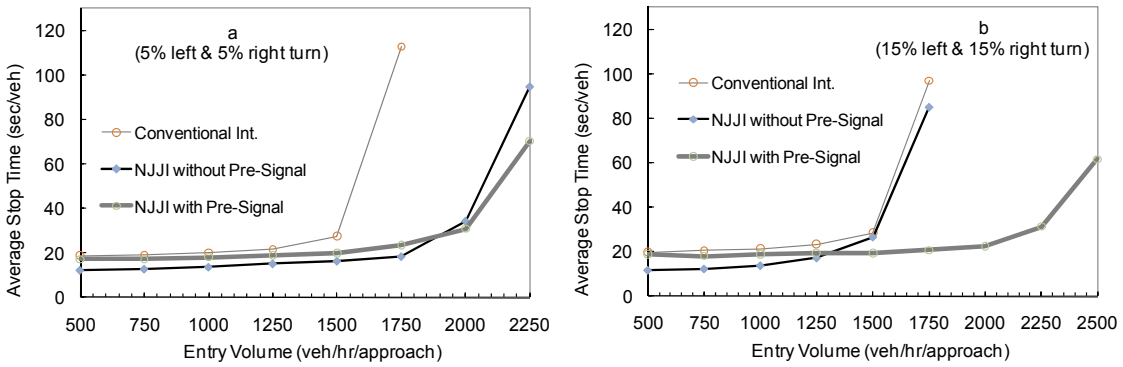



**FIGURE 5 Average Stop Time vs. Entry Volume.**

Vehicle speed, control delay and number of stops at the intersection all significantly affect the levels of vehicle emissions and fuel consumptions. As control delay and number of stops increase, average speed decreases, which eventually increases vehicle emission. Studies (*CTE/NCDOT, 2001*) have found that accelerating from a stop at the intersection produces highest vehicle emission. As delay and stop increase at the intersection, HC emission also increases. Considering how fuel consumptions and vehicle emissions may differ under the three schemes modeled at or near the capacity (e.g., for entry volume of 1500 vph) of conventional intersection, fuel consumptions (mile/gallon) and vehicle emissions (grams/mile) including HC, CO, and NO are extracted from the CORSIM outputs and average values combining 5 and 15 percent turning volume cases are computed. It is to note here that CORSIM predicts emissions and fuel consumptions based on a function of vehicle acceleration. As shown in Figure 6, fuel consumption as well as HC, CO, and NO emissions are higher under the conventional scheme as compared to the NJJI schemes with and without pre-signals. For instance, with a gallon of fuel, in an average, a vehicle can travel 13.8, 14.7 and 14.8 miles under the conventional, NJJI scheme without pre-signals and NJJI scheme with pre-signals, respectively. It is worth mentioning here that the fuel efficiencies of various schemes reported are approximate values and only applicable for a planning level comparison purpose.

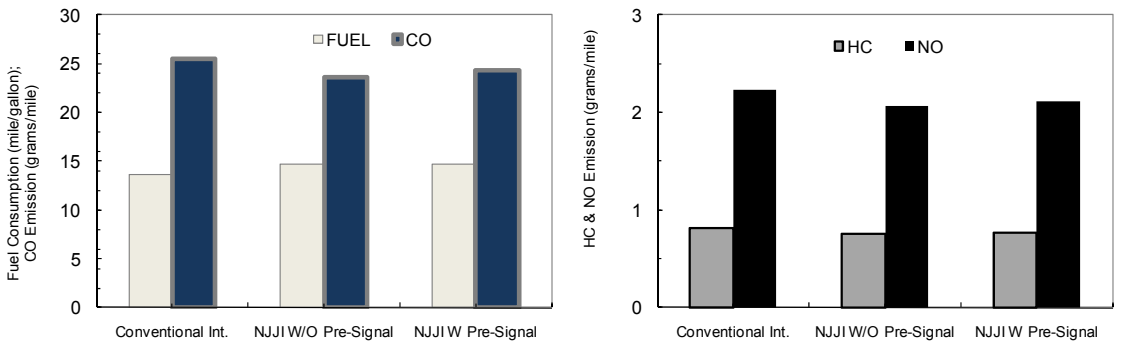



**FIGURE 6 Average Fuel Consumption and Vehicle Emissions.**

## CONCLUSIONS

In this paper, NJJI schemes with and without pre-signals have been evaluated against a conventional signalized scheme on a arterial of four-legged at-grade intersection for a wide range of volume conditions using CORSIM micro-simulation model. It is found that the NJJI scheme with pre-signals can improve arterial-to-arterial intersection operations under heavy through and turning volume conditions, while without degrading the quality of service under low volume conditions such as during off peak periods. It is also found that the NJJI scheme without pre-signals performs slightly better than the NJJI scheme with pre-signals and conventional scheme under low overall volume conditions. However, as overall volumes including turning volumes increase, NJJI scheme without pre-signals fails to perform satisfactorily. As demand increases and reaches at or beyond the capacity of conventional as well as NJJI scheme without pre-signals, signal delay increases significantly. In such a situation, with demonstrated higher capacity over both the conventional and NJJI scheme without pre-signals, NJJI scheme with pre-signals performs better by processing more traffic through the intersection, reducing vehicle delays and increasing speeds. The specific finding is that NJJI scheme with pre-signals increases intersection capacity by some 45 percent over the conventional one. Moreover, the NJJI scheme with pre-signals reduces fuel consumptions and vehicle emissions over the conventional one.

Further study should focus on examining the impact of space between the primary and secondary intersections on intersection operations; evaluating the performances of NJJI schemes under variable directional demand and lane configuration conditions. Quantifications of benefits of the NJJI schemes from arterial system operations prospective under coordinated signal conditions are also important.

## REFERENCES

AASHTO (2004). "A Policy on Geometric Design of Highways and Streets." Washington D.C.

Chowdhury, Md. S. (2006a). "Proposing and evaluating a modified Jug-handle Scheme for Arterial-to-arterial At-grade Signalized Intersections." *Unpublished working paper # 2006-1*, Jersey City, NJ

Chowdhury, Md. S. (2006b). "Examining the Operational Benefits of Freeway Ramp Metering Control Using CORSIM (CORridor Simulator)." *9th International Conference on Applications of Advanced Technology in Transportation*, ASCE/T&DI, Chicago, Illinois

CTE/NCDOT (2001), "Emissions Reduction Through Better Traffic Management: An Empirical Evaluation Based Upon On-Road Measurements," *Report No. FHWY/NC/2002-001*, North Carolina State University, Raleigh, NC

FHWA. (2004). "Signalized Intersections: Informational Guide." *Publication No. FHWA-HRT-04-091*

FHWA (2006). "CORSIM user and reference manuals." *Version 6.0*

Furtado. G., Tencha, G. and Devos, H. (2003). "Unconventional Arterial Design: Jughandle Intersection Concept for McKnight Boulevard in Calgary." *Annual Conference and Exhibition*, The Transportation Association of Canada, City of St. John's, Canada

Highway Capacity Manual. (2000). "HCM2000." *Transportation Research Board,* National Research Council, Washington, D.C.

Hummer, J., and J. Boone. (1995). "Calibrating and Validating Traffic Simulation Models for Unconventional Arterial Intersection Designs." *Transportation Research Record: Journal of Transportation Research Board, No 1500*, TRB, National Research Council, Washington, D.C., pp. 184-192.

NJDOT. (2010). "Highway Design Manual." (http://www.state.nj.us/transportation/eng/documents/RDM/), accessed June 20, 2010

Reid, J. (2000). "Using Quadrant Roadways to Improve Arterial Intersections Operations." *ITE Journal,* 70 (6), pp. 34-45.

Reid, J. and J. Hummer. (2000). "Travel Time Comparisons between Seven Unconventional Arterial Intersections." *Transportation Research Record: Journal of Transportation Research Board, No. 1751*, TRB, National Research Council, Washington, D.C., pp. 56-66.

Stevanovic, A., and P. Martin. (2006). "Evaluating the Reliability of Macroscopically Optimized Timing Plans Through Microsimulation." *CD-ROM, The 85th Annual meeting of the Transportation Research Board of National Academics*, Washington, D.C.

Tabernero, V., and T. Sayed. (2005). "Introduction and Analysis of a new unconventional Intersection scheme, the Upstream Signalized Crossover (USC) intersection." *CD-ROM, The 84th meeting of the Transportation Research Board of National Academics*, Washington, D.C.

# Non-Standard Roundabouts: Analysis of Problems

F. Pinna[1] and C. Piras[2]

[1] Assistant Professor – Civil Engineer, Department of Land Engineering, University of Cagliari, Piazza d'Armi 19 - 09123 Cagliari (ITALY), Phone-number: +39-070-6755256, Fax: +39-070-6755266, e-mail: fpinna@unica.it

[2] PhD – Civil Engineer, Department of Land Engineering, University of Cagliari, Piazza d'Armi 19 - 09123 Cagliari (ITALY), Phone-number: +39-070-6755256, Fax +39-070-6755266, e-mail: cpiras@unica.it

## ABSTRACT

Road intersections are dangerous places in the road system due to the high number of conflicting points between motorized and non-motorized vehicles. In the case of defined traffic volume, several research groups have proven that roundabouts reduce the number of injuries and fatal accident cases.
In the last years, several studies have investigated the performance of roundabouts, among which, include different models with the ability to calculate the entering flow (Qe) as a function of the circulating flow (Qc).
This research work has focused on an analysis of the performance features of Italian urban roundabouts, that have non-standard geometric characteristics.
The goal of the research work is to analyze, on the basis of geometrical and design parameters, the capacity of the "non-standard roundabouts" in order to understand in which way non-standard features influence capacity. The study underlines that in standard roundabouts the entering flow is reduced when the circulating flow increases (like in the existing international models) and in non-standard roundabouts the entering flow is not reduced when the circulating flow increases (it seems that there is not a relationship between circulating flow and entering flow).

## INTRODUCTION

The number of roundabouts has been increasing, mainly due to their exceptional performance. The use of roundabouts has also significantly increased in Italy (which is the case in many other European countries) in recent years. The majority of them are single-lane and two-lane roundabouts, and used in rural and urban areas. Roundabouts, if properly designed, provide better safety performance than other conventional intersections for certain traffic flow conditions and roadway geometries. For this reason, many intersections have been converted into roundabouts. This happens even if often the available space isn't enough. For this reason, there are a lot of non-standard roundabouts, for example with one of legs without deflection.

There is not sufficient literature available to compare the performance of standard roundabouts and non-standard ones, in order to evaluate, for example, if the performance changes from conventional to non conventional roundabouts.

For this reason this paper focuses on non-standard roundabouts in order to verify how non-standard features influence roundabout capacity, especially in terms of relationship between circulating flow and entering flow. The performance of non-standard roundabouts are analyzed and compared to that of standard roundabouts using the existing models that explain the relationship between entering flows and circulating flows.

## EXISTING MODELS

There are several models in the literature which calculate the relationship between Qe and Qc for conventional roundabouts. Some of them are considered in this paper in order to compare our results with their findings.

The NCHRP Report 572 proposed an exponential regression model, which is actually a gap-acceptance model, for single-lane and two-lane roundabouts. Entry capacity at single-lane roundabouts is expressed by the following equation [9]:

$$C = 1130 \cdot \exp^{(-0.001 \cdot Vc)} \tag{1}$$

Entry capacity at two-lane roundabouts is shown by the following formula [9]:

$$C = 1130 \cdot \exp^{(-0.0007 \cdot Vc)} \tag{2}$$

where:

C = entry capacity [pcu/h]

Vc = conflicting flow [pcu/h]

A procedure in the Highway Capacity Manual (HCM) indicates that the capacity of a roundabout can be estimated by using gap acceptance techniques with the basic parameters of the critical gap and follow-up time. The procedure underlines that there has generally been the assumption that the performance of each leg of a roundabout can be independently analyzed from the other legs and consequently,

most techniques tend to use the information from only one leg [7]. In the HCM, the approach capacity is estimated by using the following formula [7]:

$$C_a = \frac{v_c e^{-v_c t_c / 3600}}{1 - e^{-v_c t_f / 3600}} \quad (3)$$

where:
$C_a$ = approach capacity [veic/h]
$v_c$ = conflicting flow [veic /h]
$t_c$ = critical gap [s]
$t_f$ = follow-up time [s]
The recommended ranges of values of the critical gap and follow-up time are:

- upper bound: critical gap = 4.1 s and follow-up time = 2.6 s;
- lower bound: critical gap = 4.6 s and follow-up time = 3.1 s.

One of the German methodologies which calculates Qe as a function of Qc is expressed by a linear equation, which is [5]:

$$Q_e = C + DQ_c \quad (4)$$

$Q_e$ = entry flow capacity [veic/h]
$Q_c$ = circulating flow rate [veic /h]
C and D depend on the number of entry and circle lanes. For this study, we consider that:

- C = 1,218 and D = -0.74 (number of lanes entry/circle → 1/2);
- C = 1,380 and D = -0.50 (n° of lanes entry/circle → 2/2).

Finally, there are also other similar models, such as the those proposed in Switzerland, in France (SETRA), the United Kingdom (UK), and so on [1], [2], [3]. Some of them, such as the UK models, have a trend that is much higher than others and therefore, it is also out of our data set range. To keep this document at a reasonable length, such models are not mentioned in the following parts of the paper. All of the mentioned models, which are built by a statistical regression among the local collected data, shown in Figure 1, express the Qe as a function of Qc. The six models are quite similar for low Qc values (the minor differences can probably be attributed to different driver behaviors in different countries), in contrast to high Qc rates where the differences are more evident.

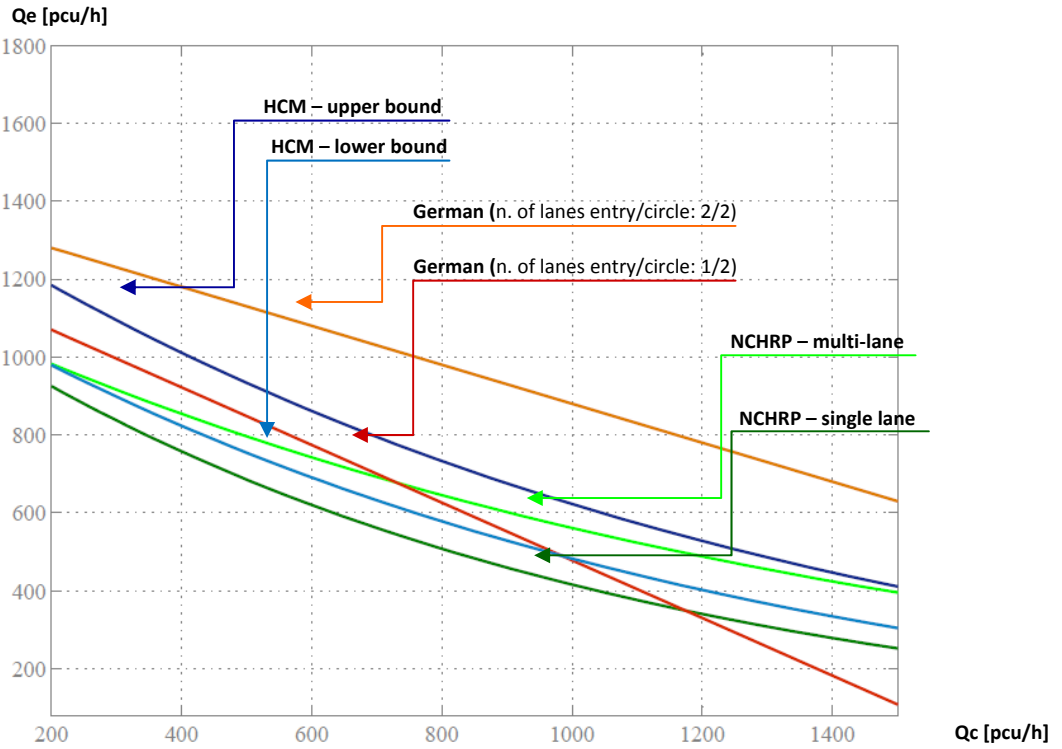



**Figure 1 A comparison of some of existing roundabout models**

## NON-STANDARD ROUNDABOUTS

In this study only non-standard roundabouts are considered, in which of them standard and non-standard legs are assessed. A roundabout is non-standard when it has one or more characteristics that don't respect design rules, for example, one of legs without deflection.

Not tangential entries allow to improve the visibility of the intersection for the approaching driver and causes reduced speed. [5]

In this regard, in the first part of the research work also speed data were collected. But in the case of study the collected speed data were low because of, during the rush hours, the speed on the studied roundabouts are influenced by other factors, like pedestrians, close accesses and so on, and not only by geometry or roundabout characteristics. In fact every roundabout, took into consideration, is an urban roundabout.

Deflection is very important in order to remove conflicts at merging points and to accommodate slow entering vehicles. The lack of deflection in the case of study is the strongest non-standard feature.

Generally the chosen roundabouts have the following geometric features:

- outer diameter between 26 m and 40 m;
- one single lane or two lanes on each of the entries;
- one single lane or two lanes on each of the exits;
- circle lane width: more than 7 m (with one or two lanes with lane marking).

## CASE OF STUDY

For the study, traffic volume data were collected for each leg of several roundabouts. Then, a procedure to estimate flow which consists of converting all traffic flows into passenger car units (pcu), in order to take into account the influence of heavy vehicles, was adopted. The empirical data were collected during rush hour. The number of entering vehicles and circulating vehicles was observed.
Empirical data from observed circulating and entry flow at non-standard and standard legs have been used also to test different capacity models.
The empirical data are collected during the rush hour; the number of entering vehicles and the number of circulating vehicles has been observed. For each time period, the observed traffic volumes have been converted to passenger car units. A total of more than 400 observations are registered.
The following figures show the empirical data respectively for standard legs and non-standard legs and both of them. Different colors represent different kind of legs. In each figure some of international similar models and our respectively best fits are showed. For these a confidence level of 95% is selected as the critical limit and the null hypothesis is rejected for a lesser than 0.005.

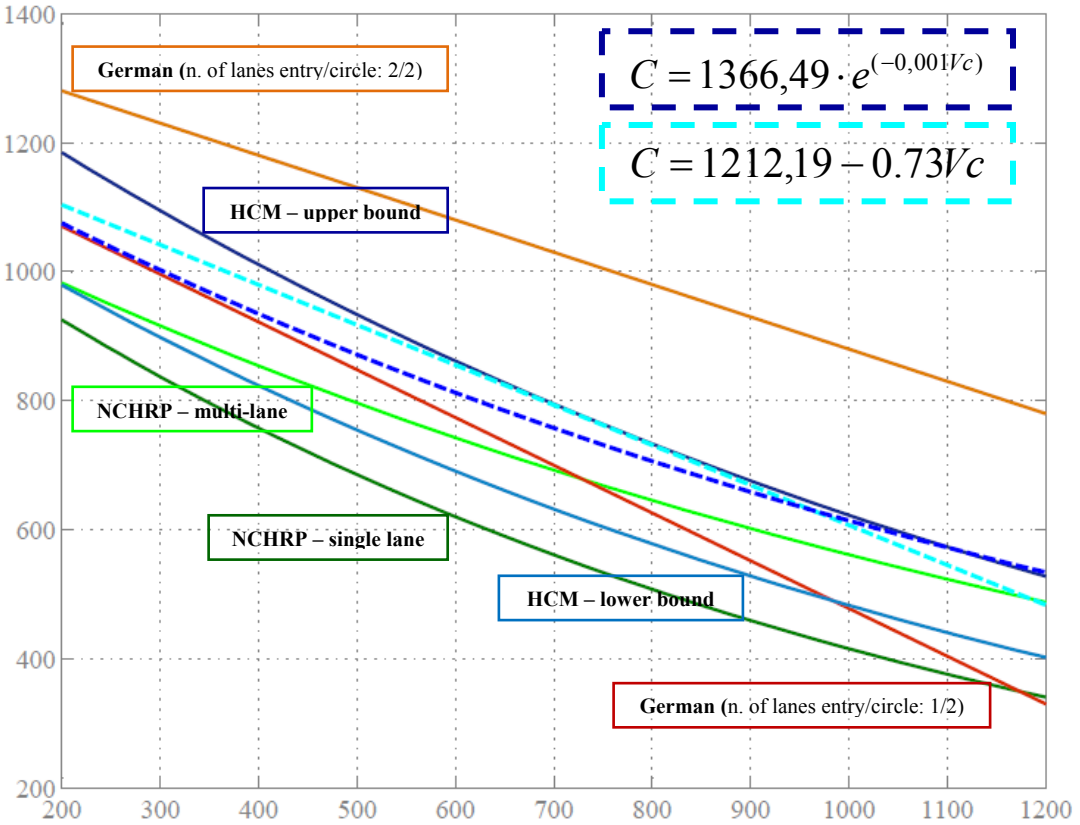



**Figure 1 - A comparison of roundabout models and models for standard legs**

In our study, the best fit is described by an exponential function. The model that expresses this relationship is as follows:

$$C = 1366{,}49 \cdot e^{(-0{,}001Vc)} \tag{5}$$

where:
C = entry capacity [pcu/h]
Vc = conflicting flow [pcu/h]
The collected data have also been fitted using a linear regression, the equation is:

$$C = 1212{,}19 - 0.73Vc \tag{6}$$

where:
C = entry capacity [pcu/h]
Vc = conflicting flow [pcu/h]
Figure 2 shows the trends of linear and exponential models that are built by using classical statistical techniques: a comparison between international models and our statistical model is so possible. The standard roundabout legs in our statistical model show a similar relationship between Qc and Qe with the existing models. The slight differences can be attributed to the different behaviors of drivers in different countries. The trend of the last model appears to be the bisecting the angle subtended by the two German models.

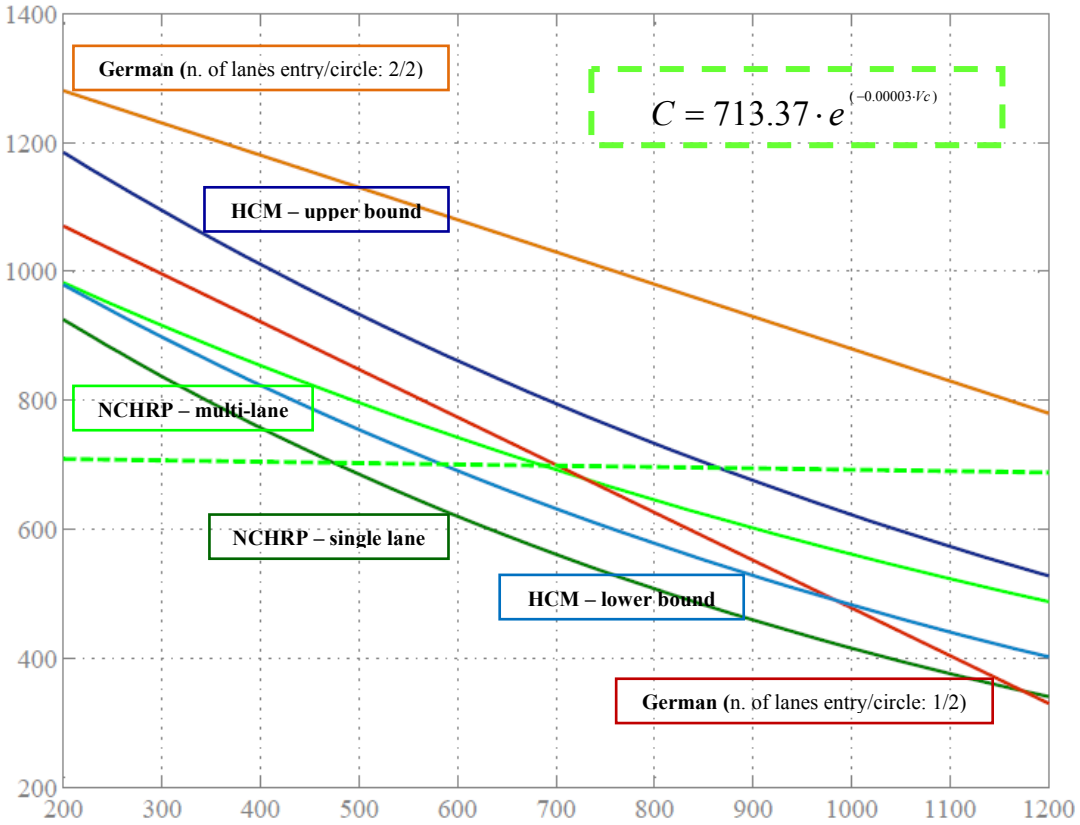



**Figure 2. A comparison of roundabout models and models for no-standard legs**

Instead the model for non-standard legs is:

$$C = 713.37 \cdot e^{(-0.00003 \cdot Vc)} \quad (7)$$

where:
C = entry capacity [pcu/h]
Vc = conflicting flow [pcu/h]
The graphical results are very strange. In fact, for low volume conditions defining a specific model appears no correct. It seems that the environmental conditions influence the relationship between circulating and entering flow. Differently from standard legs or from the international models the graphic shows that entry flow stay nearly constant for each circulating flow value. In fact for low and high volume conditions (of circulating flow) non-standard legs produce constant (low) entering flow values. It is very strange. The entering flow in the case of legs tangent should have been higher than which of standard legs. That doesn't happen.
Above all the correlation coefficient is very low, it appears that geometric features don't influence the flow values which seems linked to local characteristics. That is underlined also from the high data dispersion.
Entering flow values are constant also for high circulating flow values because of the forcing the entering maneuvers. In fact non-standard geometry, i.e. tangent legs, allows the drivers coming from tangent leg to force the entering then circulating flow stops in order to give way to entering flow.
In order to have a model that takes into account the variations in capacity due to circulating flows, with data collected on standard and non-standard legs, the equation below have been developed.

$$C = 812.41 \cdot e^{(-0.0002 \cdot Vc)} \quad (8)$$

where:
C = entry capacity [pcu/h]
Vc = conflicting flow [pcu/h]

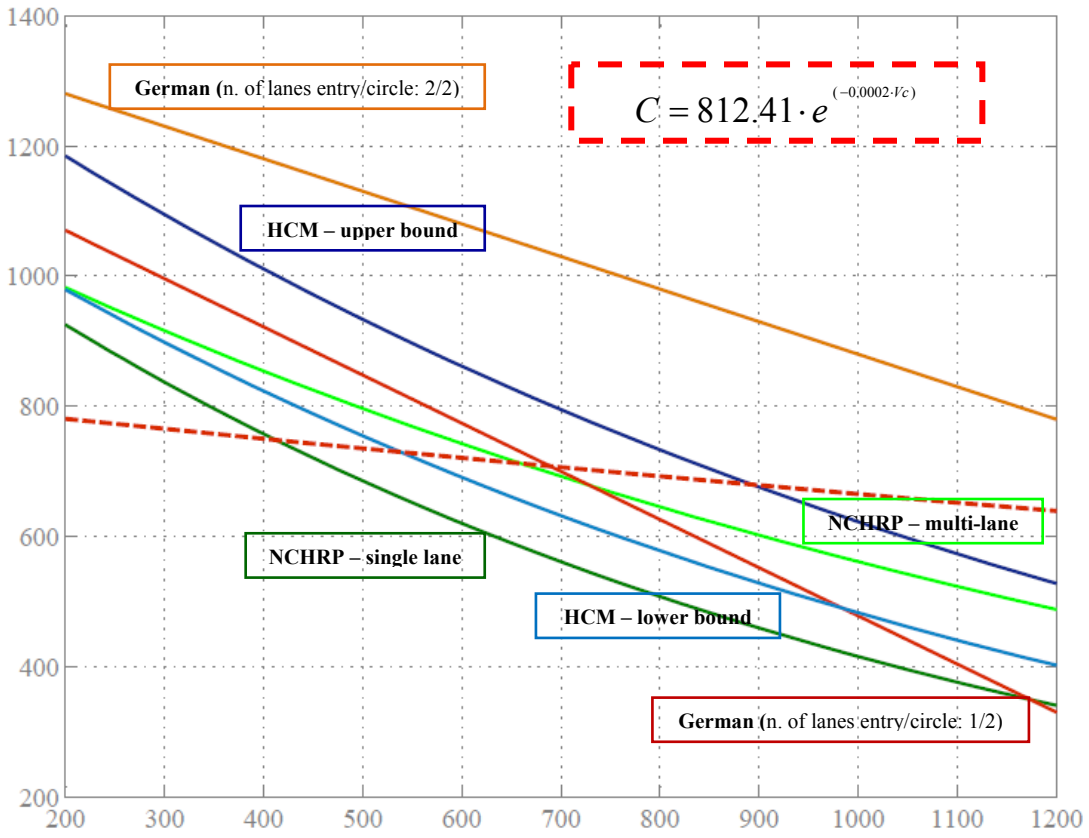



**Figure 3. A comparison of roundabout models and models for standard and non-standard legs**

The figure shows that, for low circulating flow values, the entering flow values is lower than in the previous cases. In fact the fitting is disturbed by the data collected on non-standard legs.

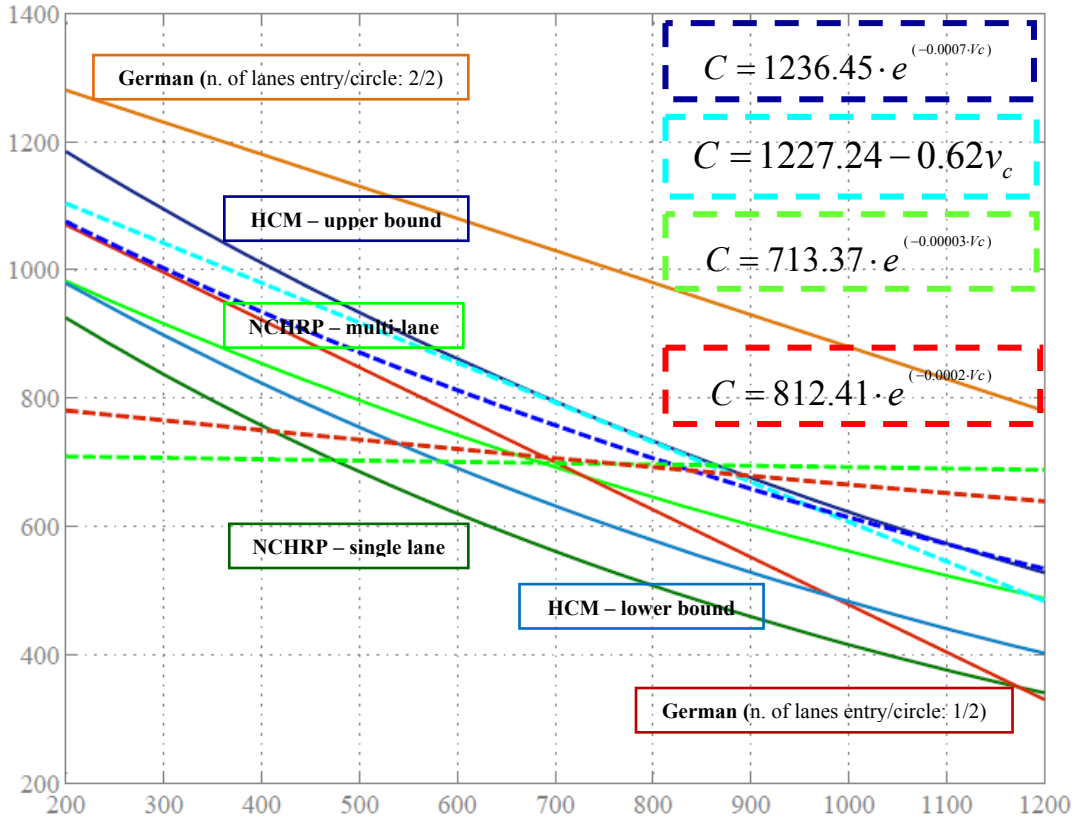



**Figure 4. A comparison of the several mentioned roundabout models**

In Figure 5 it is possible to observe that the relationship between circulating flow and entry flow for non-standard legs is lower than standard ones.
The model in the case of non-standard legs seems to underestimate the capacity at low circulating traffic flows and overestimate the capacity at high circulating traffic flow. Other empirical capacity models have been tested on the empirical data set.

## CONCLUSIONS

The first model is built on data collected on standard legs belonging to non-conventional roundabouts. In this case the trend is very similar to which of the other international models. So is possible to say that roundabout more or less uniforms driver behavior.
The second model is totally anomalous. Further our studies aim to analyze the influence of these non-standard geometric features (i.e. tangent legs) on safety levels.
Finally the trend of the third model is obviously disturbed by the data collected on no-standard legs.
It should be noted that:

- in standard legs the entering flow is reduced when the circulating flow increases (like in the existing international models);

- in non-standard legs the entering flow is not reduced when the circulating flow increases (it seems that there isn't a relationship between circulating flow and entering flow).

Moreover during data collection it has been noted that, even if the traffic entering the circle has to yield to traffic already in the circle, that does not always happens. It was found that in some of non-standard legs for high circulating and entering flow values circulating vehicles give way to entering vehicles. In this way roundabout works like a simple intersection. This happens especially when the legs are tangents: the circulating flow stops in order to give way to entering flow from tangent leg.

The work is still in progress, further studies are needed to verify and improve these models in order to understand which non-standard features influence the roundabout entering capacity and in which way this could happens. Other data are being collected on non-standard and standard roundabouts. In fact all of this study has been concentrated on non-standard roundabouts, making the comparisons only with international models. It should be interesting also making the same comparison with data collected on Italian standard roundabouts, in order to better understand what happens.

The lack of deflections weighs greatly on roundabout performance. Further studies are being done in order to estimate also the influence of the tangential entries on vehicles speed.

## REFERENCES

[1] Akcelik, R., (1997), *Lane-by-lane modeling of unequal lane use and flares at roundabouts and signalized intersections: the SIDRA solution*, Traffic Engineering and Control.

[2] Akcelik, R., (2006), *SIDRA Intersection User Guide*, 2006, Victoria, Australia.

[3] Akcelik, R., (2009), *Evaluating Roundabout Capacity, Level of Service and Performance*, ITE 2009 Annual Meeting, San Antonio, Texas, August 9-12, 2009.

[4] American Association of State Highway and Transportation Officials, (2001), *A Policy on Geometric Design Highway and Streets*, AASHTO, Washington D.C., 2001.

[5] Brilon, W., (2005), *Roundabouts: A State of Art in Germany*, paper presented at the National Roundabout Conference, Vail, Colorado, May 22-25, 2005.

[6] Greibe, P., La Cour Lund, B., (2010), *Capacity of 2-lane Roundabouts*, 4th International Symposium On Highway Geometric Design, TRB, Valencia, Spain, June 2-5, 2010.

[7] *Highway Capacity Manual*, (2000) TRB National Research Council, Washington D.C., 2000.

[8] Italian Rules About Intersections, (2006), D.M. 19/4/2006, INTERSEZIONI. NORME SULLE CARATTERISTICHE FUNZIONALI E GEOMETRICHE DELLE INTERSEZIONI STRADALI, Ministero delle Infrastrutture e Trasporti, 2006.

[9] NCHRP (National Cooperative Highway Research Program), (2007), *Roundabouts in the United States,* Report n. 572, Washington D.C., Transportation Research Board, National Research Council.

[10] Polus, A., Shifthan, Y., Shmueli-Lazar, S., (2005), *Evaluation of the Waiting-Time Effect on Critical Gaps at Roundabouts by a Logit Model*, European Journal of Transport and Infrastructure Research (EJTIR), n. 5, pp1-12.

[11] Sabyasachee, M., (2009), *Traffic Flow Characteristics Comparison between Modern Roundabouts and Intersections*, ITE.

# Exploring Impacts of Factors Contributing to Injury Severity at Freeway Exit Ramps

Zhenyu Wang[1], Jian J. Lu[2], and Zhao Zhang[3]

[1]Department of Civil and Environmental Engineering, University of South Florida, 4202 Fowler Ave, ENB118, Tampa, FL 33620; PH (813)974-8727; FAX (813)974-2957; email: zwang9@usf.edu
[2]Department of Civil and Environmental Engineering, University of South Florida, 4202 Fowler Ave, ENB118, Tampa, FL 33620; PH (813)974-5817; FAX (813)974-2957; email: lu2@usf.edu
[3]Department of Civil and Environmental Engineering, University of South Florida, 4202 Fowler Ave, ENB118, Tampa, FL 33620; PH (813)974-8727; FAX (813)974-2957; email: zzhang7@usf.edu

## ABSTRACT

Several factors of freeway mainlines, exit ramp sections, and cross-roads, such as geometric design, environment, vehicle/driver characteristics, and traffic features, may have significantly impacts on crash injury severity at freeway exit ramps. Freeway exit ramps are defined as the sections from the painted nose to the ramp terminal, and provide limit accesses from freeway mainlines to cross-roads. To explore the impacts and address the significant factors, injury severity prediction models was developed in this paper using three regression technologies: Ordered Probit, Partial Proportional Odds, and Heterogeneous Probit. The analysis results showed that ramp configuration (Diamond, Free-flow Loop, Outer Connection, and Parclo Loop) has significant impacts on crash injury severity at exit ramp sections. Diamond ramp is safer than other configurations while Outer Connection seems to increase crash injury severity than other ones. Other significant factors include lane configuration, channelization at ramp terminals, distance between ramp terminal and the first intersection on crossroads, speed limit on ramps and mainlines, shoulder width on ramps, and alcohol/drug involvement.

## INTRODUCTION

Exit ramps, the only control accesses used for exiting traffic on freeways, serve as transitions from freeways to secondary crossroads which could be freeways, major or minor arterials, or local streets. There are two concerns in the exit ramp safety study: freeway diverge area, which is defined as a specific segment that separates exiting traffic from through volume on freeway mainlines; and exit ramp section, which is defined as a section from the painted nose to the ramp terminal providing limit accesses from freeway mainlines to secondary crossroads. Several factors, such as roadway design, traffic control, environment, traffic volume, and driver/vehicle characteristics, may significantly impact the injury severity on exit ramps. Thus, to understand the impacts is a key step towards the implementation of

successful designs and/or countermeasures at exit ramp sections to improve safety performance and to obtain benefits on society, economics and cultures.

Several previous studies have been conducted to evaluate the safety performance of exit ramp sections. Lord and Bonneson *(1)* calibrated predictive models for different ramp configurations (diagonal ramps, non-free-flow loop ramps, free-flow loop ramps, and outer connection ramps) at 44 selected sites. It could be found that the non-free-flow (parclo flow loop) ramp experienced twice as many accidents as other types of ramps. Bauer and Harwood *(2)* developed a Negative Binomial regression model on the entire ramp section and concluded that diamond ramp have slight less crash frequency comparing to other ramp types when other influential variables remain constant. Khorashadi *(3)* used another method known as ANOVA test to forecast the relationship among ramp configurations, geometry parameters and crash frequencies. This study found that the geometric elements had much weaker impacts than the ramp configurations. McCartt et al. *(4)* examined 1,150 crashes occurring on heavily traveled urban interstate ramps in Northern Virginia. The three major common crash types, run-off-road, rear-end, and sideswipe, accounted for 95% of total crashes. The countermeasures suggested in the study included increasing ramp design speed, increasing curve radii, installing surveillance systems such as detectors, cameras, and advanced message signs. Abdel-Aty and Huang *(5)* explored an origin-destination survey to customers on the central Florida's expressway system. The distance traveled to exit a ramp did not depend only on the spacing between ramps, but also on other factors, such as the trip purpose, vehicle occupancy, driver's income level, and E-Pass implementation when the vehicle was equipped with an electronic toll collection system. A main finding was that the guide signs beyond the expressway exits had an important impact not only on unfamiliar travelers but also on the experienced drivers. Hunter et al. *(6)* conducted field observations on speed relationships between ramps and freeways by videotaping. Notable conclusions were drawn that vehicle speeds on exit ramps were much higher than the post speed limit. Because of the difference between the ramp post speed limit and operating speed, unfamiliar drivers might slow down speed while familiar drivers might enter the exit ramp at a high speed exceeding the limit speed. That might be a vital reason why rear-end crashes account a large percent of crashes in the ramp sections. Some studies focused on the connections between different influential factors which could be the ramp volumes, configurations, crashes, curvatures, and so on. Hunter et al. *(7)* concluded that ramp design speed should larger than 50% of freeway speed. This conclusion accommodated to Hunter et al.'s *(6)* result that operating speed on the exit ramp is higher than the post speed limit.

Other previous studies focused on the safety performance of freeway diverge areas. In 1998, Bared et al. *(8)* developed Poisson Regression models to estimate the crash frequency for the deceleration lanes plus the entire ramps as a function of ramp AADT, mainline freeway AADT, deceleration lane length and ramp configurations. Later, Sarhan et al. *(9)* designed an approach to help achieving the optimum predictive models that are related to the length of acceleration and deceleration lanes based on expected collision frequency. Garcia et al. *(10)* analyzed different deceleration lengths as functions of exit trajectory types, speeds, and localization. Cassidy et al. *(11)* noticed the problem that queuing from the segment's off-ramp

spilling over and occupying its mandatory exit lane comes up frequently. The situation delayed the mainline vehicles as well and would increase weaving conflicts. Janson *(12)* examined the relationship of ramp designs and truck accident rates in Washington State plus a comparison to limited data from Colorado and California. The results showed that loop ramps in particular have generally higher accident rates, particularly rollovers. In 2008, Chen et al. *(13)* compared crash frequency, crash rate, and crash severity between different exit ramp types which were defined as single-lane exit ramps (Type I), single-lane exit ramps without a taper (Type II), two-lane exit ramps with an optional lane (Type III), and two-lane exit ramps without optional lane (Type IV). And a crash predictive model was developed to identify the factors that contribute to the crash and to quantify the safety impacts of different freeway exit ramps. The crash data analysis results suggested that the Type I exit ramp has the best safety performance in terms of the lowest crash frequency and crash rate. In 2009, Wang et al. *(14)* developed a predictive model to identify the factors that contribute the injury severity at freeway diverge areas. The analysis results showed that the factors include length of deceleration/ramp lanes, curve/grade at diverge areas, light and weather condition, alcohol/drug involvement, heavy vehicle involvement, number of lanes on mainlines, AADT on mainlines, surface condition, land type, and crash type. Furthermore, it also could be concluded that exit ramp types (single-lane exit ramps, single-lane exit ramps with a taper, two lane exit ramps with an optional lane, and two-lane exit ramps without optional lane) have no significant effects on injury severity at freeway diverge areas.

Several statistics methods were utilized to analyze the crash injury severity. Because the injury severity, describing the most serious injury level to any person involved in a crash, is a nature ordinal variable, ordered multiple choice regressions were widely used to fit injury severity data. O'Donnell and Connor *(15)* used two econometric models, Ordered Logit models (OL) and Ordered Probit models (OP), to estimate the linkages between eleven road attributes and the probabilities of sustaining four different levels of injury. Zajac and Ivan *(16)* adopted OP models to evaluate the effect of roadway and area type features on the injury severity of pedestrian crashes in rural Connecticut. Kockelman and Kweon *(17)* examined the risk of different injury levels sustained under all crash types, two-vehicle crashes, and single-vehicle crashes by using OP models. Abdel-Aty and Keller *(18)* applied OP models to explore the factors that contribute to injuries at signalized intersections. Abdel-Aty *(19)* compared OP, Multinomial Logit models (MNL), and Nested Logit models (NL) for roadway sections, signalized intersections, and toll plazas. Besides OP and OL models, other modeling approaches were also utilized for fitting the injury severity data. The MNL approach was also applied to develop the injury severity model by Shankar and Mannering *(20)*, Carson and Mannering *(21)*, Ulfarsson and Mannering *(22)*. In addition, Yamamoto and Shankar *(23)* developed Bivariate Ordered-response Probit models (BOP) of drivers' and most severely injured passenger's severity in collision with fixed objects. Xie et al. *(24)* compared Bayesian Ordered Probit Models and the traditional OP Models. Das et al. *(25)* adopted the simultaneous estimation approach to examine the effect of gradual change in intersection influence distance on crash characteristics that explain injury severity of arterial crashes. As the extension of OP/OL models, Partial Proportional

Odds models (PPO) that allow slope coefficients to differ across levels of the ordered response were used by Wang et al *(26)* and Wang et al. *(16)*. The two studies found that PPO models consistently perform better than the ordered probit models. The Heteroscedastic Logit model (HOL) was used by Wang and Kockelman *(27)* to study the effects of various vehicle, environmental, roadway and occupant characteristics on the severity of injuries sustained by vehicle occupants, conditional on the crash occurrence.

Summarily, most of previous studies focused on the prediction of crash frequency and/or crash rate at exit ramp sections; and a study *(16)* focused on modeling the crash injury severity at freeway diverge areas. No studies on crash injury severity at exit ramps were found. Thus, the primary objective of this study is to develop predictive models to identify the significant factors that contribute to the injury severity at exit ramp sections. Three statistical approaches, Ordered Probit models, Partial Proportional Odds models, and Heterogeneous Probit models (HOP), were used and compared to develop the predictive model based on the historic crash data collected in the Florida State.

## DATA PREPARATION

In this study, totally, 221 exit ramp sections with different ramp configurations were selected by reviewing aero photos of interstate highway systems, expressways, turnpikes, and parkways in Florida. Only right exit ramps were considered in this study sites, and suggested or post speed limits are no less than 25 mph no matter the ramp configurations or ramp length. Crash data were extracted from the crash database maintained by the Florida Department of Transportation (FDOT) based on a combination of roadway ID, beginning milepost, and ending milepost, which were unique for each selected exit ramp section. Except for crash data, roadway inventory data and traffic data were also collected from the FDOT databases or in field. As the result, 24 variables and 3,964 crash observations were collected for a 4-year period (2005-2008). As shown in Table 1, the selected variables could be grouped into seven categories: response, ramp configurations, roadway design, driver, crash, environmental condition, and traffic.

Crash injury severity was divided into 4 levels at an ascending order from no injury to incapacitating/fatal injury. This ordinal variable was used as the response variable in the predictive models. Diverse ramp configuration designs have significantly dissimilar impacts on the safety performance especially for off ramps *(2)*. Four widely used ramp configurations in Florida (as shown in Figure 1) were identified in the study as diamond exit ramps, out connection exit ramps, free-flow loop exit ramps, and parclo loop exit ramps. Diamond exit ramp is a one-way road with both left and right turnings at terminals. Out connection exit ramp only supplies the single turn at the ends of exit ramps. Free-flow loop ramps are designed as full cloverleaf ramps with or without collector or distributor roads on the ramp segments. And parclo loop exit ramp is a partial cloverleaf ramp which has a preference to provide an arrangement setting the right exiting vehicles.

The variables of roadway design on exit ramps include ramp length, ramp lane number, surface width, connection type, shoulder type and width, traffic control type

at ramp terminals, channelization at ramp terminals, posted speed limits on ramps and mainlines, and locations of upstream/downstream intersections on the secondary crossroads. On some exit ramps, the lane number would change from the initial value (1 or 2) to more lanes (3 or more). In this study, a variable WIDEN was used to represent this situation. Crash related variables consisted of crash location and crash type. The terminal area was defined as a 100 feet section in the upstream of the ramp terminals. Environmental variables include the condition of weather, light, and roadway surface. In addition, traffic variables have AADT on ramp, AADT on mainline and peak hour period which was defined as 7:00 -9:00 AM and 4:00 – 6:00 PM.

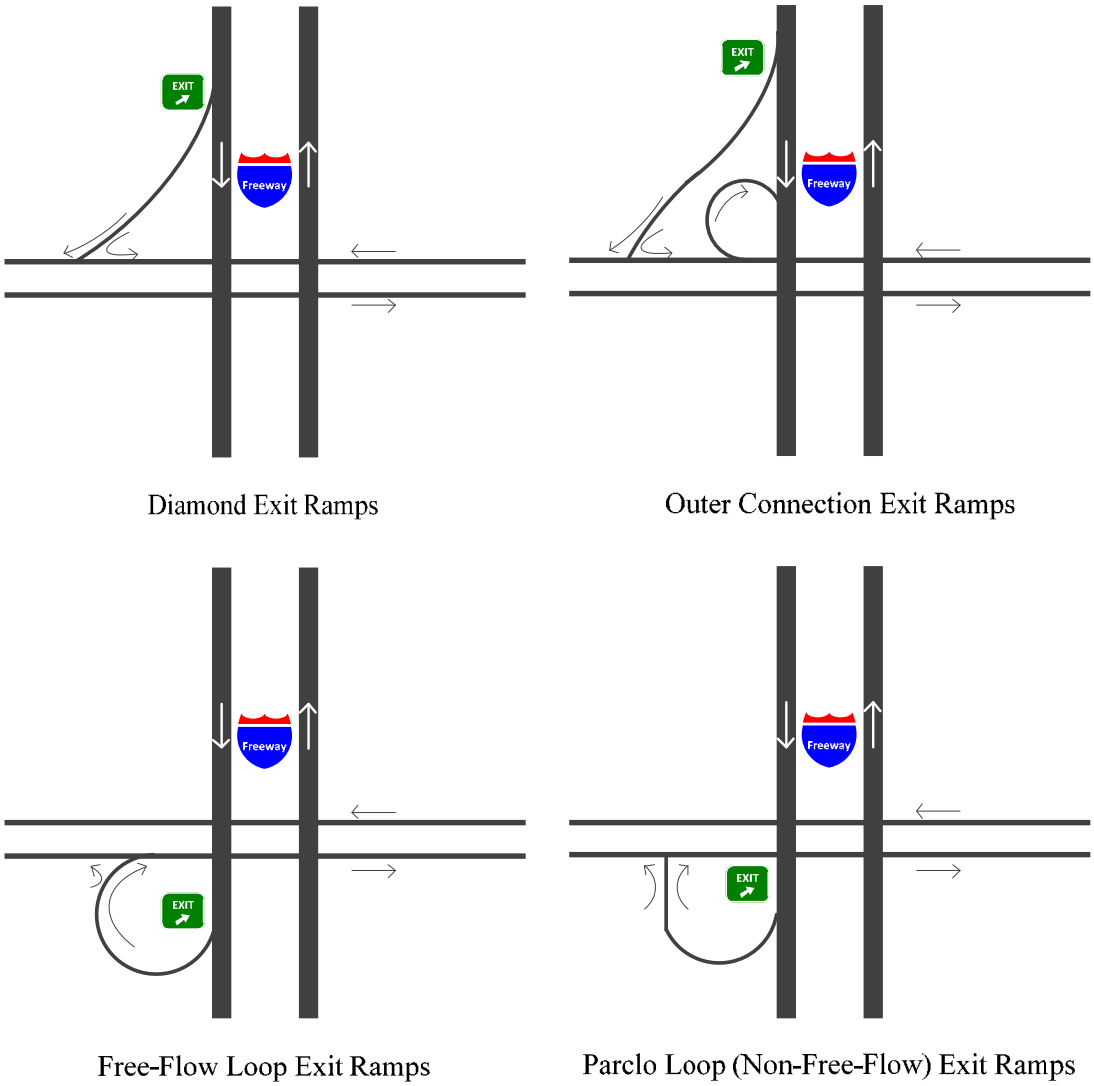



**Figure 1 Ramp Configurations**

**TABLE 1 Description of Selected Variables for Model Development**

| Category | Variable | Value (Description) | Type | Freq. | Percent |
|---|---|---|---|---|---|
| Response | ACCISEV | 1 (No Injury) | Ordinal | 2129 | 53.7 |
| | | 2 (Possible/Invisible Injury) | | 1098 | 27.7 |
| | | 3 (No-Capacitating Injury) | | 528 | 13.3 |
| | | 4 (Incapacitating/Fatal Injury) | | 209 | 5.3 |
| Ramp Configurations | DIAMOND | 1 (Diamond) | Dummy | 3051 | 77.0 |
| | LOOP | 1 (Free-flow Loop) | | 708 | 17.9 |
| | OUTER | 1 (Outer Connection) | | 27 | 0.7 |
| | PARCLO | 1 (Parclo Loop) | | 178 | 4.5 |
| Roadway Design | RAMPLEN | Ramp Length (mile) | Continuous | - | - |
| | LANENUM | 1 (2 Lanes on Ramps)<br>0 (1 Lane on Ramps) | Binary | 2500 | 63.1 |
| | CONNTYPE | 1 (Freeway to Freeway)<br>0 (Freeway to highway) | Binary | 3628 | 91.5 |
| | SURWIDH | Ramp Lane Surface Width (ft) | Continuous | - | - |
| | SHILDTYPE | 1 (Paved Shoulder)<br>0 (Not Paved Shoulder) | Binary | 2008 | 50.7 |
| | SHLDWIDH | Shoulder Width (ft) | Continuous | - | - |
| | SIGNAL | 1 (Signalized Control on Terminal)<br>0 (Unsignalized Control ) | Binary | 3142 | 79.3 |
| | WIDEN | 1 (Lane Widening on exit ramps)<br>0 (No LaneWidening) | Binary | 2884 | 72.8 |
| | CHANNELIZE | 1 (Channelization at the terminal of exit ramps connecting to cross-roads)<br>0 (No Channelization) | Binary | 1777 | 44.8 |
| | DISUPINTR | Distance between ramp terminal and the first upstream intersection (mile) | Continuous | - | - |
| | DISDWNINTR | Distance between ramp terminal and the first downstream intersection (mile) | Continuous | - | - |
| | SPEEDMAIN | Posted Speed Limit on Ramps (mph) | Continuous | - | - |
| | SPEEDRAMP | Posted Speed Limit on Ramps (mph) | Continuous | - | - |
| Driver | ALCDRUG | 1 (Alcohol/Drug involved)<br>0 (No) | Binary | 227 | 5.7 |
| Crash | LOCCRASH | 1 (Terminal Area)<br>0 (Not Terminal Area) | Binary | 859 | 21.7 |
| | REAREND | 1 (Rear-end Crash) | Dummy | 2180 | 55.0 |
| | ANGLE | 1 (Angle Crash) | | 352 | 8.9 |
| | CONCRBAR | 1 (Hit on Concrete Barrel) | | 111 | 2.8 |
| | SIDESWIP | 1 (Sideswip Crash) | | 282 | 7.1 |
| | OVERTURN | 1 (Overturn) | | 119 | 3.0 |
| | DITCH | 1 (Run into Ditch/Culvert) | | 83 | 2.1 |
| | LEFTTURN | 1 (Left turn Crash) | | 73 | 1.8 |
| | GUARDRAIL | 1 (Hit on Guardrail) | | 135 | 3.4 |
| | OTHER | 1 (Other Crash Type) | | 1015 | 25.6 |

**TABLE 1 Description of Selected Variables for Model Development (Cont'd)**

| Category | Variable | Value (Description) | Type | Freq. | Percent |
|---|---|---|---|---|---|
| Environmental Conditions | LIGHTCON | 1 (Good Light Condition)<br>0 (Not Good) | Binary | 2778 | 70.1 |
| | WEATCON | 1 (Good Weather Condition)<br>0 (Not Good) | Binary | 3466 | 87.4 |
| | SURCON | 1 (Dry Surface)<br>0 (Not Dry) | Binary | 3132 | 79 |
| TRAFFIC | AADTRAMP | AADT on Ramp | Continuous | - | - |
| | AADTMAIN | AADT on Mainline | Continuous | - | - |
| | PEAKHOUR | 1 (Peak Hour)<br>0 (Non-Peak Hour) | Binary | 803 | 20.3 |

## MOTHEDOLOGY

### Model Structure

As mentioned above, ordered choice models (OP and OL) are widely used for modeling injury severity in transportation safety analysis. Let $Y_i$ denote the injury severity for the *ith* observed crash, the ordered choice model can be written in terms of the cumulative probability as

$$\Pr(Y_i > m) = 1 - \sum_{j=1}^{m} \Pr(Y_i = j) = 1 - \mathrm{F}(\tau_j - \mathbf{x}_i\boldsymbol{\beta}) \quad j = 1,2,3,4 \tag{1}$$

where $\Pr(Y_i)$ is the probability of the response variable $Y_i$ adopting a specific severity level $j$; $\tau_j$ is the threshold parameter (cut-off points) to be estimated and satisfies the restriction $\tau_1 < \tau_2 < \cdots < \tau_{J-1}$; $\mathbf{x}_i$ is a vector containing the values of observed crash $i$ on the full set of explanatory variables; $\boldsymbol{\beta}$ is a vector of coefficients associated with the explanatory variables; F() represents the CDF (cumulative distribution function) of the standard normal distribution (OP) or the logistic distribution (OL).

In Equations 1, the slope coefficients $\boldsymbol{\beta}$ are the same across the injury severity levels while the cut-off points differ; in other words, $\boldsymbol{\beta}$ are independent to the injury severity level ( $j$ ) for ordered choice models. The identification of slope coefficients is an important assumption of in the ordered choice models, called the Parallel Regression Assumption *(28)*. However, a key problem of the OP/OL is that its assumptions are often violated; it is common for one or more coefficients to differ over the outcome levels *(28, 29, 30)*.

For relaxing the restriction that all slope coefficients are identical across injury severity levels, Partial Proportional Odds models (PPO) was proposed by Peterson and Harrel *(29)*. In the PPO, some of the slope coefficients can be identical for all outcome values if the parallel regression assumption is not violated; other coefficients can be different if the assumption is violated. The cumulative probability equation is given as following

$$\Pr(Y_i > j \mid \mathbf{x_i}) = \frac{\exp(\alpha_j + \mathbf{x_i^a}\boldsymbol{\beta}^a + \mathbf{x_i^n}\boldsymbol{\beta}_j^n)}{1+\exp(\alpha_j + \mathbf{x_i^a}\boldsymbol{\beta}^a + \mathbf{x_i^n}\boldsymbol{\beta}_j^n)}, \qquad j = 1,2,...,J-1 \tag{2}$$

where $\alpha_j$ is the *j*th constant coefficient (equal to the negatives of the cut-off points); $\mathbf{x_i^a}$ is the vector containing the values of observation $i$ on that subset of explanatory variables for which the parallel assumptions are not violated; $\boldsymbol{\beta}^a$ is the vector of coefficients associated with the non-violated variables, and is the same across values of $Y$ ; $\mathbf{x_i^n}$ is the vector containing the values of observation $i$ on that subset of explanatory variables for which the parallel assumption are violated; $\boldsymbol{\beta}_j^n$ is the vector of coefficients associated with the violated variables, and differs across the response values.

PPO models can relax the overly restriction owned by ordered choice models while are more parsimonious than non-ordinal alternatives, such as multinomial logit models. However, there are two concerns when PPO is used for fitting injury severity data. First, unlike other categorical models, estimated probabilities of PPO may be negative since its non-parallel properties. This issue was addressed by McCullaph and Nelder *(31)*. Second, PPO models may have multiple coefficients for the variables which are non-paralleled. Sometimes, it is difficult to explain the models according to the coefficients if the variety of coefficients is opposite across the output values.

Another extension to ordered choice models is the Heterogeneous Choice Model (also known as the Location-Scale model) which can be used to fit binary and ordinal response *(31,32)*. To estimate traditional ordered choice models, an assumption that the error term of the latent variable $y^*$ is homoskedastic or constant across individuals must to be made, while the variance of error terms is often differ across individuals actually *(31,32)*. The violation of the assumption results in a biased, inconsistent, and inefficient parameter estimation in ordered choice models. Heterogeneous Choice models adopt a multiplicative functional form for the variance of error terms ( $\varepsilon_i$ ):

$$\begin{aligned} &\mathrm{Var}(\varepsilon_i) = \sigma_i^2 = \exp(\mathbf{z}_i\gamma)^2 \\ &\sigma_i = \exp(\mathbf{z}_i\gamma) \end{aligned} \tag{3}$$

where $\mathbf{z}_i$ is a vector of covariates of the *ith* observation that define groups with different error variances in the underlying latent variable, $\gamma$ is a vector of parameters to be estimated. The cumulative probability form of Heterogeneous Choice models is given as

$$\Pr(Y_i > j) = \mathrm{F}\left(\frac{\tau_j - \mathbf{x}_i\boldsymbol{\beta}}{\sigma_i}\right) = \mathrm{F}\left(\frac{\tau_j - \mathbf{x}_i\boldsymbol{\beta}}{\exp(\mathbf{z}_i\gamma)}\right) \tag{4}$$

where F() is the cumulative distribution function of the standard normal distribution (Heterogeneous Ordered Probit models) or the logistic distribution (Heterogeneous Ordered Logit models).

Summarily, Order Probit models are widely used in crash injury severity analysis but with two assumptions that often are violated. Heterogeneous Choice models can correct the invalid assumption of the variance of error terms; meanwhile Partial Proportional Odds models can be called for the violation of the Parallel Assumption. In this study, Ordered Probit model (OP), Heterogeneous Probit model (HOP), and Partial Proportional Odds models (PPO) were used for fitting the injury severity data on exit ramp sections.

**Model Interpretation and Goodness-of-Fit**

The slope coefficients of fitted models can be used to interpret the impacts of variables on the injury severity. A positive coefficient indicate the value changing from 0 to 1 (for discrete variable) or a unit increase in the variable value (for continuous variable) is more likely increase the injury severity; while a negative coefficient means the value change tends to decrease the injury severity.

Another useful approach to interpret the coefficients is marginal effect, more commonly, defined as the slope of the probability curve relating $x_k$ to $\Pr(Y = j|X)$ at the mean values of all variables, holding all other explanatory variables constant.

$$\begin{aligned}\frac{\partial \Pr(y=j)}{\partial x_k} &= \frac{\partial F(\tau_j - \bar{\mathbf{x}}\boldsymbol{\beta}_j)}{\partial x_k} - \frac{\partial F(\tau_{j-1} - \bar{\mathbf{x}}\boldsymbol{\beta}_{j-1})}{\partial x_k} \\ &= \beta_{k,j}\left[f(\tau_{j-1} - \bar{\mathbf{x}}\boldsymbol{\beta}_{j-1}) - f\left(\tau_j - \bar{\mathbf{x}}\boldsymbol{\beta}_j\right)\right]\end{aligned} \tag{5}$$

where $F()$ and $f()$ are the cumulative distribution function and the probability density function respectively; $\bar{\mathbf{x}}$ is the mean values of all variables; and $\beta_{k,j}$ is the slop coefficient of the *kth* variable for the *jth* outcome level. In this study, for continues variables, the marginal effects are calculated as the derivative of the curve; and for dummy/binary independent variables, the marginal effects are computed as the difference of probabilities due to the discrete change in the variables from 0 to 1. Equation 5 is suitable for OP model and PPO model. For HOP model, the marginal effect form is given as

$$\frac{\partial \Pr(y=j)}{\partial x_k} = -\left(f\left(\frac{\tau_j - \bar{\mathbf{x}}\boldsymbol{\beta}}{\bar{\sigma}}\right) - f\left(\frac{\tau_{j-1} - \bar{\mathbf{x}}\boldsymbol{\beta}}{\bar{\sigma}}\right)\right)\cdot\frac{\beta_k}{\bar{\sigma}} \tag{6}$$

where $\beta_k$ is the slop coefficient of the *kth* variable; $\bar{\sigma}$ is the mean variance for across observations.

In this study, Pseudo - $R^2$ was used to evaluate the goodness-of-fit of the fitted models. If the unconstrained model (with all slope coefficients) does much better than the constrained model (constant coefficients only) the value Pseudo - $R^2$ is closer to 1.

## ESTIMATION RESULTS

### Comparison of Fitted Models

In this study, the STATA software was used for estimating the predictive models. At first, the stepwise method was applied to select the explanatory variables which have significant contributes to the outcome probabilities from the initial variable list (Table 1) with a confidence level (95%). Successively, the Ordered Probit model was fitted by the *oprobit* command and the results are given in Table 2. The *gologit2* command was utilized to estimate the Partial Proportional Odds model with the explanatory variables selected by the step wise, meanwhile the variable that violated the Parallel Assumption were identified by the Brant Test *(30)*. The estimation results are given in Table 3. The Heterogeneous Probit model was also fitted by the *oglm* command with the explanatory variables selected by the stepwise method. The variables identified by the Brant Test were believed have different error term variance across individuals. The fitted model and estimated variances for these variables are given in Table 4. Finally, marginal effects were calculated for each fitted model and results were given in Table 5.

From the results of three fitted models, the Pseudo-$R^2$ values of PPO model and HOP model (0.0482 and 0.0451 respectively) are closer to 1 than that of OP model (0.0350). This result showed that PPO model and HOP model are better than OP model in fitting injury severity data. OUTER, ALCDRUG, REAREND, and CHANNEL were identified by the Brant Test that the Parallel Assumption were violated. Unlike in OP model, their slope coefficients in PPO model are different across injury severity levels while coefficients of other explanatory variables are the same. The four variables were taken as the variables with heterogeneous variance of error terms. It should be noticed that *gologit2* command found a sample case has negative prediction outcome of the PPO model (Table 3).

### Ramp Configurations

Since all slope coefficients of the ramp configuration variables are positive in the three fitted models, it could be suggested that Free-Flow Loop, Outer Connection, and Parclo Loop are more likely to increase injury severity than Diamond. In other words, diamond configuration has the highest safety performance in terms of crash injury severity. In Table 5, all marginal effects of ramp configuration variables in the three models for $j = 1$ (no injury) are negative; that means the three configurations tends to decrease the probability of no injury when a crash happened on exit ramp sections. Except for Outer Connection, both other two configurations have positive marginal effects for $j = 2,3,4$ in three models. It can be concluded that the two configurations tends to increase the probability of possible/invisible injury, no-capacitating injury, and incapacitating/fatal injury comparing to Diamond. For Outer Connection, there is a little difference between the OP model and other two models for $j = 2$. But it also can be concluded that Outer Connection is more likely to increase injury severity on exit ramp sections since its marginal effects are all positive for $j = 3,4$. In addition, for $j = 4$, Outer Connection in all three models has the

highest values of marginal effects. Thus, Outer Connection seems to increase the probability of a severe injury mostly.

**Roadway Design**

In the fitted models, roadway design factors included number of ramp lanes, shoulder width, speed limits on ramp and mainline, channelization at ramp terminal, and distance between ramp terminal and upstream intersection on crossroads. Other factors in Table 1, such as ramp length, shoulder type, connection type, surface width, traffic control type at terminal, lane widening on ramp, and distance between ramp terminal and downstream intersection on crossroads, are excluded from the models due to insignificant impacts on injury severity on exit ramp sections.

There are two common lane configurations on exit ramp sections in Florida: one lane or two lanes. Based on the slop coefficients and marginal effects in Tables 2 to 5, it can be concluded that two-lane configuration on exit ramp sections is more likely to increase the injury severity than one-lane configuration. One possible reason is that two-lane configuration increases the risk of angle crashes which is more dangerous to vehicle occupants than other crash type on exit ramps. The parallel assumption was rejected by the Brant Test for the channelization variable. Thus, the marginal effects in OP model are reversed to those in PPO model and HOL model for incapacitating/fatal injury. In OP model, the impact of channelization is to increase injury severity, while a more reasonable result, which the impact is expected to decrease the probability of incapacitating/fatal injury, is provided in the PPO and HOL models. Distance between ramp terminal and the first upstream intersection on crossroads is another significant factor that influences the injury severity on exit ramp sections. A large distance between the first upstream intersection to ramp terminals, which induces more conflicts between through traffic on crossroads and the exit traffic on ramps, is more likely to increase severity on exit ramp sections. It can be understood that high speed limit on ramp is easy to result in more severe crash on exit ramps, and this conclusion can be supported by the three fitted models. However, the models also suggested that high speed limit on mainlines is more likely to decrease the probability of severe injury. It is reversed to the common sense that high speed difference tends to result in severe crashes. Another interesting factor is shoulder width on exit ramps. The conclusion that an increase in shoulder width is more likely to result in severe injury on exit ramp sections is suggested by the fitted models. A possible explanation is that a wider shoulder width is more likely to induce a high operating speed which is a major factor that tends to increase probability of severe injury on ramps.

**Crash Type**

Crash type is an important category of factors that influence injury severity on roadways. For $j = 2,3,4$ , the marginal effects of rear-end crash and sideswip crash in three models are positive. These two crash types are more likely decrease the injury severity on exit ramp sections. On the hand, the negative marginal effects of hit-on concrete barrel and overturn for $j = 2,3,4$ showed these two crash types are more likely to increase injury severity on exit ramp sections.

**Environmental Condition, Traffic, and Driver**

Environmental conditions, including the conditions of weather, light, and roadway surface, are excluded from the fitted models because their contributions to injury severity are not significant in statistics. AADT on mainline, AADT on ramp, and peak hour are also abandoned due to their insignificance. For alcohol and/or drug involvement, the marginal effects are positive for $j = 3,4$ in the three fitted models. It can be concluded that alcohol and/or drug tends to increase the probability of severe injury (No-Capacitating Injury and Incapacitating/Fatal Injury).

## CONCLUSIONS

This study evaluated the impacts of various factors (including ramp configuration, roadway design, driver, environment, and traffic) on the injury severity on exit ramp sections. Three predictive models were developed based on the crash data collected at 221 exit ramps in Florida using Ordered Probit model, Partial Proportional Odds model, and Heterogeneous Ordered Probit model respectively. Based on the prediction models, the following conclusions can be obtained:

- Ramp configurations (Diamond, Free-flow Loop, Outer Connection, and Parclo Loop) have significant impacts on the injury severity on exit ramp sections. Diamond exit ramp is most likely to decrease the injury severity on exit ramp sections than other ramp configurations.
- Two-lane configuration on exit ramp tends to increase the injury severity; while channelization at ramp terminals is more likely to decrease the injury severity. And a longer distance between ramp terminal and the first intersection on crossroads is another factor that increases the probability of severe injury on exit ramp sections.
- Speed limit on ramps and mainlines are both significantly influence the injury severity on exit ramp sections, but their impacts are reversed. A higher speed limit on ramps is more likely to increase the injury severity.
- Alcohol and drug is the most significant factor that increases the probability of server injury on exit ramp sections.
- Hit-on concrete barrel and overturn are the two crash types that are more likely to increase injury severity, while rear-end and sideswipe are the crash types that are more likely to decrease injury severity on exit ramp sections.
- Both PPO model and HOP model can provide better goodness-of-fit than OP model and more reasonable interpretations of explanatory variables. However, PPO model may produce negative prediction outcome, and complex interpretations of coefficients. Thus HOP model is an attractive alternative to PPO model and OP model for modeling injury severity.

**TABLE 2 Ordered Probit Model Estimation**

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | Number of obs=10946 | | |
| AIC= 8549.872 | | | | LR chi2(11) = 693.3 | | |
| BIC= 8656.717 | | | | Prob > chi2 = 0.0000 | | |
| Pseudo $R^2$= 0.0350 | | | | Log likelihood = -4257.9358 | | |
| | | | | | | |
| Explanatory Variable | Coef. | Std. Err. | $z$ | $p > \|z\|$ | [95% Conf. Interval] | |
| LOOP | 0.11297 | 0.05744 | 1.97000 | 0.04900 | 0.00039 | 0.22555 |
| OUTER | 0.50370 | 0.22656 | 2.22000 | 0.02600 | 0.05965 | 0.94775 |
| PARCLO | 0.26964 | 0.09418 | 2.86000 | 0.00400 | 0.08506 | 0.45422 |
| LANENUM | 0.12683 | 0.03989 | 3.18000 | 0.00100 | 0.04865 | 0.20501 |
| SHLDWIDH | 0.00909 | 0.00454 | 2.00000 | 0.04500 | 0.00018 | 0.01799 |
| DISUPINTR | 0.17883 | 0.04831 | 3.70000 | 0.00000 | 0.08415 | 0.27351 |
| SPEEDRAMP | 0.02536 | 0.00737 | 3.44000 | 0.00100 | 0.01092 | 0.03980 |
| SPEEDMAIN | -0.02154 | 0.00530 | -4.06000 | 0.00000 | -0.03193 | -0.01115 |
| ALCDRUG | 0.19551 | 0.07829 | 2.50000 | 0.01300 | 0.04206 | 0.34896 |
| CONCRBAR | 0.24362 | 0.11076 | 2.20000 | 0.02800 | 0.02653 | 0.46071 |
| REAREND | -0.20343 | 0.04235 | -4.80000 | 0.00000 | -0.28644 | -0.12043 |
| OVERTURN | 0.61152 | 0.10471 | 5.84000 | 0.00000 | 0.40628 | 0.81675 |
| SIDESWIP | -0.97208 | 0.09014 | -10.78000 | 0.00000 | -1.14876 | -0.79541 |
| CHANNELIZE | 0.11153 | 0.03942 | 2.83000 | 0.00500 | 0.03427 | 0.18879 |
| /cut1 | -0.19316 | 0.18434 | | -0.55445 | 0.16814 | -0.19316 |
| /cut2 | 0.64916 | 0.18452 | | 0.28752 | 1.01081 | 0.64916 |
| /cut3 | 1.41934 | 0.18562 | | 1.05553 | 1.78315 | 1.41934 |

**TABLE 3 Partial Proportional Odds Model Estimation***

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | Number of obs=3964 | | |
| AIC= 8449.716 | | | | LR chi2(44)=425.05 | | |
| BIC= 8606.841 | | | | Prob > chi2=0 | | |
| Pseudo R2=0.0482 | | | | Log likelihood = -4199.858 | | |
| | | | | | | |
| Variable | Coef. | Std. Err. | $z$ | $p > \|z\|$ | [95% Conf. Interval] | |
| $j = 1$ (No Injury) | | | | | | |
| LOOP | 0.21119 | 0.09688 | 2.18000 | 0.02900 | 0.02131 | 0.40108 |
| OUTER | 0.41372 | 0.40769 | 1.01000 | 0.31000 | -0.38533 | 1.21277 |
| PARCLO | 0.42396 | 0.16002 | 2.65000 | 0.00800 | 0.11032 | 0.73759 |
| LANENUM | 0.22837 | 0.06728 | 3.39000 | 0.00100 | 0.09650 | 0.36024 |
| SHLDWIDH | 0.01531 | 0.00767 | 2.00000 | 0.04600 | 0.00028 | 0.03033 |
| DISUPINTR | 0.32807 | 0.08154 | 4.02000 | 0.00000 | 0.16825 | 0.48789 |
| SPEEDRAMP | 0.04040 | 0.01254 | 3.22000 | 0.00100 | 0.01583 | 0.06498 |
| SPEEDMAIN | -0.03444 | 0.00899 | -3.83000 | 0.00000 | -0.05207 | -0.01681 |
| ALCDRUG | 0.09455 | 0.14251 | 0.66000 | 0.50700 | -0.18476 | 0.37387 |
| CONCRBAR | 0.36093 | 0.18234 | 1.98000 | 0.04800 | 0.00355 | 0.71831 |
| REAREND | -0.10846 | 0.07453 | -1.46000 | 0.14600 | -0.25454 | 0.03763 |
| OVERTURN | 0.94208 | 0.17545 | 5.37000 | 0.00000 | 0.59821 | 1.28596 |
| SIDESWIP | -1.59287 | 0.16408 | -9.71000 | 0.00000 | -1.91446 | -1.27127 |
| CHANNELIZE | 0.26763 | 0.07006 | 3.82000 | 0.00000 | 0.13031 | 0.40495 |
| _cons | 0.12286 | 0.31019 | 0.40000 | 0.69200 | -0.48509 | 0.73081 |

**TABLE 3. Partial Proportional Odds Model Estimation (Cont'd)**

| | | | | | | |
|---|---|---|---|---|---|---|
| $j = 2$ (Possible Injury) | | | | | | |
| LOOP | 0.21119 | 0.09688 | 2.18000 | 0.02900 | 0.02131 | 0.40108 |
| OUTER | 1.05346 | 0.44174 | 2.38000 | 0.01700 | 0.18766 | 1.91926 |
| PARCLO | 0.42396 | 0.16002 | 2.65000 | 0.00800 | 0.11032 | 0.73759 |
| LANENUM | 0.22837 | 0.06728 | 3.39000 | 0.00100 | 0.09650 | 0.36024 |
| SHLDWIDH | 0.01531 | 0.00767 | 2.00000 | 0.04600 | 0.00028 | 0.03033 |
| DISUPINTR | 0.32807 | 0.08154 | 4.02000 | 0.00000 | 0.16825 | 0.48789 |
| SPEEDRAMP | 0.04040 | 0.01254 | 3.22000 | 0.00100 | 0.01583 | 0.06498 |
| SPEEDMAIN | -0.03444 | 0.00899 | -3.83000 | 0.00000 | -0.05207 | -0.01681 |
| ALCDRUG | 0.57039 | 0.15180 | 3.76000 | 0.00000 | 0.27287 | 0.86791 |
| CONCRBAR | 0.36093 | 0.18234 | 1.98000 | 0.04800 | 0.00355 | 0.71831 |
| REAREND | -0.72283 | 0.09303 | -7.77000 | 0.00000 | -0.90517 | -0.54050 |
| OVERTURN | 0.94208 | 0.17545 | 5.37000 | 0.00000 | 0.59821 | 1.28596 |
| SIDESWIP | -1.59287 | 0.16408 | -9.71000 | 0.00000 | -1.91446 | -1.27127 |
| CHANNELIZE | 0.15560 | 0.08802 | 1.77000 | 0.07700 | -0.01692 | 0.32813 |
| _cons | -0.95745 | 0.31172 | -3.07000 | 0.00200 | -1.56841 | -0.34649 |
| $j = 3$ (No-Capacitating Injury) | | | | | | |
| LOOP | 0.21119 | 0.09688 | 2.18000 | 0.02900 | 0.02131 | 0.40108 |
| OUTER | 1.72169 | 0.51537 | 3.34000 | 0.00100 | 0.71159 | 2.73180 |
| PARCLO | 0.42396 | 0.16002 | 2.65000 | 0.00800 | 0.11032 | 0.73759 |
| LANENUM | 0.22837 | 0.06728 | 3.39000 | 0.00100 | 0.09650 | 0.36024 |
| SHLDWIDH | 0.01531 | 0.00767 | 2.00000 | 0.04600 | 0.00028 | 0.03033 |
| DISUPINTR | 0.32807 | 0.08154 | 4.02000 | 0.00000 | 0.16825 | 0.48789 |
| SPEEDRAMP | 0.04040 | 0.01254 | 3.22000 | 0.00100 | 0.01583 | 0.06498 |
| SPEEDMAIN | -0.03444 | 0.00899 | -3.83000 | 0.00000 | -0.05207 | -0.01681 |
| ALCDRUG | 0.58971 | 0.22655 | 2.60000 | 0.00900 | 0.14567 | 1.03374 |
| CONCRBAR | 0.36093 | 0.18234 | 1.98000 | 0.04800 | 0.00355 | 0.71831 |
| REAREND | -1.07592 | 0.16765 | -6.42000 | 0.00000 | -1.40451 | -0.74732 |
| OVERTURN | 0.94208 | 0.17545 | 5.37000 | 0.00000 | 0.59821 | 1.28596 |
| SIDESWIP | -1.59287 | 0.16408 | -9.71000 | 0.00000 | -1.91446 | -1.27127 |
| CHANNELIZE | -0.16344 | 0.15338 | -1.07000 | 0.28700 | -0.46405 | 0.13717 |
| _cons | -2.20760 | 0.32065 | -6.88000 | 0.00000 | -2.83605 | -1.57915 |

* WARNING! 1 in-sample cases have an outcome with a predicted probability that is less than 0.

**TABLE 4. Heterogeneous Ordered Probit Estimation**

| | | | | Number of obs=3964 | | |
|---|---|---|---|---|---|---|
| AIC= 8469.094 | | | | LR chi2(11) = 397.67 | | |
| BIC= 8601.08 | | | | Prob > chi2 = 0.0000 | | |
| Pseudo $R^2$= 0.0451 | | | | Log likelihood = -4213.5472 | | |
| Explanatory Variable | Coef. | Std. Err. | $z$ | $p > \|z\|$ | [95% Conf. Interval] | |
| Choice | | | | | | |
| LOOP | 0.07757 | 0.04684 | 1.66000 | 0.09800 | -0.01422 | 0.16937 |
| OUTER | 0.09677 | 0.39823 | 0.24000 | 0.80800 | -0.68376 | 0.87729 |
| PARCLO | 0.19557 | 0.08665 | 2.26000 | 0.02400 | 0.02574 | 0.36540 |
| LANENUM | 0.09296 | 0.03111 | 2.99000 | 0.00300 | 0.03199 | 0.15393 |
| SHLDWIDH | 0.00665 | 0.00352 | 1.89000 | 0.05900 | -0.00025 | 0.01355 |
| DISUPINTR | 0.13966 | 0.03727 | 3.75000 | 0.00000 | 0.06661 | 0.21270 |
| SPEEDRAMP | 0.02010 | 0.00629 | 3.20000 | 0.00100 | 0.00778 | 0.03242 |
| SPEEDMAIN | -0.01699 | 0.00443 | -3.84000 | 0.00000 | -0.02566 | -0.00831 |
| ALCDRUG | 0.06026 | 0.08780 | 0.69000 | 0.49200 | -0.11181 | 0.23234 |
| CONCRBAR | 0.20716 | 0.10887 | 1.90000 | 0.05700 | -0.00623 | 0.42055 |
| REAREND | -0.06457 | 0.04012 | -1.61000 | 0.10800 | -0.14321 | 0.01406 |
| OVERTURN | 0.56425 | 0.10275 | 5.49000 | 0.00000 | 0.36287 | 0.76563 |
| SIDESWIP | -0.86601 | 0.08684 | -9.97000 | 0.00000 | -1.03622 | -0.69581 |
| CHANNELIZE | 0.13884 | 0.03430 | 4.05000 | 0.00000 | 0.07161 | 0.20607 |
| Variance | | | | | | |
| OUTER | 0.52455 | 0.33100 | 1.58000 | 0.11300 | -0.12420 | 1.17329 |
| ALCDRUG | 0.21880 | 0.09138 | 2.39000 | 0.01700 | 0.03970 | 0.39789 |
| REAREND | -0.29853 | 0.04348 | -6.87000 | 0.00000 | -0.38374 | -0.21331 |
| CHANNELIZE | -0.15535 | 0.04170 | -3.73000 | 0.00000 | -0.23708 | -0.07363 |
| /cut1 | -0.08851 | 0.14227 | -0.62000 | 0.53400 | -0.36735 | 0.19033 |
| /cut2 | 0.58155 | 0.14220 | 4.09000 | 0.00000 | 0.30285 | 0.86025 |
| /cut3 | 1.23539 | 0.14556 | 8.49000 | 0.00000 | 0.95009 | 1.52069 |

**TABLE 5. Marginal Effects for Fitted Models**

| Explanatory Variable | Ordered Probit Model | Partial Proportion Odds Model | Heterogeneous Ordered Probit |
|---|---|---|---|
| $j = 1$ *No Injury* | | | |
| LOOP | -0.04496 | -0.05262 | -0.03836 |
| OUTER | -0.19663 | -0.10302 | -0.04563 |
| PARCLO | -0.10721 | -0.10557 | -0.09676 |
| LANENUM | -0.05040 | -0.05677 | -0.04590 |
| SHLDWIDH | -0.00361 | -0.00380 | -0.00328 |
| DISUPINTR | -0.07101 | -0.08146 | -0.06889 |
| SPEEDRAMP | -0.01007 | -0.01003 | -0.00991 |
| SPEEDMAIN | 0.00855 | 0.00855 | 0.00838 |
| ALCDRUG | -0.07787 | -0.02354 | -0.03286 |
| CONCRBAR | -0.09693 | -0.08999 | -0.10245 |
| REAREND | 0.08071 | 0.02693 | 0.04347 |
| OVERTURN | -0.23604 | -0.22711 | -0.26763 |
| SIDESWIP | 0.33019 | 0.32516 | 0.35420 |
| CHANNELIZE | -0.04429 | -0.06643 | -0.06333 |
| $j = 2$ *Possible/Invisible Injury* | | | |
| LOOP | 0.01499 | 0.02181 | 0.01285 |
| OUTER | 0.03924 | -0.09477 | -0.09473 |
| PARCLO | 0.03023 | 0.03866 | 0.02790 |
| LANENUM | 0.01735 | 0.02426 | 0.01577 |
| SHLDWIDH | 0.00128 | 0.00166 | 0.00115 |
| DISUPINTR | 0.02512 | 0.03568 | 0.02425 |
| SPEEDRAMP | 0.00356 | 0.00439 | 0.00349 |
| SPEEDMAIN | -0.00303 | -0.00375 | -0.00295 |
| ALCDRUG | 0.02368 | -0.06973 | -0.03732 |
| CONCRBAR | 0.02781 | 0.03380 | 0.02886 |
| REAREND | -0.02803 | 0.07667 | 0.04801 |
| OVERTURN | 0.04056 | 0.05650 | 0.03795 |
| SIDESWIP | -0.17084 | -0.18496 | -0.18589 |
| CHANNELIZE | 0.01551 | 0.04459 | 0.05585 |

**TABLE 5. Marginal Effects for Fitted Models (Cont'd)**

| | Ordered Probit Model | Partial Proportion Odds Model | Heterogeneous Ordered Probit |
|---|---|---|---|
| $j = 3$ *No-Capacitating Injury* | | | |
| LOOP | 0.01885 | 0.02222 | 0.01673 |
| OUTER | 0.08748 | 0.05081 | 0.01341 |
| PARCLO | 0.04631 | 0.04750 | 0.04347 |
| LANENUM | 0.02096 | 0.02353 | 0.01989 |
| SHLDWIDH | 0.00149 | 0.00155 | 0.00141 |
| DISUPINTR | 0.02934 | 0.03325 | 0.02966 |
| SPEEDRAMP | 0.00416 | 0.00410 | 0.00427 |
| SPEEDMAIN | -0.00353 | -0.00349 | -0.00361 |
| ALCDRUG | 0.03324 | 0.06443 | 0.02212 |
| CONCRBAR | 0.04177 | 0.04002 | 0.04619 |
| REAREND | -0.03347 | -0.05886 | -0.03816 |
| OVERTURN | 0.10537 | 0.11570 | 0.12559 |
| SIDESWIP | -0.11339 | -0.10521 | -0.12420 |
| CHANNELIZE | 0.01834 | 0.02803 | 0.01593 |
| $j = 4$ *Incapacitating/Fatal Injury* | | | |
| LOOP | 0.01112 | 0.00860 | 0.00878 |
| OUTER | 0.06991 | 0.14699 | 0.12694 |
| PARCLO | 0.03068 | 0.01941 | 0.02539 |
| LANENUM | 0.01208 | 0.00898 | 0.01024 |
| SHLDWIDH | 0.00084 | 0.00058 | 0.00071 |
| DISUPINTR | 0.01654 | 0.01253 | 0.01497 |
| SPEEDRAMP | 0.00235 | 0.00154 | 0.00215 |
| SPEEDMAIN | -0.00199 | -0.00132 | -0.00182 |
| ALCDRUG | 0.02095 | 0.02885 | 0.04807 |
| CONCRBAR | 0.02736 | 0.01617 | 0.02739 |
| REAREND | -0.01921 | -0.04474 | -0.05332 |
| OVERTURN | 0.09011 | 0.05491 | 0.10408 |
| SIDESWIP | -0.04596 | -0.03498 | -0.04411 |
| CHANNELIZE | 0.01043 | -0.00620 | -0.00844 |

## REFERENCE


1. Lord, D. and Bonneson, J., *Calibration of Predictive Models for Estimating Safety of Ramp Design Configurations*, Transportation Research Record, Transportation Research Board of the National Academies, Washington, D.C., Vol. 1908, 2005, p. 88-95.
2. Bauer, K.M. and Harwood, D.W., *Statistical Models of Accidents on Interchange Ramps and Speed-Change Lanes*, FHWA-RD-97-106, FHWA, U.S. Department of Transportation, 1998.
3. Khorashadi, A., *Effect of Ramp Type and Geometry on Accidents*, FHWA/CA/TE-98/13, California Department of Transportation, 1998
4. McCartt, A.T., Northrup, V.S., and Retting, R.A., *Types and Characteristics of Ramp-Related Motor Vehicle Crashes On Urban Interstate Roadways In Northern Virginia*, Journal of Safety Research, 35, 2004, p 107–114
5. Abdel-Aty, M. and Huang, Y., *Exploratory Spatial Analysis of Expressway Ramps and its Effect on Route Choice*, Journal of Transportation Engineering, Vol. 130, No. 1, 2004, pp. 104-112
6. Hunter, M., Machemehl, R., and Tsyganov, A., *OperationalEvaluation of Freeway Ramp Design,* Transportation Research Record, Transportation Research Board of the National Academies, Vol. 1751, 2001, p. 90-100
7. Hunter, M., Machemehl R., and Tsyganov, A., *Reevaluation of Ramp Design Speed Criteria: Summary Report*, Center for Transportation Research, The University of Texas at Austin, Texas Department of Transportation Research and Technology Transfer Section/Construction Division, No. 1732-S.
8. Bared, J., Giering, G.L., and Warren, D.L., *Safety of Evaluation of Acceleration and Deceleration Lane Lengths*, ITE Journal, May 1999, p 50-54
9. Sarhan, M., Hassan, Y., and Halim, A.E., *Design of Freeway Speed Change Lanes: Safety-Explicit Approach*, Transportation Research Board, Washington, D.C., 2006
10. Garcia, A. and Romero, M.A., *Experimental Observation of Vehicle Evolution on a Deceleration Lane with Different Lengths*, Transportation Research Board, Washington D.C., 2006
11. Cassidy, M.J., Anani, S.B., and Haigwood, J.M., *Study of Freeway Traffic Near an Off-Ramp*, California PATH Working Paper, 2000, UCB-ITS-PWP-2000-10.
12. Janson, B.N., Awad, W., and Robles, J., *Truck Accidents at Freeway Ramps: Data Analysis and High-Risk Site Identification*, Journal of Transportation and Statistics, January 1998, p75-92.
13. Chen, H., Liu, P., Behzadi, B., and Lu, J.J., *Impacts of Exit Ramp Type on the Safety Performance of Freeway Diverge Areas*, Transportation Research Board, Washington D.C., 2008
14. Wang, Z., Chen, H., and Lu, J.J., *Exploring Impacts of Factors Contributing to Injury Severity at Freeway Diverge Areas*, Transportation Research Board, Washington D.C., 2009

15. O' Donnell, C. and Connor, D., *Predicting the Severity of Motor Vehicle Accident Injuries Using Models of Ordered Multiple Choice*. Accident Analysis and Prevention, Vol. 28, No. 6, 1996, p 739-753.
16. Zajac, S.S. and Ivan., J.N., *Factors influencing injury severity of motor vehicle–crossing pedestrian crashes in rural Connecticut*. Accident Analysis and Prevention, Vol. 35, No. 3, 2003, p 369-379.
17. Kockelman, K.M. and Kweon, Y.J., *Driver injury severity: an application of ordered probit models*. Accident Analysis and Prevention, Vol. 34, No. 3, 2002, p 313-321.
18. Abdel-Aty, M. and J. Keller., *Exploring the Overall and Specific Crash Severity Levels at Signalized Intersections*. Accident Analysis and Prevention, Vol. 37, 2005, pp. 417-425.
19. Abdel-Aty, M., *Analysis of Driver Injury Severity Levels at Multiple Locations Using Ordered Probit Models*. Journal of Safety Research, Vol. 34, 2003, p 597-603.
20. Shankar, V., and Mannering, F. *An Exploratory Multinomial Logit Analysis of Single Vehicle Motorcycle Accident Severity*. Journal of Safety Research, Vol 27, No 3, 1996, p 183-194.
21. Carson J., and Mannering, F., *The Effect of Ice Warning Signs on Ice-accident Frequencies and Severities*. Accident Analysis and Prevention, Vol. 33, Issue 1, 2001, p 99-109.
22. Ulfarsson, G., and Mannering, F., *Differences in Male and Female Injury Severities in Sport-Utility Vehicle, Minivan, Pickup and Passenger Car Accidents*. Accident Analysis and Prevention, Vol. 36, No. 2, 2004, p 135-147.
23. Yamamotoa, T. and Shankar, V.N., *Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects*. Accident Analysis and Prevention, Vol. 36, No. 2, 2004, p 869-876.
24. Xie, Y., Zhang, Y., and Liang F., *Crash Injury Severity Analysis Using a Bayesian Ordered Probit Model*. Transportation Research Board, Washington, D.C., 2007
25. Das, A., Pande, A., Abdel-Aty, M., and Santos, J., *Urban Arterial Crash Characteristics Related to Proximity to Intersections and Injury Severity*. Transportation Research Board, Washington, D.C., 2008
26. Wang, X. and Abdel-Aty, M., *Left-Turn Crash Injury Severity Analyses Using Partial Proportional Odds Models: Significant Factors and Varying Effects*. Transportation Research Board, Washington, D.C., 2008
27. Wang. X. and Kockelman, K.M., *Use of Heteroscedastic Ordered Logit Model to Study Severity of Occupant Injury: Distinguishing Effects of Vehicle Weight and Type*, Transportation Research Record, 2005, pp. 195–204
28. Williams, R., *Generalized Ordered Logit/Partial Proportional Odds Models for Ordinal Dependent Variables*. *The Stata Journal*, Vol. 6, No. 1, 2006, p 58-82.
29. Peterson, B. and Harrell F.E., *Partial Proportional Odds Models for Ordinal Response Variables*. *Applied Statistics*, 1990, 39(2), p 205-217.
30. Long, J. S., *Regression Models for Categorical and Limited Dependent Variables.* Sage Publications, Inc. Thousand Oaks, CA., 1996

31. McCullagh, P. and J.A. Nelder., *Generalized Linear Models*. Second Edition. New York: Chapman and Hall, 1989
32. Williams, R., *Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients across Groups.* Sociological Methods & Research 37(4), 2009, p. 531-559

# Characteristics of the Work Zone Crashes

Sreekanth Reddy Akepati[1] and Sunanda Dissanayake, Ph.D., P.E.[2]

[1]Graduate Research Assistant, Department of Civil Engineering, Kansas State University, Manhattan, Kansas 66506-5000; srireddy@ksu.edu
[2]Associate Professor, Department of Civil Engineering, Kansas State University, 2118 Fiedler Hall, Manhattan, Kansas 66506-5000; tel: 785-323-1540, fax: 785-532-7717, email: sunanda@ksu.edu

## ABSTRACT

Work zone crashes have accounted for 9,900 fatalities in the United States in the last ten years. As the traffic demand on the highways is increasing, management of such traffic flows in work zones has become a challenging task for many highway agencies and understanding crash characteristics would be valuable for highway agencies in setting up proper traffic management plans at work zones based on the prevailing conditions. Accordingly, this study investigated characteristics of work zone crashes in Iowa, Kansas, Missouri, Nebraska and Wisconsin, states currently included in the Smart Work Zone Deployment Initiative (SWZDI) region. Work zone crash data were obtained from respective departments of transportations and crash data from years 2002 to 2006 were considered in this study**.** The characteristics related to work zone crashes such as environmental conditions, vehicles, crashes, driver, and roadway factors were analyzed for the combined states. Results showed that most of the work zone crashes occurred under clear environmental conditions as during daylight, no adverse weather, etc. Multiple-vehicle crashes were more predominant than single-vehicle crashes in work zone crashes. Primary driver-contributing factors of work zone crashes were inattentive driving, following too close for conditions, failure to yield right of way, driving too fast for conditions, and exceeding posted speed limits within work zones. Middle aged drivers were more involved in work zone crashes and most of the work zone crashes lead to Property Damage Only (PDO) crashes. In order to find the relation between crash severity and other variables a test of independency was performed between them using chi-square test methodology.

## INTRODUCTION

As the construction of most of the major highway networks in the United States has already been completed, the majority of current highway work includes maintenance and rehabilitation of those highways, during which work zones are generated. The safety in these work zones has become a major concern among many highway agencies. In the past ten years i.e., for the period 1998-2007 nearly 413,461 fatalities occurred in the US out of which nearly 9,900 (2.4%) fatalities occurred near work zones and there trends are shown in Figure 1.

In order to improve the safety and efficiency of traffic operations in work zones, the states of Iowa (the leading state), Kansas, Missouri and Nebraska created Smart Work Zone Deployment Initiative (SWZDI). Through this pool-funded study, researchers investigate better ways of controlling traffic through work zones, therefore, improving

the safety and efficiency of traffic operations and highway workers. This study identifies the contributory causes and characteristics of the work zone crashes using data from SWZDI region for the five year period from 2002 to 2006.

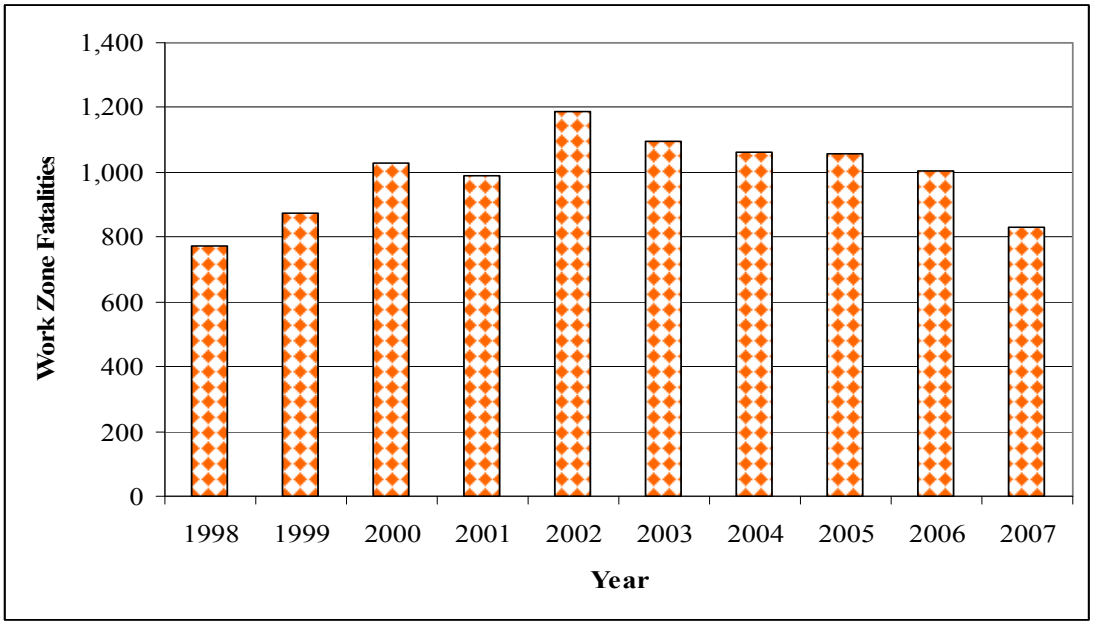



**Figure 1. Work zone fatalities trend in the United States (FARS 2009)**

## LITERATURE REVIEW

The research studies that were related to characteristic analysis of work zone crashes by different authors have been discussed briefly as follows.

Many Studies were conducted to identify the predominant locations within work zone areas where crashes occurred (Garber and Zhao 2002, Hargrove's 1981, Nemeth and Migletz 1978, Nemeth and Rathi 1983). In their study, the entire work zone was divided into different areas such as (i) advance warning area, (ii) transition area, (iii) longitudinal buffer area, (iv) activity area, and (v) termination area. Garber and Zhao (2002) found that 70% of work zone crashes in Virginia occurred at the activity area. Hargrove's (1981) also found in his study that most of the crashes occurred in work area (combining the longitudinal buffer area and the activity area) with 44.7% of total work zone crashes. Nemeth and Migletz (1978) concluded that 39.1% and 16.6% of accidents occurred in the longitudinal buffer area and activity areas respectively. In another study by Nemeth and Rathi (1983), a different set of location categories was used such as advance zone, taper zone, crossover zone and bidirectional zone. Most of the crashes were found to have occurred in crossover and bidirectional (two-lane, two-way operation) zones.

Comparative studies between work zone crash characteristics and non work zone characteristics were done by many researchers (Pigment and Agent 1990, Hall and Lorenz 1989, Chambless, et al. 2002, Daniel et al. 2000). Pigment and Agent (1990) found that in Kentucky, 25.7% of work zone crashes were involved by trucks when compared with 9.6% of non-work zone highway crashes and most work zone crashes

occurred on interstate routes. Results also showed that the percentage of rear-end and same direction sideswipe crashes in work zone crashes was almost three times the percentage of the same types of crashes in the statewide non-work zone crashes. Hall and Lorenz (1989) concluded that in New Mexico, the proportion of rear-end collisions increased from 9.4% before construction to 13.8% during construction. In addition, the researchers concluded 1) the proportion of crashes caused mainly by following too close was much higher in during-work-zone periods than in before-work-zone periods; 2) in comparison with the identical period in the prior year, crashes in construction areas increased 33% on the rural interstate system. Chambless, et. al. (2002) also compared the work zone and non work zone characteristics for Alabama, Michigan and Tennessee and concluded that 63% of work zone crashes take place on interstate, U.S. and state roads, as compared to only 37% of non work zone crashes. Misjudging stopping distance and following too closely accounted for 27% of work zone crashes whereas 15% of these types of crashes took place in non work zone areas. The crashes occurring at the speed limits 45 and 55mph were more predominant (48%) when compared non work zone crashes (24%). Daniel et al. (2000) also found that passenger vehicles are highly involved in both the fatal crashes, whereas the involvement of the trucks in work zone fatal crashes (20%) is significantly higher in proportion when compared to non-work zone (13%) locations. A higher proportion of fatal crashes occurred on rural roadways when compared to urban roadways for both work zone and non-work zone locations. The primary contributing factors to fatal crashes in work zones are driver lost control, failed to yield and too fast for conditions which accounted for nearly 38% of all fatal crashes within work zones.

## DATA

Work zone crash data for the SWZDI region states were obtained from the respective departments of transportations for the period 2002-2006. The data that was gathered divided into different categories such as environmental related factors, vehicle related factors, driver related contributory factors, crash related factors, road characteristics, and other contributing factors which prevail or contribute to crashes in work zones. Crash files from each state were merged by matching the unique crash identification codes using Statistical Analysis System (SAS) software.

As data for this project came from five different states, considerable complications were encountered while comparing or combining similar parameters among the five states. Characteristics considered from the data sets were not always described elaborately creating difficulty in understanding their precise definitions. Lack of exposure-related factors in the data sets, such as the number of vehicles passing through the work zones during daytime and nighttime, length and duration of work zone, status of the work whether active or inactive at the time of crash, etc. limited the study in terms of analyzing the work zone crashes more precisely.

## RESULTS AND DISCUSSION

Characteristics of work zone crashes obtained using crash data from SWZDI region are presented here.

Crashes occurring under different environmental conditions were analyzed to identify characteristics of work zone crashes as shown in Table 1. Based on the total, majority of crashes occurred on daylight conditions (75.3%) with no adverse weather conditions (68.9%).

**Table1. Environmental Related Work Zone Crash Characteristics**

| Description | Iowa | | Kansas | | Missouri | | Nebraska | | Wisconsin | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. | % | No. | % | No. | % | No. | % | No. | % | No. | % |
| **Light Condition** | | | | | | | | | | | | |
| Daylight | 2,915 | 79.0% | 6,617 | 73.1% | 14,792 | 76.5% | 2,064 | 71.7% | 6,753 | 74.7% | 33,141 | 75.3% |
| Dawn or Dusk | 99 | 2.7% | 331 | 3.7% | 0 | 0.0% | 123 | 4.3% | 268 | 3.0% | 821 | 1.9% |
| Lighted | 380 | 10.3% | 1,062 | 11.7% | 2,163 | 11.2% | 339 | 11.8% | 1,157 | 12.8% | 5,101 | 11.6% |
| Dark | 273 | 7.4% | 1,010 | 11.2% | 2,105 | 10.9% | 327 | 11.4% | 824 | 9.1% | 4,539 | 10.3% |
| Unknown | 22 | 0.6% | 32 | 0.4% | 280 | 1.4% | 25 | 0.9% | 43 | 0.5% | 402 | 0.9% |
| Total | 3,689 | 100% | 9,052 | 100% | 19,340 | 100% | 2,878 | 100% | 9,045 | 100% | 44,004 | 100% |
| **Weather Condition** | | | | | | | | | | | | |
| Clear | 2,154 | 58.4% | 7,986 | 88.2% | 12,996 | 67.2% | 2,058 | 71.5% | 5,133 | 56.7% | 30,327 | 68.9% |
| Cloudy | 1,124 | 30.5% | 0 | 0.0% | 4,356 | 22.5% | 531 | 18.5% | 2,894 | 32.0% | 8,905 | 20.2% |
| Rain | 308 | 8.3% | 762 | 8.4% | 1,055 | 5.5% | 119 | 4.1% | 722 | 8.0% | 2,966 | 6.7% |
| Snow | 26 | 0.7% | 115 | 1.3% | 139 | 0.7% | 73 | 2.5% | 144 | 1.6% | 497 | 1.1% |
| Winds | 22 | 0.6% | 77 | 0.9% | 0 | 0.0% | 19 | 0.7% | 21 | 0.2% | 139 | 0.3% |
| Unknown/Other | 55 | 1.5% | 112 | 1.2% | 794 | 4.1% | 78 | 2.7% | 131 | 1.4% | 1,170 | 2.7% |
| Total | 3,689 | 100% | 9,052 | 100% | 19,340 | 100% | 2,878 | 100% | 9,045 | 100% | 44,004 | 100% |

Crash-related work zone characteristics are shown in Table 2. Crash statistics showed that the majority of work zones crashes in the five states are PDO crashes. However, nearly 296 persons died in work zones during the five-year period considered and 27.2% of the total work zone crashes led to injuries. Collision with other moving vehicle is one of the most predominant crashes which consisted of 73.3% of total work zone crashes. When considering collisions with other motor vehicle, rear-end collision (42.7%) was the most frequent type of crash in work zones followed by angle (14.4%) collisions. This might be due to reduced traffic lanes creating more congestion in work zones, which tends to increase interaction between the vehicles possibly leading to rear-end collisions.

It is very important to analyze the area within a work zone, and type of work zone, where most of the crashes occurred. As only the Iowa and Nebraska data sets had these work zone-related details, the analyzed characteristics of these variables for the given two states are shown in Table 3. Results showed that in these two states, the majority of the crashes occurred in a lane-closure (37%) type of work zone. In terms of location within work zone areas, the highest proportion (47.6%) of crashes occurred in the activity area where the actual work was done supporting other research works (Garber and Zhao 2002, Hargrove's 1981, Nemeth and Migletz 1978, Nemeth and Rathi 1983).

**Table 2. Crash Related Work Zone Crash Characteristics**

| Description | Iowa | | Kansas | | Missouri | | Nebraska | | Wisconsin | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. | % | No. | % | No. | % | No. | % | No. | % | No. | % |
| **Crash Severity** | | | | | | | | | | | | |
| Fatal | 26 | 0.7% | 70 | 0.8% | 100 | 0.5% | 41 | 1.4% | 59 | 0.7% | 296 | 0.7% |
| Injury | 1,259 | 34.1% | 2,112 | 23.3% | 4,342 | 22.5% | 1,183 | 41.1% | 3,059 | 33.8% | 11,955 | 27.2% |
| PDO | 2,404 | 65.2% | 6,870 | 75.9% | 14,898 | 77.0% | 1,654 | 57.5% | 5,927 | 65.5% | 31,753 | 72.2% |
| Total | 3,689 | 100% | 9,052 | 100% | 19,340 | 100% | 2,878 | 100% | 9,045 | 100% | 44,004 | 100% |
| **Crash Class** | | | | | | | | | | | | |
| Overturn | 90 | 2.4% | 269 | 3.0% | 273 | 1.4% | 187 | 6.5% | 196 | 2.2% | 1,015 | 2.3% |
| Parked Motor Vehicle | 102 | 2.8% | 255 | 2.8% | 529 | 2.7% | 29 | 1.0% | 184 | 2.0% | 1,099 | 2.5% |
| Animal | 14 | 0.4% | 528 | 5.8% | 112 | 0.6% | 111 | 3.9% | 33 | 0.4% | 798 | 1.8% |
| Vehicle in Transit | 2,738 | 74.2% | 6,359 | 70.2% | 14,676 | 75.9% | 2,079 | 72.2% | 6,422 | 71.0% | 32,274 | 73.3% |
| Fixed Object | 319 | 8.6% | 1,124 | 12.4% | 2,558 | 13.2% | 260 | 9.0% | 1,248 | 13.8% | 5,509 | 12.5% |
| Other | 426 | 11.5% | 517 | 5.7% | 1,192 | 6.2% | 212 | 7.4% | 962 | 10.6% | 3,309 | 7.5% |
| Total | 3,689 | 100% | 9,052 | 100% | 19,340 | 100% | 2,878 | 100% | 9,045 | 100% | 44,004 | 100% |
| **Collision Manner** | | | | | | | | | | | | |
| Head On | 48 | 1.3% | 87 | 1.0% | 185 | 1.0% | 14 | 0.5% | 126 | 1.4% | 460 | 1.0% |
| Rear End | 1,796 | 48.7% | 3,741 | 41.3% | 8,571 | 44.3% | 1,145 | 39.8% | 3,547 | 39.2% | 18,800 | 42.7% |
| Angle | 145 | 3.9% | 1,481 | 16.4% | 2,693 | 13.9% | 380 | 13.2% | 1,652 | 18.3% | 6,351 | 14.4% |
| Sideswipe | 589 | 16.0% | 824 | 9.1% | 2,966 | 15.3% | 353 | 12.3% | 1,227 | 13.6% | 5,959 | 13.5% |
| No Collision | 714 | 19.4% | 75 | 0.8% | 4,100 | 21.2% | 798 | 27.7% | 2,405 | 26.6% | 8,092 | 18.4% |
| Unknown/Other | 397 | 10.8% | 2,844 | 31.4% | 825 | 4.3% | 188 | 6.5% | 88 | 1.0% | 4,342 | 9.9% |
| Total | 3,689 | 100% | 9,052 | 100% | 19,340 | 100% | 2,878 | 100% | 9,045 | 100% | 44,004 | 100% |

**Table 3. Work Zone Area Related Crash Characteristics**

| Description | Iowa | | Nebraska | | Total | |
|---|---|---|---|---|---|---|
| | No. | % | No. | % | No. | % |
| **Within Work Zone Area** | | | | | | |
| Advance Warning Area | 251 | 6.8% | 112 | 3.9% | 363 | 5.5% |
| Between advance warning sign and work area | 563 | 15.3% | 418 | 14.5% | 981 | 14.9% |
| Within transition area for lane shift | 627 | 17.0% | 513 | 17.8% | 1,140 | 17.4% |
| Within or adjacent to work activity | 1,486 | 40.3% | 1,642 | 57.1% | 3,128 | 47.6% |
| Between end of work area and "End Work Zone" sign | 109 | 3.0% | 175 | 6.1% | 284 | 4.3% |
| Other work zone area | 496 | 13.4% | 18 | 0.6% | 514 | 7.8% |
| Unknown | 157 | 4.3% | 0 | 0.0% | 157 | 2.4% |
| Total | 3,689 | 100% | 2,878 | 100% | 6,567 | 100% |
| **Work Zone Type** | | | | | | |
| Lane closure | 1,567 | 42.5% | 862 | 30.0% | 2,429 | 37.0% |
| Lane shift/crossover/head-to-head traffic | 442 | 12.0% | 540 | 18.8% | 982 | 15.0% |
| Work on Shoulder or Median | 554 | 15.0% | 630 | 21.9% | 1,184 | 18.0% |
| Intermittent or Moving Work | 185 | 5.0% | 384 | 13.3% | 569 | 8.7% |
| Other type of work zone | 739 | 20.0% | 439 | 15.3% | 1,178 | 17.9% |
| Unknown | 202 | 5.5% | 23 | 0.8% | 225 | 3.4% |
| Total | 3,689 | 100% | 2878 | 100% | 6,567 | 100% |

Descriptive information about the crashes is shown in Table 4. This included vehicle maneuvers before the crash, vehicle body type, and number of vehicles involved.

**Table 4. Vehicle Related Work Zone Crash Characteristics**

| Description | Iowa | | Kansas | | Missouri | | Nebraska | | Wisconsin | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. | % | No. | % | No. | % | No. | % | No. | % | No. | % |
| **Vehicle Maneuvering** | | | | | | | | | | | | |
| Going Straight | 4,027 | 54.6% | 8,531 | 50.9% | 27,364 | 71.0% | 3,227 | 59.3% | 7,526 | 46.9% | 50,675 | 60.2% |
| Turning Left | 383 | 5.2% | 889 | 5.3% | 1,031 | 2.7% | 409 | 7.5% | 1,225 | 7.6% | 3,937 | 4.7% |
| Turning Right | 158 | 2.1% | 381 | 2.3% | 504 | 1.3% | 101 | 1.9% | 561 | 3.5% | 1,705 | 2.0% |
| Making U-Turn | 26 | 0.4% | 63 | 0.4% | 46 | 0.1% | 22 | 0.4% | 58 | 0.4% | 215 | 0.3% |
| Overtaking | 43 | 0.6% | 159 | 0.9% | 256 | 0.7% | 67 | 1.2% | 158 | 1.0% | 683 | 0.8% |
| Changing Lanes | 267 | 3.6% | 593 | 3.5% | 708 | 1.8% | 176 | 3.2% | 585 | 3.6% | 2,329 | 2.8% |
| Backing | 104 | 1.4% | 303 | 1.8% | 397 | 1.0% | 43 | 0.8% | 344 | 2.1% | 1,191 | 1.4% |
| Slowing or Stopping | 1,323 | 17.9% | 2,125 | 12.7% | 1,571 | 4.1% | 0 | 0.0% | 2,554 | 15.9% | 7,573 | 9.0% |
| Stopped in Traffic | 315 | 4.3% | 2,547 | 15.2% | 4,655 | 12.1% | 1,137 | 20.9% | 1,596 | 9.9% | 10,250 | 12.2% |
| Merging | 293 | 4.0% | 339 | 2.0% | 0 | 0.0% | 73 | 1.3% | 386 | 2.4% | 1,091 | 1.3% |
| Parked | 146 | 2.0% | 31 | 0.2% | 74 | 0.2% | 4 | 0.1% | 286 | 1.8% | 541 | 0.6% |
| Unknown | 295 | 4.0% | 797 | 4.8% | 1,925 | 5.0% | 181 | 3.3% | 767 | 4.8% | 3,981 | 4.7% |
| Total | 7,380 | 100% | 16,758 | 100% | 38,531 | 100% | 5,440 | 100% | 16,046 | 100% | 84,155 | 100% |
| **Crash Type** | | | | | | | | | | | | |
| Single Vehicle | 691 | 18.7% | 2,631 | 29.1% | 3,626 | 18.7% | 772 | 26.8% | 2,174 | 24.0% | 9,894 | 22.5% |
| Two Vehicles | 2,483 | 67.3% | 5,420 | 59.9% | 13,438 | 69.5% | 1,751 | 60.8% | 5,855 | 64.7% | 28,947 | 65.8% |
| >2 Vehicles | 515 | 14.0% | 1,001 | 11.1% | 2,276 | 11.8% | 355 | 12.3% | 1,016 | 11.2% | 5,163 | 11.7% |
| Total | 3,689 | 100% | 9,052 | 100% | 19,340 | 100% | 2,878 | 100% | 9,045 | 100% | 44,004 | 100% |
| **Vehicle Body Type** | | | | | | | | | | | | |
| Automobile | 3,946 | 53.5% | 8,775 | 52.4% | 18,855 | 48.9% | NA | NA | 11,220 | 69.9% | 42,796 | 54.4% |
| Motor Cycle | 52 | 0.7% | 119 | 0.7% | 205 | 0.5% | NA | NA | 221 | 1.4% | 597 | 0.8% |
| Light-Duty Vehicle | 2,471 | 33.5% | 6,324 | 37.7% | 13,667 | 35.5% | NA | NA | 2,115 | 13.2% | 24,577 | 31.2% |
| Heavy Duty Vehicle | 685 | 9.3% | 1,257 | 7.5% | 4,422 | 11.5% | NA | NA | 1,781 | 11.1% | 8,145 | 10.3% |
| Unknown/Other | 226 | 3.1% | 283 | 1.7% | 1,382 | 3.6% | NA | NA | 709 | 4.4% | 2,600 | 3.3% |
| Total | 7,380 | 100% | 16,758 | 100% | 38,531 | 100% | NA | NA | 16,046 | 100% | 78,715 | 100% |

NA-Not Available

As a result of construction and maintenance work activity on highways, lane widths were reduced to less than normal width, which increases the interaction between vehicles leading to multiple-vehicle crashes. Results showed the majority (65.8%) of the work zone crashes are multiple-vehicle crashes. These multiple-vehicle crashes occurred when the vehicles were going straight (60.2%) in work zones. Critical maneuvers such as left turns, right turns, and u-turns in work zones contribute to a small percentage of crashes, but a predominant percent (21.2%) of crashes occurred when the vehicles are slowing and stopped in traffic due to work activity. Based on the data availability, vehicle body type was categorized into three types such as automobile, light-duty vehicles, and heavy-duty vehicles. More than 50% of work zone crashes involved passenger cars, as

the major portion of traffic consists of passenger cars. Although it was not possible to normalize the results, they showed that a majority of work zone crashes involved passenger cars. In addition to passenger cars, light-duty vehicles such as pickup trucks, vans, and SUVs contributed to the second highest percentage of work zone crashes. In terms of heavy-duty vehicles such as trucks, these require additional consideration in work zones as their characteristics are different from other vehicles. According to the Federal Highway Administration (FHWA), almost 30% of work zone crashes involved trucks. Analysis showed that 10.3% of work zone crashes involved heavy-duty vehicles. The vehicle body type variable was incomplete in the data obtained from Nebraska Department of Roads.

**Table 5. Driver Related Work Zone Crash Characteristics**

| Description | Iowa | | Kansas | | Missouri | | Nebraska | | Wisconsin | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. | % | No. | % | No. | % | No. | % | No. | % | No. | % |
| **Driver Age** | | | | | | | | | | | | |
| Young Age | 1,911 | 25.9% | 4,924 | 29.5% | 8,209 | 21.3% | 1,424 | 26.2% | 3,827 | 23.9% | 20,295 | 24.1% |
| Middle Age | 4,523 | 61.3% | 10,462 | 62.6% | 25,059 | 65.0% | 3,418 | 62.8% | 10,073 | 62.8% | 53,535 | 63.6% |
| Old Age | 615 | 8.3% | 1,330 | 8.0% | 2,847 | 7.4% | 458 | 8.4% | 1,305 | 8.1% | 6,555 | 7.8% |
| Unknown | 331 | 4.5% | 0 | 0.0% | 2,416 | 6.3% | 140 | 2.6% | 841 | 5.2% | 3,728 | 4.4% |
| Total | 7,380 | 100% | 16,716 | 100% | 38,531 | 100% | 5,440 | 100% | 16,046 | 100% | 84,113 | 100% |
| **Driver Gender** | | | | | | | | | | | | |
| Male | 4,170 | 56.5% | 9,837 | 58.8% | 22,318 | 57.9% | 3,299 | 60.6% | 0 | 0.0% | 39,624 | 58.2% |
| Female | 2,890 | 39.2% | 6,456 | 38.6% | 13,564 | 35.2% | 1,985 | 36.5% | 0 | 0.0% | 24,895 | 36.6% |
| Unknown | 320 | 4.3% | 423 | 2.5% | 2,649 | 6.9% | 156 | 2.9% | 0 | 0.0% | 3,548 | 5.2% |
| Total | 7,380 | 100% | 16,716 | 100% | 38,531 | 100% | 5,440 | 100% | 0 | 0.0% | 68,067 | 100% |
| **Driver Contributing Circumstance** | | | | | | | | | | | | |
| Disregarded Traffic Controls | 140 | 1.9% | 513 | 5.0% | 331 | 1.7% | 91 | 3.5% | 340 | 3.1% | 1,415 | 2.8% |
| Exceeded Posted Speed Limit | 37 | 0.5% | 111 | 1.1% | 455 | 2.3% | 17 | 0.6% | 246 | 2.2% | 866 | 1.7% |
| Driving Too Fast for Conditions | 295 | 4.0% | 926 | 9.0% | 3,038 | 15.3% | 102 | 3.9% | 874 | 8.0% | 5,235 | 10.2% |
| Made Improper Turn | 67 | 0.9% | 244 | 2.4% | 394 | 2.0% | 14 | 0.5% | 223 | 2.0% | 942 | 1.8% |
| Following Too Close | 713 | 9.7% | 1,763 | 17.1% | 4,397 | 22.1% | 339 | 12.9% | 1,265 | 11.5% | 8,477 | 16.6% |
| Inattention | 68 | 0.9% | 4,183 | 40.6% | 4,292 | 21.6% | 234 | 8.9% | 1,961 | 17.9% | 10,738 | 21.0% |
| Failed to Yield Right of Way | 593 | 8.0% | 759 | 7.4% | 1,725 | 8.7% | 195 | 7.4% | 1,486 | 13.5% | 4,758 | 9.3% |
| Other | 4,839 | 65.6% | 1,772 | 17.2% | 4,710 | 23.7% | 1,552 | 59.1% | 2,009 | 18.3% | 14,882 | 29.1% |
| Unknown | 628 | 8.5% | 26 | 0.3% | 514 | 2.6% | 83 | 3.2% | 2,578 | 23.5% | 3,829 | 7.5% |
| Total | 7,380 | 100% | 10,297 | 100% | 19,856 | 100% | 2,627 | 100% | 10,982 | 100% | 51,142 | 100% |

The driver plays a key role in involvement in a crash, and identification of driver contribution to crashes is highly important in suggesting possible countermeasures. Work zone crashes based on driver-contributing circumstances is shown in Table 5. Driver ages

are divided into three age groups such as young age (≤ 24 Years), Middle Age (> 24 to ≤ 64 Years) and Old age (≥ 65 years or Older). Results showed the majority (63.6%) of work zone crashes involved males aged 25 to 64 years. This may be due to males tending to drive more than females. Older age people were involved in a small but predominant percent (7.8%) of work zone crashes.

Of all work zone crashes considered for the five states, inattentive driving (21%) in work zones was the leading cause of crash occurrence. This might be due to the fact that most of the drivers were unaware of the general problems associated with work zones. Among other factors, following too close was responsible for 16.6% of total work zone crashes, which might be due to interruption of regular traffic flows caused by closed lanes in work zone areas. Driving too fast for conditions and exceeding posted speed limits were other predominant contributing factors in work zone crashes.

**Chi-Square Test of Independency**

This method tests the relation between two variables using chi-square distribution. The hypotheses for this test of independence are as follows:

$H_o$: The two variables are "independent" of each other; and

$H_a$: The two variables are "dependent" on each other

Where $H_o$ is the null hypothesis and $H_a$ is the alternative hypothesis.

The test is performed for light condition vs accident severity. The observed frequencies are shown in Table 6.

**Table 6. Observed Values for Light Condition vs Crash Severity**

| Light Conditions | Crash Severity | | | Total |
|---|---|---|---|---|
| | Fatal | Injury | PDO | |
| Daylight | $n_{11}$=175 | $n_{12}$ = 8,787 | $n_{13}$ = 24,179 | $n_{1+}$= 33,141 |
| Poor Visible Conditions | $n_{21}$=121 | $n_{22}$ = 3,168 | $n_{23}$ = 7,574 | $n_{2+}$ = 10,863 |
| Total | $n_{+1}$=296 | $n_{+2}$ =11,955 | $n_{+3}$ = 1,753 | n = 44,004 |

**Table 7. Expected Values for Light Condition vs Crash Severity**

| Light Condition | Fatal | Injury | PDO | Total |
|---|---|---|---|---|
| Daylight | 222.9 | 9,003.7 | 23,914.3 | 33,141 |
| Poor Visibility Conditions | 73.1 | 2,951.3 | 7,838.7 | 10,863 |
| Total | 296 | 11955 | 31753 | 44,004 |

Expected values are calculated based on the assumption that the null hypothesis is true. The expected frequency for the $n_{11}$ can be calculated as shown in equation 1 and the calculated expected values are shown in Table 7.

$$n_{11} = \frac{(n_{1+}) \times (n_{+1})}{(n)} \quad (1)$$

The Chi-Square value is calculated using the following formula shown in equation 2.

$$\chi^2 = \sum \frac{(\text{Observed Frequency} - \text{Expected Frequency})^2}{\text{Expected Frequency}} \quad (2)$$

Calculated chi-square value is compared with the tabular values corresponding to the degree of freedom and user-defined confidence levels. The degree of freedom can be obtained by multiplying (Number of rows-1)* (Number of columns -1). For the values shown in Table 7, the value of the test statistic is $\chi^2$= 74.7. At a 95% confidence level, the value shown in the table for two degrees of freedom is 5.991. Since the calculated $\chi^2$ > the table value, the null hypothesis is rejected and it can be concluded that light condition and crash severity are related. Similarly, test of independence was carried out for all other variables considered in this study and the results are presented in Table 8.

**Table 8. Dependency Relation of Crash Severity with Different Variables**

| Category | Degree of Freedom | Chi-Square Calculated | Table Value | P-Value | Statistical Significance |
|---|---|---|---|---|---|
| Light Conditions | 2 | 74.7 | 6 | P < 0.01 | Yes |
| Weather Condition | 8 | 215.05 | 15.51 | P < 0.01 | Yes |
| Posted Speed Limit | 12 | 431.55 | 21.03 | P < 0.01 | Yes |
| Surface Condition of Road | 4 | 6.3 | 9.5 | P > 0.01 | No |
| Road Surface Type | 4 | 31 | 9.5 | P < 0.01 | Yes |
| Traffic Controls | 12 | 173.4 | 21 | P < 0.01 | Yes |
| Driver Gender | 2 | 59 | 6 | P < 0.01 | Yes |
| Day of Crash | 12 | 65.6 | 21 | P < 0.01 | Yes |
| Age of Driver | 12 | 34.9 | 21 | P < 0.01 | Yes |
| Vehicle Maneuver Before Crash | 10 | 199 | 18.3 | P < 0.01 | Yes |
| Alcohol Involvement | 2 | 478.3 | 6 | P < 0.01 | Yes |
| Number of Vehicle Involved | 4 | 1148 | 9.5 | P < 0.01 | Yes |
| Manner of Collision | 10 | 726.9 | 18.3 | P < 0.01 | Yes |
| Vehicle Body Type | 14 | 1056.2 | 23.7 | P <0.01 | Yes |
| Driver Contributing Circumstances | 22 | 795.7 | 33.9 | P < 0.01 | Yes |

## CONCLUSIONS

According to analysis results, in all five states, most of the work zone crashes occurred under clear environmental conditions. Multiple-vehicle crashes were more predominant in work zone crashes when compared to crashes involving a single vehicle. A majority of the work zone crashes led to PDO crashes and a few but noticeable percentage of fatal crashes occurred in work zones. At the time of occurrence of a crash, a majority of vehicles involved were going straight. Further, a predominant percentage of vehicles were stopped in traffic or slowing down for a signal. Passenger cars were more involved in work zone crashes when compared to light-duty and heavy-duty vehicles. Rear-end was the most predominant type of collision in work zone areas when compared to other collisions. As of 2006, only two states have tracked work zone-related variables such as type of work zone and location of crash within work zone areas. Results showed that nearly 50% of work zone crashes occurred in the activity area of the work zone where the actual work goes on. The safest zone within work zones was before the work zone warning sign, i.e., advance warning area which warns the traffic what to expect

ahead. The lane-closure work zone type was the one where the highest percentage of crashes occurred, followed by work on the shoulder or median type of work zone. In driver-contributory factors, inattentive driving and following too close for conditions were some of the factors contributing to work zone crashes. Male drivers aged between 25 to 64 years were more involved in work zone crashes when compared to female drivers, as they might be the ones who drive more.

## REFERENCES

Chambless, J., A. M. Ghadiali., J. K. Lindly., J. McFadden (2002). "Multistate Work Zone Crash Characteristics." *Institute of Transportation Engineers (ITE) Journal*, May.

Daniel, J., K. Dixon., D. Jared (2000). "Analysis of Fatal Crashes in Georgia Work Zones." In *Transportation Research Record: Journal of Transportation Research Board,* No.1715, Transportation Research Board of the National Academies, Washington D.C., pp. 18-23.

Fatality Analysis Reporting System (FARS), NHTSA (2009). http://www-fars.nhtsa.dot.gov/Main/index.aspx . Accessed March 22, 2009.

Garber, N. J. and Zhao, M (2002). "Crash Characteristics at Work Zones." Research Report VTRC 02-R12, Virginia Transportation Research Council, Charlottesville, Virginia.

Hall, J. W., and V. M. Lorenz (1989). "Characteristics of Construction Zone Crashes." In *Transportation Research Record: Journal of Transportation Research Board,* No. 1230, Transportation Research Board of the National Academies, Washington, D.C., pp. 20-27.

Hargroves, B. T (1981). "Vehicle Accidents in Highway Work Zones." *Journal of Transportation Engineering*, Vol.107, No.5, pp.525-539.

Iowa Department of Transportation (IDOT) (2009). http//www.ctre.iastate.edu/smartwz/index.cfm. Accessed March 22, 2009.

Nemeth, Z. A., and A. Rathi (1983). "Freeway Work Zone Accident Characteristics." *Transportation Quarterly*, Vol.37, No.1, pp.145-159.

Nemeth, Z. A., and D. J. Migletz (1978). "Accident Characteristics Before, During, and After Safety Upgrading Projects on Ohio's Rural Interstate System." In *Transportation Research Record: Journal of Transportation Research Board,* No. 672, Transportation Research Board of the National Academies, Washington, D.C., pp. 19 – 24.

Pigman, J. G., and K. R. Agent (1990). "Highway Crashes in Construction and Maintenance Work Zones." In *Transportation Research Record: Journal of Transportation Research Board,* No. 1270, Transportation Research Board of the National Academies, Washington D.C., pp. 12 – 21.

# Identification of Risk Factors Associated With Injury Severity of Work Zone Crashes

Sreekanth Reddy Akepati[1] and Sunanda Dissanayake, Ph.D., P.E.[2]

[1]Graduate Research Assistant, Department of Civil Engineering, Kansas State University, Manhattan, Kansas 66506-5000; srireddy@ksu.edu
[2]Associate Professor, Department of Civil Engineering, Kansas State University, 2118 Fiedler Hall, Manhattan, Kansas 66506-5000; Tel: 785-532-1540, Fax: 785-532-7717, Email: sunanda@ksu.edu

## ABSTRACT

Work zone safety has become a major concern to many highway agencies and the highway safety community is looking at various ways of improving the situation. Accordingly, this paper describes an application of ordered probit modeling methodology to identify the factors associated with the injury severity of work zone crashes using a state level dataset. According to the findings of the severity model, work zone crashes involving trucks, following too close, non deployment of airbags, sideswipe collision of the same direction vehicles, crashes occurring on roadway, and crashes occurred while vehicles were taking left/right turns in work zone area are some of the factors that are contributing towards higher injury severity in work zone crashes.

## INTRODUCTION

In the United States, approximately 1,100 people die and 40,000 people are injured annually as a result of motor vehicle crashes in work zones (FHWA, 2005). These numbers may be a result of interruption to regular traffic flow caused by closed traffic lanes, general misunderstanding of problems associated with work zones, or improper usage of traffic control devices. In this regard a number of studies have addressed various aspects of work zone safety issues; however the results are not always consistent with respect to each study. A clear understanding of the risk factors associated within work zones will help in mitigating the severity of work zone crashes. The study mainly investigates the risk factors associated with the injury severity of work zone crashes by using ordered probit model analysis. The model analysis was done for the Iowa work zone crash data for the period 2002-2006. At the end of the analysis the risk factors associated with the injury severities were presented.

## LITERATURE REVIEW

Many studies have been carried out using ordered probit model analysis and there results as summarized briefly even though they are not directly related to work zone safety.

Kockelman and Kweon (2002) found that in terms of the severity of injuries sustained by drivers, the manner of collision, number of involved vehicles, driver

gender, vehicle type, and driver alcohol were associated with more severe injuries. Another study conducted by Ma and Kockelman (2004) found that speeding, following too close, females, older persons and those in passenger cars are more prone to increase in injury severity.

Khattak, Schneider, and Targa (2003) found that dangerous truck-driver behaviors, particularly speeding, reckless driving, alcohol and drug use, non-use of restraints, and traffic control violations are the factors which increase the injury severities in large truck rollover crashes. Khattak, Pawlovich, Souleyrette, and Hallmark (2002) also conducted a study using ordered probit modeling to isolate factors that contribute to more severe injuries to older drivers involved in traffic crashes and found that alcohol-related crashes and crashes involving farm vehicles were more likely to cause serious injuries to older drivers.

Ratnayake (2004) carried out an analysis using Kansas Accident Reporting System (KARS) crash data considering all ages who met with a crash during 1999 to 2002. Ordered probit modeling was used to investigate the critical factors contributing towards higher crash severity in rural/urban highway crashes. According to the author, alcohol involvement, excessive speed, driver ejection, curved and graded roads, etc. was contributory factors for high-severity crashes.

It is the usual practice to report crash or injury severity in three or more categories such as fatal, incapacitating, property damage only, etc. This makes it possible to order the severity level from most severe to less severe. In other words, the severity, the response variable in the model, could be considered as an ordinal variable. This type of variable can be modeled using ordered choice models. This phenomenon has been applied to model injury severity using both ordered probit and ordered logit models by O'Donnell and Conner (1996). In this study, they considered comparatively higher number of factors to model injury severity. They found that factors such as alcohol involvement, lack of seatbelt usage, occupant being a female and excessive speed were significant towards increased injury severities. According to their conclusion, both ordered probit and ordered logit methods produced similar results in modeling injury severity, although the magnitudes of the estimations were different.

## METHODOLOGY

### Data

Work zone crash data obtained from the Iowa Department of Transportation were used in this study. For the ordered probit modeling analysis, work zone crash data from years 2002 to 2006 were considered. Every occupant involved in a crash/ crashes during the considered period was taken into account with respective injury severity and other related information. This included single-vehicle crashes as well as multi-vehicle crashes. Accordingly, there were about 7,158 occupants involved in work zone crashes during the five-year period.

For the ordered probit analysis, some data lines were deleted where data were missing in at least in one variable. After doing that, about 3,769 crashes involving occupants remained for analysis.

**Ordered Probit Modeling**

When the variable can be ranked or ordered but the difference between two levels are unknown that variable is called an ordinal variable. The response variable in this study, injury severity, can also be ordered as fatal, disabling/incapacitating, non-incapacitating, possible and no injury and thus it can be considered as an ordinal response variable Long (Long, 1997) has discussed the applicability of ordered logit and probit models in detail in his publication.

The derivation of ordered probit model is based on the measurement model as

$$yi = m \qquad if\ \tau_{m-1} \le y^* < \tau_m \qquad for\ m = 1\ to\ J \tag{1}$$

Where y* is the injury risk, which is an unobserved continuous variable called latent variable ranging from -∞ to ∞, which is mapped to an observed variable *y*. The values are called thresholds or cut off points and the extreme categories at *m*=1 and *m*= J are defined by open ended intervals with $\tau_0 = -\infty$ and $\tau_J = \infty$. According to the measurement model the variable y is thought of as providing incomplete information about an underlying y*.

Then the structural model can be considered as,

$$y_i^* = x_i\beta + \varepsilon_i \tag{2}$$

Where $x_i$ is a row of a vector of explanatory variables with a 1 in the first column for the intercept and the $i^{th}$ observation for $x_k$ in column $k+1$. β is a vector of parameters to be estimated and $\varepsilon_i$ is the error term which is assumed to be normally distributed. However, the database does not comprise any information on injury risk, *y** as it is unobserved, but it includes details on the variable *y* which is observed at different levels of *y** at which, y =1 if there are no evident injuries, y =2 if the crash results only possible injuries, y=3 when the crash is non-incapacitating, y = 4 if it is a incapacitating crash and y=5 when crash is fatal. Thus, the measurement model (1) can be illustrated as,

$$y_i = \begin{cases} 1 \rightarrow No\ injury & if\ \tau_0 = -\infty \le y^* < \tau_1 \\ 2 \rightarrow Possible & if\ \tau_1 \le y^* < \tau_2 \\ 3 \rightarrow Non-incapacita\ ting & if\ \tau_2 \le y^* \tau_3 \\ 4 \rightarrow Incapacita\ ting & if\ \tau_3 \le y^* < \tau_4 \\ 5 \rightarrow Fatal & if\ \tau_4 \le y^* < \tau_5 = \infty \end{cases} \tag{3}$$

Where the threshold values $\tau_1, \tau_2, \tau_3$, and $\tau_4$ are parameters to be estimated. According to the measurement model the probability that the i[th] victim of crash, suffer injury severity level of *m* (*m* = 1 to 5) is the probability that the injury propensity *y** takes a value between two cut off points. That is,

$$Pr(y_i = m \mid x_i) = F(\tau_m - x_i\beta) - F(\tau_{m-1} - x_i\beta) \tag{4}$$

Where,

$F$(x) is the cumulative distribution function of the unobserved error term $\varepsilon_i$ evaluated at given x under the assumption that $\varepsilon_i$s are normally distributed with mean zero and constant variance as mentioned previously. For example, the probability that the victim *i* sustain a fatal injury due to the crash is,

$$Pr(y_i = 5 \mid x_i) = 1 - F(\tau_4 - x_i\beta) \tag{5}$$

It should be noted that to these probabilities be positive the thresholds values should satisfy the order, $\tau_1 < \tau_2 < \tau_3 < \tau_4$. The estimation of these model parameters can be carried out through the method of maximum likelihood. The log likelihood, which is the logarithm of the likelihood function, can be written as,

$$\ln L(\beta, \tau \mid y, X) = \sum_{j=1}^{J} \sum_{y_i = j} \ln[F(\tau_j - x_i\beta) - F(\tau_{j-1} - x_i\beta)] \tag{6}$$

Where is the vector of parameters from the structural model, first column consisting of the intercept and $\tau$ is the vector of threshold parameters. The procedure consists of maximizing this equation using numerical methods. To make the model estimable either one threshold value, possibly $\tau_1$ or the intercept is constrained to be some arbitrary value usually zero. The software used in this analysis assumes the intercept *as equal to zero* and estimate the other parameters. For more details on parameter estimation of ordered models using maximum likelihood procedure reader is directed to *Regression Models for Categorical and Limited Dependent Variables* (Long, 1996).

***Goodness of Fit Measure***

In linear regression models, the goodness of fit is usually measured by the $R^2$ value whereas there is no such straightforward measure to evaluate model fitness of ordered probit models. McFadden (1974) suggested using a likelihood ratio index (LRI) that is analogous to the $R^2$ in the linear regression model**.**

$$R^2{}_M = 1 - [\ln L / (\ln L_0)] \tag{7}$$

Where,
$L$ = the value of the maximum likelihood function, and
*Lo* = likelihood function when regression coefficients, except for the intercept term, are zero.

The $R^2_M$ value is bounded by zero and one, where one denotes perfect fit of the model. Similarly, a few other values are given in the SAS output such as Estrella, Adjusted Estrella, Veall-Zimmermann, and McKelvey-Zovoina, which can also be considered in evaluating goodness of fit of a model.

## RESULTS AND DISCUSSION

The ordered probit model was developed to assess the injury severity of work zone crashes by considering nearly 38 explanatory variables using statistical modeling software, SAS version 9.1. The response variable was taken as injury severity. The predictor variable names, description about how variables are determined, and corresponding mean values are given in Table 1.

As the selection criteria for the variables to be included in the model, a 95% confidence level was used in which the probability should be less than 0.05. Co-linearity of individual variables was also checked before considering variables into the model, and if such relationship existed, one of the two correlated variables was discarded based on the lowest mean value criterion.

The parameter estimates of the variables considered in the model are presented in Table 2. Coefficients were estimated using the maximum likelihood method as explained in methodology section. Likelihood ratio indexes (LRI) is presented for the model along with Estrella value and log likelihood value. In this model, the significant variables are denoted by asterisk (*). By looking at the goodness of fit values obtained for the model, it can be stated that the injury severity model for work zone crashes has an acceptable level of fit compared past studies. The likelihood ratio index value for the injury severity model is 0.1250. Thus, the injury severity model has a better capability of explaining injury severities caused to occupants with a selected set of explanatory variables compared to other models developed in the past. Past studies based on ordered probit modeling have shown that the goodness of fit value is typically low. In the model developed by Ma and Kockelman (2004), it was around 0.05 and in the models developed by Kockelman and Kweon (2002), the highest LRI value was around 0.08. There are many other studies in the past which had similar results (O'Donnell and Connor, 1996). Therefore, the reliability of the overall model can be considered as acceptable.

Variables considered in this analysis can be broadly classified under five sections: work zone related, driver related, crash related, roadway related, and environment related. Thus, the discussion of model results is also presented under the same sections for better understanding. A positive estimated coefficient in the model implies increasing injury severity with increasing values of the explanatory variables. Independent variables from each category that were significantly contributing to injury severity are discussed in the following sections.

### *Work Zone Related*

None of the work zone-related variables (location of crash within work zone areas and work zone types) were significant except the variable (WZ_type8) "other work zone" type. This implies if a crash occurs in another work zone type (exact name of work zone was not specified in the database), severity of the resulting crash is going to be less, since the variable had a negative estimated parameter.

**Table 1. Description of Variables Used in the Model**

| Variables | Variable Name | Description | Mean |
|---|---|---|---|
| First Harmful Event | Overturn | If overturn/rollover=1, otherwise=0 | 0.01 |
| | Fixedobj | If collided with fixed object=1, otherwise=0 | 0.06 |
| Manner of Collision | Headon | If it is a head-on collision=1, otherwise=0 | 0.01 |
| | Broad | If it is a broadside collision=1, otherwise=0 | 0.1 |
| | Sideswipe_same | If it is a sideswipe-same direction=1, otherwise=0 | 0.14 |
| | Sideswipe_opp | If it is a sideswipe-opposite direction=1, otherwise=0 | 0.01 |
| Location of First Harm | Onrdway | If a crash occurred on roadway=1, otherwise=0 | 0.95 |
| Weather Conditions | Weathercond | If a work zone crash occurred under no adverse weather conditions =1, otherwise=0 | 0.58 |
| Light Conditions | Lightcond | If a crash occurred in day light conditions=1, otherwise=0 | 0.82 |
| Surface Conditions | Surfcond | If crash occurred on dry road conditions of road=1, otherwise=0 | 0.85 |
| Type of Roadway | Intersectn | If a crash occurred at intersection =1, otherwise=0 | 0.32 |
| Traffic Controls | Trafcntrl | If no traffic controls present =1, otherwise=0 | 0.48 |
| Location within Work Zone | WZ_Loc1 | If crash occurred before work zone warning sign=1, otherwise=0 | 0.08 |
| | WZ_Loc2 | If crash occurred in advance warning area=1, otherwise=0 | 0.17 |
| | WZ_Loc3 | If crash occurred in transition area =1, otherwise=0 | 0.18 |
| | WZ_Loc4 | If crash occurred in activity area =1, otherwise=0 | 0.42 |
| | WZ_Loc5 | If crash occurred in termination area=1, otherwise=0 | 0.03 |
| Work Zone Type | WZ_type2 | If it is lane shift/crossover work zone type=1, otherwise=0 | 0.12 |
| | WZ_type3 | If the work is on shoulder or median=1, otherwise=0 | 0.16 |
| | WZ_type8 | If it is other type of work zone=1, otherwise=0 | 0.19 |
| Workers | Workers | If workers are present=1, otherwise=0 | 0.38 |
| Occupant Protection | Occprotect | If occupant protection is used =1, otherwise=0 | 0.95 |
| Airbag | Airbag_1 | If airbag is not deployed=1, otherwise=0 | 0.72 |
| Vehicle Configuration | Ligdtyveh | If it is a light-duty vehicle=1, otherwise=0 | 0.55 |
| | Truck | If it is a truck ( > 3 Axles) =1, otherwise=0 | 0.09 |
| Vehicle Action | Critmaneu | If the vehicle is making left/right turn=1, otherwise=0 | 0.07 |
| | Passing | If the vehicle is overtaking/passing=1, otherwise=0 | 0.01 |
| | Merging | If the vehicle is changing lanes/merging=1, otherwise=0 | 0.08 |
| | Stopped | If the vehicle is stopped/slowed in traffic=1, otherwise=0 | 0.23 |
| Driver Age | Youngage | If the driver age is in between 0-24 years=1, otherwise=0 | 0.27 |
| Driver Gender | Drivgender | If the driver is male=1, otherwise=0 | 0.59 |
| Driver-Contributing Circumstances | DrivCC_1 | If the driver exceeded posted speed limit=1, otherwise=0 | 0.05 |
| | DrivCC_2 | If the driver is following too close=1, otherwise=0 | 0.1 |
| | DrivCC_3 | If the driver is taking other action=1, otherwise=0 | 0.49 |
| Posted Speed Limit | Speedlimit | Posted speed limit in mph | 45.31 |

**Table 2. Parameter Estimates of the Variables Used in the Model**

| Parameter | Estimate | Standard Error | t Value | Approx Pr > \|t\| |
|---|---|---|---|---|
| Intercept | 2.535957 | 0.207631 | 12.21 | <.0001 |
| Overturn | -0.62822 | 0.144738 | -4.34 | <.0001* |
| Fixedobj | -0.249555 | 0.103922 | -2.4 | 0.0163* |
| Headon | -0.510623 | 0.162692 | -3.14 | 0.0017* |
| Broad | -0.102719 | 0.07925 | -1.3 | 0.1949 |
| Sideswipe_same | 0.491945 | 0.085065 | 5.78 | <.0001* |
| Sideswipe_opp | -0.285423 | 0.186144 | -1.53 | 0.1252 |
| Onrdway | 0.424076 | 0.105473 | 4.02 | <.0001* |
| Weathercond | 0.122153 | 0.047372 | 2.58 | 0.0099* |
| Lightcond | -0.005181 | 0.060819 | -0.09 | 0.9321 |
| Surfcond | -0.01884 | 0.066643 | -0.28 | 0.7774 |
| Intersectn | -0.145374 | 0.053422 | -2.72 | 0.0065* |
| Trafcntrl | 0.045306 | 0.046406 | 0.98 | 0.3289 |
| WZ_Loc1 | 0.030489 | 0.107723 | 0.28 | 0.7772 |
| WZ_Loc2 | 0.061595 | 0.086252 | 0.71 | 0.4751 |
| WZ_Loc3 | 0.124856 | 0.088054 | 1.42 | 0.1562 |
| WZ_Loc4 | 0.072204 | 0.132212 | 0.55 | 0.5850 |
| WZ_Loc5 | 0.088682 | 0.074674 | 1.19 | 0.2350 |
| WZ_type2 | -0.057789 | 0.070798 | -0.82 | 0.4144 |
| WZ_type3 | 0.096443 | 0.06856 | 1.41 | 0.1595 |
| WZ_type8 | -0.21211 | 0.063343 | -3.35 | 0.0008* |
| Workers | -0.03381 | 0.047143 | -0.72 | 0.4733 |
| Occprotect | -0.913787 | 0.088484 | -10.33 | <.0001* |
| Airbag_1 | 0.639337 | 0.04923 | 12.99 | <.0001* |
| Ligdtyveh | 0.101087 | 0.049271 | 2.05 | 0.0402* |
| Truck | 0.834399 | 0.102663 | 8.13 | <.0001* |
| Critmaneu | 0.210727 | 0.095955 | 2.2 | 0.0281* |
| Passing | -0.871643 | 0.250026 | -3.49 | 0.0005* |
| Merging | 0.140812 | 0.095137 | 1.48 | 0.1388 |
| Stopped | 0.068576 | 0.05943 | 1.15 | 0.2485 |
| Youngage | 0.158399 | 0.052115 | 3.04 | 0.0024* |
| Drivgender | 0.233336 | 0.046134 | 5.06 | <.0001* |
| DrivCC_1 | 0.052207 | 0.102294 | 0.51 | 0.6098 |
| DrivCC_2 | 0.567017 | 0.097047 | 5.84 | <.0001* |
| DrivCC_3 | -0.102133 | 0.054379 | -1.88 | 0.0604 |
| Speedlimit | 0.013828 | 0.001967 | 7.03 | <.0001* |
| _Limit2 | 0.910437 | 0.092537 | 9.84 | <.0001 |
| _Limit3 | 1.742421 | 0.097744 | 17.83 | <.0001 |
| _Limit4 | 2.451284 | 0.099206 | 24.71 | <.0001 |
| Estrella | 0.2076 | | | |
| Adjusted Estrella | 0.1886 | | | |
| McFadden's LRI | 0.1250 | | | |
| AIC | 5788 | | | |
| Log Likelihood | -2855 | | | |

*Variables are significant at 0.05 levels

***Driver Related***

In the model, the driver age has been divided into three categories. If driver age ≤24 yrs as young age drivers, if age>24 yrs and age ≤64 yrs as middle age drivers and if age>64 yrs then that drivers were considered as old age drivers (Perera, 2009). The positive estimated parameter statistically significant at a 95% confidence level for the variable 'Youngage' indicates crashes involving young age drivers increase the propensity of more injury severity in work zone crashes. The variable associated with gender 'Drivgender' has a positive estimate, indicating when male drivers are involved in crashes there is a tendency for high injury severity compared to female drivers involved in crashes. This could be due to the fact that males tend to drive more, compared to females, which increases their chances of being involving in a crash.

Whether occupant protection at the time of a crash was used or not was also investigated by including an indicator variable 'Occprotect.' Results showed that occupant protection usage has reduced injury severity. The nondeployment of airbags at the time of a crash increased injury severity of the crash since the variable 'Airbag_1'has a positive estimated coefficient.

When driver-contributing circumstances were analyzed, the variable 'DrivCC_2' showed a positive estimated coefficient. This indicates when the drivers are following too close to each other; there is a tendency towards having high injury severity. A careful observation of estimates gives more specific details about how far this affects injury severity.

***Roadway Related***

According to the model estimates, work zone crashes occurring on roadways (Onrdway) have a tendency towards high severe injuries, whereas intersection-related work zone crashes have an opposite effect on injury severity.

High injury severities on roadway crashes could be due to higher speed limits and lack of facilities available on the roadside such as guard rails, shoulder lanes, lighting, etc. However, at intersections, speeds are a little lower with better facilities, due to which the chances are lower for such type of crashes. Speed is one of the most important parameters capable of generating different levels of injury severity. Speed limit variable "Spdlimit" was included in the model specification to evaluate its effect on injury severity of work zone crashes. Results indicated speed has a proportional relationship with injury severity by which if speed increases injury severity increases.

***Crash Related***

Among different types of vehicles involved in work zone crashes, the variable trucks (Truck) and light-duty vehicles (Ligdtyveh) such as pickup trucks, vans, and SUV's indicate statistically significant influence towards injury severity in work zone crashes. This implies when trucks and light duty vehicles are involved in work zone crashes, injury severity of those crashes is expected to be high. Trucks had a higher positive estimated parameter than light-duty vehicles which indicates a higher probability of a high severity crash if a truck is involved in a crash than light-duty vehicles. This might be due to the fact that trucks occupy more space in work zones, leading to multiple-vehicle collisions which end in high injury severity.

When the vehicle is taking a left turn or right turn before the crash, the resulting crash leads to increased injury severity, as the variable 'Critmaneu' has a positive estimated parameter. However, when the vehicle is passing another vehicle before the crash, the probability of injury severity is less, as the variable "passing" showed a negative estimated parameter.

In case of multiple-vehicle collisions, sideswipe collision (Sideswipe_same) in the same direction results in more severe injuries to vehicle occupants than head-on collisions. This might be because in work zones, reduced traffic lane widths will increase the interaction between the vehicles travelling in the same direction, which tends to result in more sideswipe collisions. Reduced injury severity in the case of head-on collisions might be because work zones were present in urban areas where there are low speed limits. Similarly, the variables "overturn" and "collision with fixed object" showed a decreasing injury severity, as the usage of seat belts and deployment of airbags might have reduced injury severity.

### *Environment Related*

The variable related to weather conditions (Clearweacond) had a positive estimated parameter. This shows that, when a crash occurs in clear weather conditions, severity of the crash could be expected to be more, compared to crashes that occur in adverse weather conditions. It doesn't show that all work zone crashes occurring under clear weather conditions are more severe. This variable can be better explained once details such as number of vehicles passing through work zones in daytime and nighttime, length of work zone, active and idle times of work zones etc. are known. This was not possible in this study due to limitations in the electronic data set.

## CONCLUSIONS

In order to identify risk factors associated with work zone crashes, the ordered probit model was developed for the Iowa work zone crash data set for the period 2002-2006. The objective of this type of modeling was to see the combined effect of variables contributing to higher injury severity.

Based on the study, work zone crashes involving trucks, light-duty vehicles following too close, non-deployment of airbags, sideswipe collision of same-direction vehicles, crashes occurring on roadways, posted speed limits and crashes occurring while vehicles were taking left/right turns in a work zone area showed a higher propensity for severe injuries. Work zone crashes involving male drivers had a tendency for higher injury severities compared to female drivers. Middle-age drivers were more prone to severe injuries than old age and young age drivers. Injury severity was high in crashes occurring on on-roadway work zone areas. Vehicles colliding sideways while travelling in the same direction showed significant results with respect to higher injury severity when compared to head-on collisions. Compared to other vehicle types, involvement of trucks in work zone crashes tended to have high injury severity. Further, it was found that vehicles following too close in work zone areas tended to increase the injury severity of the occupants. Finally, it can be concluded the study has found many important parameters where occupants are at

risk in work zone areas, and these findings can be used in the future to improve safety in work zones.

## REFERENCES

Federal Highway Administration (FHWA) (2005), U.S. Department of Transportation, http//www.fhwa.dot.gov/index.html. Accessed March.14, 2010.

Kockelman, K.M., and Y. Kweon (2002). "Driver Injury Severity: An Application of Ordered Probit Models."*Accident Analysis and Prevention*, Vol.34, pp. 313-321.

Ma, J., and K.M. Kockelman (2004). "Anticipating Injury and Death: Controlling for New Variables on Southern California Highways." Presented at 83rd Annual Meeting of the Transportation Research Board, Washington D.C.

Khattak, A.J., J.R. Schneider, and F. Targa (2003). "Risk Factors in Large Truck Rollovers and Injury Severity: Analysis of Single-Vehicle Collisions." CD-ROM. *Transportation Research Board* of the National Academies, Washington, D.C., pp. 03-2331.

Khattak, A.J., M. D. Pawlovich, R.R. Souleyrette, and S.L. Hallmark (2002). "Factors Related to More Severe Older Driver Traffic Crash Injuries." *Journal of Transportation Engineering*, Vol. 128, No. 3, pp 243-249.

Ratnayake, I (2004). "Identification of Factors Related to Urban and Rural Highway Crashes." Submitted for presentation at Fall Student Conference, Midwest Transportation Consortium.

O'Donnell, C.J., and D.H. Connor (1996). "Predicting the Severity of Motor Vehicle Accident Injuries Using Models of Ordered Multiple Choice." *Accident Analysis and Prevention*, Vol.28, pp. 739-753.

Long, J.S. (1997). "Regression Models for Categorical and Limited Dependent Variables." SAGE Publications, Inc., California.

Perera, H. L. K. *An Analysis of Older-Driver Involvement in Crashes and Injury Severity In Kansas.* Graduate Thesis, Kansas State University, Manhattan, 2009.

SAS Online Doc 9.1.3, SAS Institute Inc., Cary, NC, 2008.

Greene, W.H. (1997). Econometric Analysis, 3rd Edition, Prentice Hall, Upper Saddle River, N.J.

# Crash Estimation Models for Intersections

**Srinivas S. Pulugurtha[1] and Anusha P. Nujjetty[2]**

## ABSTRACT

The focus of this paper is to develop models to estimate the number of crashes at intersections. Data collected at 150 intersections in the City of Charlotte, North Carolina were used in the development of models. Factors such as demographic (population and households), socio-economic (income and employment), network (number of lanes, speed limit, signalized or not, skewed, number of left-turn and right-turn lanes and traffic volume by turning movement) and land use (commercial and residential) characteristics at each intersection were considered as independent variables. The numbers of crashes at each intersection was used as a dependent variable. Generalized linear models were developed to estimate risk at intersections. The role of spatial proximity in capturing data to develop models was also examined. These models provide valuable insights on characteristics that contribute to crashes at intersection. They could be used by practitioners to estimate potential risk at new intersections and at intersections near new developments so as to proactively apply appropriate treatments.

## INTRODUCTION

Traffic collisions are a major cause of mortality in the United States (Mokdad et al., 2004) and the leading cause of death for persons aged 4 - 34 (Subramanian, 2006). Intersections are often recognized as the most hazardous locations than any other locations on the roads. Every year more than 2,500 crashes are recorded at 150 high crash intersections in the Charlotte metropolitan area. Speeding, failure to yield right-of-way, inattention, distractions, and driving under the influence of alcohol and drugs are some of the major contributing factors of crashes at intersections. Rear end collisions, angle collisions, sideswipe collisions, right-turn collisions and left-turn collisions are the most common types of crashes at intersections. These contributing

[1] Associate Professor of Civil & Environmental Engineering, Assistant Director of Center for Transportation Policy Studies, The University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001, USA; Phone: (704) 687-6660, Fax: (704) 687-6953, E-mail: sspulugurtha@uncc.edu

[2] Graduate Student of Civil & Environmental Engineering, The University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001, USA; Phone: (704) 687-2769, Fax: (704) 687-6953, E-mail: anujjett@uncc.edu

factors and collision types account for more than 70 percent of crashes at intersections in the Charlotte metropolitan area. It is anecdotally believed that characteristics such as demographic, socio economic, network and land use play a vital role in explaining the causes of crashes and developing models to estimate the number of crashes.

The prime focus of this paper is to develop models to estimate the number of crashes at intersections. The developed models could be used by practitioners to estimate potential risk at new intersections and at intersections near new developments so as to proactively apply appropriate treatments.

## LITERATURE REVIEW

Persaud et al. (2002) developed crash estimation models to explain the complexity of calibrating these models for intersections. The models were developed using data from Toronto, Canada for 3- / 4 legged signalized and unsignalized intersections and compared to Vancouver and California. These models were developed to evaluate the relationship between Average Annual Daily Traffic (AADT) and number of crashes.

Wood (2002) described underlying mechanism of generalized linear crash models and practical resolution of the 'low mean value' problem. Chin et al. (2003) developed a Random Effect Negative Binomial (RENB) model to investigate the relationship between crash occurrence and geometric, traffic and control characteristics at signalized intersections in Singapore.

Greibe (2003) researched on crash estimation models in order to develop models for urban junctions and links using generalized linear modeling technique. Crash frequencies were related to road design and road geometry variables. They found that while road environment variables, minor side roads, parking facilities and speed limits are significant for road links, vehicle traffic flow is significant for road junctions.

Oh et al. (2004) developed crash estimation models for rural highway intersections using Poisson and negative binomial distribution techniques. This study mainly considered signalized highway intersections and multilane stop-controlled highway intersections. Data from the States of Georgia, California and Michigan were collected to use in their crash estimation models.

Wood (2005) explained errors that were mainly related to crash rates and variables using generalized linear models with logarithmic function. Lee and Abdel-Aty (2005) explained how they used log linear and ordered-probit models to correlate pedestrian crashes with pedestrian and driver demographic characteristics, traffic, and environmental characteristics at intersections. The study found that the above characteristics are closely related to the frequency and severity of pedestrian crashes. They recognized that intoxicated drivers and pedestrians were correlated to nighttime crashes, and the absence of traffic signals in rural areas was correlated to more crashes.

Mitra and Washington (2007) evaluated the nature of over-dispersion in motor vehicle crash estimation models. This study was motivated to corroborate the findings of Miaou and Lord (2003) regarding the variance structure in over-dispersed crash models. Ma et al. (2008) researched creating multivariate Poisson lognormal regression (MVPLN) models using Bayesian statistics, the Gibbs sampler and the

Metropolis-Hastings (M-H) algorithms for different severity levels. Crash data was collected from Washington State through the Highway Safety Information System (HSIS) for roadway segments. They also evaluated the correlation structure and over-dispersion. Crash data was used as a dependent variable while roadway geometric characteristics were used as independent variables.

Haleem et al. (2009) used a reliability process to reduce uncertainty in predicting crashes at intersections without a traffic signal. The authors examined the effect of updating the parameters of covariates in the fitted negative binomial model using Bayesian updating reliability method to predict the crashes more accurately. Roadway geometric, traffic and control fields data were used in this study.

### Limitations of Past Research

Demographic, socio-economic, network and land use characteristics were considered individually to estimate the number of crashes at intersections in the past. A few researchers have considered a combination of two characteristics. However, not all characteristics were considered in developing estimation models in the past.

Researchers considered variables within the proximity of an intersection. However, the effect of spatial proximity in capturing data and results obtained was not examined. Crashes at intersections may involve characteristics within 0.25-mile, 0.5-mile, 1-mile and 2-mile buffers. Data within these buffers need to be extracted, examined and used in recommending the models that provide better and statistically meaningful estimates.

The research presented in this paper attempts to address the above limitations.

## METHODOLOGY

The GIS based methodology to estimate the number of crashes at intersections comprises the following steps.

i. Selection of study intersections
ii. Data collection
iii. Using Geographic Information Systems (GIS) to extract data
iv. Development of correlation matrices and models to estimate the number of crashes

### Selection of Study Intersections

Intersections should be geographically distributed all over the study area to minimize any spatial bias. Study intersections selected should contain a well defined data set to explain the crash occurrence in the study area. In this research, 150 randomly selected intersections in the City of Charlotte, North Carolina were used to collect data and develop crash estimation models. The selected intersections comprise 3- / 4-legged, signalized / unsignalized, and skewed intersections.

## Data Collection

Crash data were collected from the City of Charlotte Department of Transportation (CDOT) for years 2006, 2007 and 2008. The average number of crashes was calculated and used in the development of models.

Traffic volume data collected by CDOT from 7:00 AM to 7:00 PM (12-hour counts) for each study intersection was obtained and used as an independent variable. Census block data was used to extract the demographic and socio-economic characteristics within the vicinity of each study intersection. Demographic characteristics used include the total numbers of population and total numbers of households. Socio-economic characteristics considered are average household income and the total number of civilians employed within the vicinity of each study intersection.

The network characteristics considered in the development of crash estimation models include the following.

- Intersection control type (signalized or unsignalized)
- Skewed or not skewed
- Total number of legs (or approaches)
- Number of left-turn lanes on major street approaches
- Number of through lanes on major street approaches
- Number of right-turn lanes on major street approaches
- Speed limit on major street
- Number of left-turn lanes on minor street approaches
- Number of through lanes on minor street approaches
- Number of right-turn lanes on minor street approaches
- Speed limit on minor street

Land use characteristics may also have an effect on the number of crashes. Those considered in this paper include residential and commercial land use categories.

## Using Geographic Information System (GIS) to Extract Data

Spatial features in Geographic Information Systems (GIS) were used to capture demographic, socio-economic and land use characteristics within the vicinity of each study intersections. Buffers of width 0.25 miles, 0.5 miles, 1 mile and 2 miles were generated around each study intersection to examine the effect of spatial proximity in capturing data for development of crash estimation models.

The demographic, socio-economic characteristics and land use data are overlaid on the generated buffers to extract data pertaining to each characteristic identified in the previous step. The attribute table of the intersected demographic / socio-economic data was adjusted based on the proportion of each census block in the generated buffer. Data were summed if multiple census blocks fell in the generated buffer.

### Development of Correlation Matrices and Models to Estimate the Number of Crashes

The focus of this step was to eliminate variables that are correlated to each other. Variables with correlation -0.3 to 0.3 were only selected for the development of models. Statistical analysis software SPSS® (SPSS Inc., 2005) was used to develop the models.

The Poisson log linear relationship holds well when the mean is equal to the variance. If this assumption is not valid, the standard errors, usually estimated by the maximum likelihood (ML) method will be biased and test statistics derived from the model will be incorrect (Chin et al., 2003). In order to overcome the problem of over-dispersion, generalized linear models based on negative binomial distribution (negative binomial log-link) were used to develop the models. Corrected Quasi Likelihood under Independence Model Criterion (QICC) statistic was used to select the best model. In general, the crash estimation model with lower QICC value was selected as the best model.

## ANALYSIS & RESULTS

Data for 150 randomly selected intersections in the City of Charlotte, North Carolina were used to develop crash estimation models. Demographic, socio-economic and land use data within 0.25-mile, 0.5-mile, 1-mile, and 2-mile buffer widths were extracted to examine the effect of spatial proximity.

Correlation coefficients were computed for data based on each buffer width and network characteristics. These were used to identify independent variables that are not related to each other. As an example, the correlation matrix for 0.25-mile buffer width data is shown in Table 1.

Models were initially developed using all variables with correlation coefficients between -0.3 to 0.3. Independent variable with the highest significance value was identified and eliminated. The model was re-run (developed) using the new list of independent variables. This process was repeated until all the independent variables with significance value greater than 0.05 (95% confidence level) were eliminated. Hence, the final model had independent variables with significance values less than or equal to 0.05.

Table 2 summarizes the final models developed for different buffer widths using SPSS® statistical analysis software. B and p in the table are coefficient and significance value for the corresponding variable. The results shown in the table can used to estimate the number of crashes. As an example, the number of crashes based on crash estimation model for 0.25-mile buffer width can be estimated using the following equation (based on results shown in Table 2).

$$\text{Ln (Crashes)} = 2.39 + 0.000012\ \text{Traffic volume} - 0.000009\ \text{Income} + 0.08\ \text{Skew} + 0.02\ \#\ \text{Minor approach right-turn lanes} - 6.86\ \text{Residential 3} + 3.46\ \text{Institutional}$$

In general, it was observed that crashes increase with traffic volume but decrease with population and household income within the vicinity of an intersection. Skewed

intersections and number of turn lanes generally tend to increase the number of crashes at an intersection.

It can be observed that the computed QICC is lowest for model based on 0.25-mile buffer width. This indicates that crash estimation models based on 0.25-mile buffer width data would yield the best results (estimate).

## CONCLUSIONS

This paper presents a methodology and crash estimation models to estimate the number of crashes at intersections. Data were extracted using 0.25-mile, 0.5-mile, 1-mile and 2-mile buffer widths to study the role of spatial proximity in capturing demographic, socio-economic and land use data in developing models. It was observed from the goodness of fit tests that using 0.25-mile buffer width would yield statistically better results than compared to other buffer widths used in this paper.

Crashes can be estimated using predictor variables such as traffic volume, average household income, whether the intersection is skewed or not, number of minor approach right-turn lanes, 0.5 - 2 acres residential area and institutional area. The results obtained from this research could be pro-actively used to estimate crashes and enhance safety at new intersections and at intersections near new developments.

Crash severity was not considered in this research. Identifying critical variables and developing models based on crash severity needs investigation. Variables such as gender, race, human machine interactions, and driver behaviors were not considered. These need to be considered to study their role and develop better crash estimation models.

## REFERENCES

1. Chin, H. C., and Quddus, M. A. (2003). "Applying the Random Effect Binomial Model to Examine Traffic Accident Occurrence at Signalized Intersections." *Accident Analysis & Prevention*, Vol. 35(2), 253-259.
2. Greibe, P. (2003). "Accident Prediction Models for Urban Roads." *Accident Analysis & Prevention*, Vol. 35(2), 173-185.
3. Haleem, K., Abdel-Aty, and M., Mackie, K. (2009). "Using a Reliability Process to Reduce Uncertainty in Predicting Crashes at Unsignalized Intersections." *Accident Analysis & Prevention*, Vol. 42(2), 654-666.
4. Lee, C., and Abdel-Aty, M. (2005). "Comprehensive Analysis of Vehicle-Pedestrian Crashes at Intersections in Florida." *Accident Analysis & Prevention*, Vol. 37(4), 775-786.
5. Ma, J., Kockelman, K. M., and Damien, P. (2008). "A Multivariate Poisson-Lognormal Regression Model for Prediction of Crash Counts by Severity." Using Bayesian Methods, *Accident Analysis & Prevention*, Vol. 40(3), 964-975.
6. Miaou, S. P., and Lord, D. (2003). "Modeling Traffic Crash Flow Relationships for Intersections: Dispersion Parameter, Functional Form, and Bayes versus Empirical Bayes Methods." Transportation Research Record # 1840, *Journal of Transportation Research Board*, 31-40.

7. Mitra, S., and Washington, S. (2007). "On the Nature of Over-dispersion in Motor Vehicle Crash Prediction Models." *Accident Analysis & Prevention*, Vol. 39(3), 459-468.
8. Mokdad, A. H., Marks, J. S., Stroup, D F., and Gerberding, J. L. (2004). "Actual causes of Death in the United States- 2000." *Journal of American Association*, Vol. 291(10), 1238-1245.
9. Oh, J., Lyon, C., Washington, S., Persaud, B., and Bared, J. (2003). "Validation of FHWA Crash Models for Rural Intersections: Lessons Learned." Transportation Research Record # 1840, *Journal of Transportation Research Board*, 41-49.
10. Persuad, B., Lord, D., and Palmisano, J. (2002). "Calibration and Transferability of Accident Prediction Models for Urban Intersections." Transportation Research Record # 1784, *Journal of Transportation Research Board*, 57-64.
11. SPSS Inc. (2005). SPSS® 14.0 *Brief Guide*, Copyright © 2005 by SPSS Inc., 233 South Wacker Drive, 11th Floor, Chicago, IL.
12. Subramanian, R. (2006). "Motor vehicle Traffic Crashes as a Leading Cause of Death in the United States, 2003." NHTSA Traffic Safety Facts Research Note.
13. Wood, G. R. (2002). "Generalized Linear Accident Models and Goodness of Fit Testing." *Accident Analysis & Prevention*, Vol. 34(4), 417-427.
14. Wood, G. R. (2005). "Confidence and Prediction Intervals for Generalized Linear Accident Models." *Accident Analysis & Prevention*, Vol. 37(2), 267-273.

## ACKNOWLEDGMENTS

The authors acknowledge and sincerely thank Charlie Jones, Ashton Watson, Steven Castongia, and Anna Gallup of the City of Charlotte Department of Transportation (CDOT) for their help with crash data, land use data, and demographics / socio-economic characteristics data.

## DISCLAIMER

The contents of this paper reflect the views of the author(s) and not necessarily the views of the University or the City of Charlotte Department of Transportation (CDOT). The author(s) are responsible for the facts and the accuracy of the data presented herein.

**TABLE 1 Correlation Matrix for 0.25-mile Buffer Width**

| Variables | TV | Pop | HH | Inc | Empl | Signal | Skew | Nofl | Ma_S | Ma_L | Ma_T | Ma_R | Mi_S | Mi_L | Mi_T | Mi_R | R1 | R2 | R3 | R4 | LC | LI | HC | HI | Inst |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TV | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| Pop | -0.22 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| HH | -0.26 | 0.88 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| Inc | 0.04 | 0.15 | 0.04 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| Empl | -0.21 | 0.96 | 0.93 | 0.16 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| Signal | 0.10 | 0.01 | 0.05 | -0.14 | -0.01 | 1.00 | | | | | | | | | | | | | | | | | | | |
| Skew | 0.13 | -0.01 | -0.11 | -0.03 | -0.03 | -0.06 | 1.00 | | | | | | | | | | | | | | | | | | |
| Nofl | 0.10 | 0.18 | 0.18 | -0.06 | 0.16 | 0.41 | 0.03 | 1.00 | | | | | | | | | | | | | | | | | |
| Ma_S | 0.27 | -0.17 | -0.33 | -0.01 | -0.18 | -0.13 | 0.26 | -0.15 | 1.00 | | | | | | | | | | | | | | | | |
| Ma_L | 0.40 | -0.24 | -0.32 | -0.03 | -0.23 | 0.28 | 0.13 | 0.16 | 0.26 | 1.00 | | | | | | | | | | | | | | | |
| Ma_T | 0.35 | -0.12 | -0.16 | 0.13 | -0.13 | 0.17 | -0.09 | 0.09 | 0.00 | 0.31 | 1.00 | | | | | | | | | | | | | | |
| Ma_R | 0.34 | -0.22 | -0.29 | 0.05 | -0.22 | 0.17 | 0.24 | 0.23 | 0.22 | 0.56 | 0.15 | 1.00 | | | | | | | | | | | | | |
| Mi_S | 0.16 | -0.22 | -0.31 | -0.14 | -0.23 | -0.01 | 0.26 | -0.06 | 0.53 | 0.26 | 0.00 | 0.27 | 1.00 | | | | | | | | | | | | |
| Mi_L | 0.40 | -0.22 | -0.26 | -0.02 | -0.23 | 0.28 | 0.13 | 0.19 | 0.11 | 0.64 | 0.33 | 0.60 | 0.22 | 1.00 | | | | | | | | | | | |
| Mi_T | 0.15 | 0.07 | 0.08 | -0.03 | 0.06 | 0.29 | -0.02 | 0.46 | -0.21 | 0.29 | 0.17 | 0.17 | -0.10 | 0.20 | 1.00 | | | | | | | | | | |
| Mi_R | 0.09 | 0.03 | 0.02 | -0.03 | 0.07 | -0.17 | -0.06 | -0.14 | 0.13 | 0.01 | 0.09 | 0.03 | 0.05 | 0.08 | -0.08 | 1.00 | | | | | | | | | |
| R1 | -0.03 | 0.41 | 0.33 | 0.17 | 0.43 | 0.05 | -0.12 | 0.21 | -0.12 | -0.13 | 0.01 | -0.06 | -0.14 | 0.02 | 0.10 | 0.01 | 1.00 | | | | | | | | |
| R2 | 0.15 | 0.18 | -0.01 | 0.50 | 0.14 | -0.12 | -0.02 | -0.06 | -0.03 | 0.08 | 0.18 | 0.05 | -0.06 | 0.07 | -0.04 | -0.03 | 0.03 | 1.00 | | | | | | | |
| R3 | 0.01 | -0.25 | -0.27 | -0.07 | -0.25 | -0.08 | 0.00 | -0.16 | 0.12 | 0.05 | -0.06 | -0.01 | 0.07 | 0.04 | -0.08 | 0.03 | -0.14 | -0.01 | 1.00 | | | | | | |
| R4 | -0.12 | -0.21 | -0.27 | -0.11 | -0.23 | -0.12 | 0.26 | -0.12 | 0.23 | 0.10 | -0.20 | 0.19 | 0.34 | -0.05 | -0.18 | -0.03 | -0.24 | -0.25 | -0.13 | 1.00 | | | | | |
| LC | 0.17 | 0.02 | 0.00 | 0.16 | 0.04 | 0.05 | -0.04 | 0.02 | 0.01 | 0.21 | 0.24 | -0.01 | -0.01 | 0.16 | 0.13 | 0.05 | 0.01 | 0.05 | -0.11 | -0.09 | 1.00 | | | | |
| LI | -0.09 | 0.13 | 0.11 | -0.01 | 0.12 | 0.14 | 0.13 | 0.06 | -0.07 | -0.06 | 0.06 | -0.13 | -0.02 | -0.08 | 0.04 | 0.03 | -0.10 | -0.06 | -0.12 | -0.03 | -0.07 | 1.00 | | | |
| HC | -0.14 | 0.25 | 0.48 | -0.30 | 0.26 | 0.13 | -0.13 | 0.15 | -0.32 | -0.32 | -0.07 | -0.15 | -0.24 | -0.17 | 0.18 | -0.03 | -0.02 | -0.31 | -0.34 | -0.30 | -0.14 | -0.12 | 1.00 | | |
| HI | 0.35 | -0.13 | -0.14 | -0.04 | -0.14 | 0.01 | -0.02 | -0.03 | 0.20 | -0.01 | 0.02 | -0.07 | 0.06 | -0.02 | -0.06 | -0.02 | -0.09 | -0.11 | -0.10 | -0.09 | 0.01 | 0.16 | -0.16 | 1.00 | |
| Inst | -0.01 | -0.10 | -0.17 | -0.03 | -0.16 | 0.05 | 0.08 | 0.12 | -0.05 | 0.12 | 0.03 | 0.07 | -0.08 | 0.07 | 0.19 | -0.04 | -0.08 | -0.06 | 0.01 | 0.02 | 0.02 | -0.08 | -0.02 | -0.07 | 1.00 |

Note: TV= Traffic volume, Pop= Population, Inc= Income, Empl= Employment, Nofl= Number of legs/approaches, Ma_S= Major approach speed limit, Ma_L= # major approach left-turn lanes, Ma_T= # major approach through lanes, Ma_R= # major approach right-turn lanes, Mi_S= Minor approach speed limit, Mi_L= # minor approach left-turn lanes, Mi_T= # minor approach through lanes, Mi_R= # minor approach right-turn lanes, R1= 0.25 acre residential, R2 = 0.25 - 0.5 acre residential, R3= 0.5 - 2 acres residential, R4= > 2 acres residential, LC= Light commercial, LI= Light industrial, HC= Heavy commercial, HI= Heavy industrial, Inst=Institutional

**TABLE 2 Summary of Models**

| Parameter | 0.25-mile buffer | | 0.50-mile buffer | | 1-mile buffer | | 2-mile buffer | |
|---|---|---|---|---|---|---|---|---|
| | B | p | B | p | B | p | B | p |
| Intercept[1,2,3] | 2.39 | 0.00 | 2.45 | 0.00 | 2.04 | 0.00 | 2.07 | 0.00 |
| Traffic volume [1,2,3] | 0.000012 | 0.00 | 0.00001 | 0.00 | 0.00001 | 0.00 | 0.00001 | 0.00 |
| Population [1,2,3,4] | | | -0.00011 | 0.00 | -0.00002 | 0.00 | 0.000006 | 0.01 |
| Household income (average) [1,2,3,4] | -0.000009 | 0.00 | -0.00001 | 0.01 | -0.00002 | 0.00 | | |
| Signalized or Unsignalized [1,2,3,4] | | | 0.27 | 0.00 | 0.23 | 0.00 | | |
| Skewed or not [1,2,3,4] | 0.08 | 0.03 | | | 0.10 | 0.01 | | |
| Major approach speed limit [1,] | | | | | | | | |
| Minor approach through lanes [1,2,3,4] | | | 0.12 | 0.00 | | | 0.1 | 0.00 |
| Minor approach right-turn lanes [1,2,3,4] | 0.02 | 0.00 | 0.02 | 0.00 | 0.03 | 0.00 | 0.02 | 0.00 |
| 0.25 - 0.5 acre Residential [4] | | | | | | | 0.002 | 0.00 |
| 0.5 - 2 acres Residential [1,4] | -6.86 | 0.00 | | | | | | |
| > 2 acres Residential/ Open Space [1,3,4] | | | | | 0.19 | 0.00 | 0.09 | 0.00 |
| Light commercial [1,2] | | | | | | | | |
| Light industrial [1] | | | | | | | | |
| Heavy industrial [2] | | | -2.86 | 0.00 | | | | |
| Institutional [1,2,3,4] | 3.46 | 0.01 | -4.91 | 0.00 | 3.51 | 0.00 | | |
| QICC | 59.58 | | 71.31 | | 71.57 | | 71.57 | |

Note: B = Coefficient, p = Significance level, [1] = Variable with correlation coefficient between -0.3 to 0.3 for 0.25-mile buffer width, [2]= Variable with correlation coefficient between -0.3 to 0.3 for 0.50-mile buffer width, [3] = Variable with correlation coefficient between -0.3 to 0.3 for 1-mile buffer width, [4] = Variable with correlation coefficient between -0.3 to 0.3 for 2-mile buffer width

# Geographically-Weighted Regression Models for Improved Predictability of Urban Intersection Vehicle Crashes

Zongzhi Li*[1], Yongdoo Lee[2], Sang Hyuk Lee[3], Eirini Valiou[4]

* Corresponding author. Tel.: (312) 567-3556; fax: (312) 567-3519. E-mail address: lizz@iit.edu.

[1]Associate Professor, Department of Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, Chicago, IL 60616

[234] Graduate Research Assistants, Department of Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, Chicago, IL 60616

## ABSTRACT

Most of current intersection vehicle crash models are calibrated using global regression analysis methods that are often inaccurate in crash predictions as some localized crash contributing effects are not explicitly addressed. This paper employs the Geographically-Weighted Regression (GWR) technique to calibrate statistical models for predicting intersection injury, property damage only (PDO), and total crashes using data on 245 intersections in City of Chicago for period 2001-2008. In the calibrated GWR models, factors contributing to intersection vehicle crashes identified include major and minor road daily traffic, number of major and minor road through and left-turn lanes, and household income level. The analysis of variance (ANOVA) test reveals that improved model predictability is achieved from all crash models developed using the GWR technique compared with those models calibrated based on the Ordinary Least Squares (OLS) technique. The Monte Carlo test identifies significance of spatial variability of explanatory variables in the GWR models.

## INTRODUCTION

Motor vehicle crashes are among the leading causes of death in the United States. In each year, more than 40,000 people lose their lives on public roads (USDOT, 2009). Higher number of vehicle crashes occurs in urban areas because of higher intensity of system usage, where approximately one-third of the total vehicle crashes occur at intersections (USDOT, 2009). Over the last several decades, highway safety engineers and researchers have been trying to identify factors causing intersection vehicle crashes and to develop statistical models for predicting intersection vehicle crashes. Such efforts serve for two purposes. First, it facilitates estimation of the likelihood of a motor vehicle crash given a set of explanatory variables associated with intersection geometric design, traffic control, and usage. Second, it helps establish the potential for safety improvements resulted from implementing safety countermeasures. The following subsections discuss intersection crash model types, model calibration techniques, and some notable studies, limitations, and motivation of current study.

### Intersection Crash Model Types

The intersection crash models can be generally classified into Poisson, negative binomial, and lognormal model types. Poisson models are a well-accepted method of modeling discrete rare events such as highway vehicle crashes. It is assumed that vehicle crashes occurring at a particular intersection are independent of one another and that average number of crashes per unit time is the characteristic of a given intersection site and of other sites with same properties. The Poisson model takes the form (Ang and Tang, 2006):

$$P(k_i) = \frac{e^{-\mu_i} . \mu_i^{k_i}}{k_i!} \quad (1)$$

where

$k_i$ = Expected number of crashes at intersection $i$,
$\mu_i$ = Average number of vehicle crashes per unit time at intersection $i$, $\mu_i = E[k_i]$,
$P(k_i)$ = Probability of having $k_i$ crashes at intersection $i$, and
$i$ = Number of intersections with filed crash observations, $i = 1, 2, \ldots, n$.

A limitation of the assumption in Poisson models is that the mean equals to the variance of the distribution. The actual field observed crashes might not always support this assumption. For the case where the variance of the observed crashes exceeds or falls below the estimated average of vehicle crashes, the crashes are then said to be overdispersed or underdispersed and the underlying assumption of variance being equal to the mean for the Poisson distribution is violated. The negative binomial distribution, which is a discrete distribution, provides an alternative model to deal with overdispersion or underdispersion circumstances. The negative binomial model is of the following specification (Ang and Tang, 2006):

$$P(k_i) = \frac{\Gamma(k_i + \frac{1}{K})}{k_i! \Gamma(\frac{1}{K})} . (\frac{K.\mu_i}{1 + K.\mu_i})^{k_i} . (\frac{1}{1+K.\mu_i})^{\frac{1}{K}} \quad (2)$$

where
$k_i$ = Expected number of crashes at intersection *i*,
$\mu_i$ = Average number of vehicle crashes per unit time at intersection *i*, $\mu_i = E[k_i]$,
$P(k_i)$ = Probability of having $k_i$ crashes at intersection *i*,
*K* = Overdispersion parameter, and
*i* = Number of intersections with field crash observations, $i = 1, 2, \ldots, n$.

The mean and variance of negative binomial distribution are $\mu_i$ and $\mu_i + K.\mu_i^2$. If K equals to 0, the negative binomial model reduces to the Poisson model.

The third model type is the lognormal regression models assuming that the natural logarithm of vehicle crashes $k_i$ follows a normal distribution with mean $\mu_i$, i.e., $E[k_i]$, and variance $V[k_i]$. This model is a reasonable choice whenever the field crash data are inherently non-negative and maintain large mean value as well as positive skewness. The log-normal distribution has density function (Ang and Tang, 2006):

$$f(k|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma k} e^{-(\ln k-\mu)^2 / 2\sigma^2} \quad (3)$$

where
k = Expected number of crashes,
$\mu$ = Average number of vehicle crashes per unit time, $\mu = E[k]$, and
$\sigma$ = Standard deviation of the variable *k*'s logarithm, $\sigma^2 = V[k]$.

For $k > 0$, mean and variance of the variable *k*'s logarithm are $E[k]=e^{\mu+\sigma^2/2}$ and $V[k]=( e^{\sigma^2}-1) e^{2\mu+\sigma^2}$.

**Current Model Calibration Techniques**

For either Poisson, negative binomial or lognormal model type, the average number of vehicle crashes per unit time at intersection *i* is correlated with intersection design and usage variables, $X_1$, $X_2$, ..., $Xq$, of a generalized linear form given by

$$\ln(\mu_i) = \beta_0 + \beta_1.X_{1i} + \beta_2.X_{2i} + \ldots + + \beta_p.X_{pi} \quad (4)$$

where
$\mu_i$ = Average number of vehicle crashes per unit time at intersection *i*, $\mu_i = E[k_i]$,
$X_{pi}$ = The $p^{th}$ highway variables that contributes to crashes at intersection *i*,
$\beta_p$ = Linear regression coefficient of variable $X_{pi}$ to be estimated by model calibration, and
*i* = Number of intersections with field crash observations, $i = 1, 2, \ldots, n$.

Modeling techniques used in calibrating the coefficients in Equation (4) usually include the Ordinary Least Squares (OLS) technique where parameter coefficients are estimated by minimizing the sum of squared residuals; and the maximum likelihood estimation technique where parameter coefficients are determined by maximizing the log-likelihood function. The maximum likelihood technique is broadly applied to

calibrate the best linear unbiased estimator (BLUE) with smallest variance (Wooldridge, 2003).

**Some Related Studies, Limitations, and Motivation of Current Study**

Due to differences in intersection geometric design and traffic control, vehicle crash models are separately calibrated for 3-leg stop controlled, 4-leg stop controlled, and 4-leg signalized intersections. Because of relaxing the constraint of equality between mean and variance of crash frequencies, the negative binomial model has been used more often than the Poisson model. Since the negative binomial model does not require to have a relatively large value and positive skewness of crash counts, it is also more frequently used than the loglinear model. The most comprehensive negative binomial regression models are found in the FHWA's Interactive Highway Safety Design Manual (Bauer and Harwood, 1996) and the Highway Safety Manual (AASHTO, 2010). In addition, negative binomial models with fewer highway variables were developed in a number of states, including California, Iowa, Illinois, Louisiana, Minnesota, North Carolina, Nebraska, Oregon, and Virginia (Bauer and Harwood, 1999; and Harwood et al., 2002). Relatively fewer models were calibrated using lognormal distributions. One example is the lognormal regressions models developed in California (Bauer and Harwood, 1999). In the existing intersection crash models, statistically significant highway variables pertaining to geometric design and traffic control elements and intensity of intersection usage include the following: sight distance, horizontal and vertical alignments, intersection angle, lane width, speed, traffic volume, traffic channelization, and safety hardware characteristics.

In viewing the existing models, both time series or cross-section data were used for model calibration. However, very little attention has been given to the spatial distribution of the data. A rational approach would be to assign data weights according to its proximity at the point of individual intersections and to calibrate a model for each intersection point in space of interest. Such a model would allow for spatially varying model coefficients. That is to say regardless of Poisson, negative binomial or lognormal assumptions, model coefficients should capture variability of localized intersection geometric design, traffic control, and usage that contribute to vehicle crashes to improve model prediction power.

In this respect, the geographically-weighted regression (GWR) technique offers an alternative model calibration technique that permits the model coefficient estimates to vary locally rather than globally (Fotheringham et al., 2002), which is well suited for predicting intersection vehicle crashes. The GWR model has the general form

$$\ln(\mu_i) = \beta_0(x, y) + \beta_1(x, y).X_{1i} + \beta_2(x, y).X_{2i} + ... + +\beta_p(x, y).X_{pi} \quad (5)$$

where

$\mu_i$ = Mean number of vehicle crashes per unit time at intersection $i$, $\mu_i = E[k_i]$;
$X_{pi}$ = The $p^{th}$ highway variables that contributes to crashes at intersection $i$;
$\beta_p(x, y)$= Spatially varied regression coefficient of variable $X_{pi}$; and
$i$ = Number of intersections with field crash observations, $i = 1, 2, \ldots, n$.

The main advantage of a GWR model is that the model coefficient estimates vary spatially. This is a very essential trait of a model as it adds uniqueness in spatial

location of an event, such as vehicle crashes. It allows for a spatially varying parameter representative of the particular intersection in space. Another important aspect of the GWR model is that it also allows the set of spatially varying parameter to be used along with spatially stationary parameter, as mixed GWR model. The GWR technique is just at its early stage in application to highway safety analysis. The only research found in existing vehicle crash model development literature was a study in estimating vehicle crashes in Toronto metropolitan area (Hadayeghi et al., 2003).

This study aims to calibrate intersection injury, property damage only (PDO), and total crash models by applying the GWR technique to multiyear data on intersections in City of Chicago. Statistical tests are performed systematically to demonstrate the superiority of models calibrated using the GWR technique over those models developed utilizing the OLS technique. The paper is organized as follows: the next section introduces basic analytical steps of the proposed study methodology. The subsequent section discusses methodology application. The last section summarizes and concludes the study.

## PROPOSED METHODOLOGY

### Study Design

*Selection of Response Variables*. Vehicle crash frequency and vehicle crash rate are two popular measures for intersection safety analysis. Vehicle crash frequency is generally defined as number of vehicle crashes of specified type in a specified period, such as number of fatal crashes per year. Vehicle crash rate is defined as number of vehicle crashes of specified type for unit traffic exposure in a specified period, such as number of intersection fatal crashes per million entering vehicles per year. The curvilinear function of vehicle crashes per year to traffic flow reflects the fact that drivers behave differently in sparse and heavy traffic and crash frequency depends on traffic flow, speed, and density. This reveals that vehicle crash rate at a higher traffic volume becomes lower, but we cannot claim the intersection is safer. This is because crash frequency at higher traffic volume is actually higher, which is in deteriorated safety condition (Hauer, 1999). Hence, crash frequency in terms of number of total vehicle crashes per year is a valid variable to model intersection safety. The response variables for crash model development in current study are injury, PDO, and total crashes per year, respectively.

*Selection of Explanatory Variables*. A number of variables have been found to provide explanations for intersection vehicle crashes. The primary importance of traffic as an explanatory variable for intersection crashes relative to other highway variables has long been acknowledged, as found in existing crash prediction models. Design aimed at improving intersections safety is always in demand. Design variables found to be statistically significant in connection to intersection crashes include left-turn and right-turn traffic channelization, sight distance, horizontal and vertical alignments, intersection angle, median, shoulder type and width, roadside condition, and safety hardware conditions. Other variables noted involve truck percentage in the traffic stream, speed, and weather condition. Table 1 lists response variables and key explanatory variables considered in current study.

Table 1. Detailed Data Items Considered for Intersection Crash Model Calibration

| Data Category | Data Item | Unit |
|---|---|---|
| Vehicle Crashes | Longitude coordinate | Degree |
| | Latitude coordinate | Degree |
| | Total crashes per year | Number |
| | Fatal crashes per year | Number |
| | Class A injury crashes per year | Number |
| | Class B injury crashes per year | Number |
| | Class C injury crashes per year | Number |
| | Property Damage Only (PDO) crashes per year | Number |
| Intersection Geometric Design | Number of through lanes for each approach | Number |
| | Number of left-turn lanes for each approach | Number |
| | Number of right-turn lanes for each approach | Number |
| | Through lane width for each approach | ft |
| | Left-turn lane width for each approach | ft |
| | Right-turn lane width for each approach | ft |
| | Pavement condition (dry/wet/snow/ice/mud/other) | Dummy variable |
| Traffic Control | Signal type (pretimed/semi actuated/actuated) | Dummy variable |
| | Signal cycle (green time/yellow time/red time) | Dummy variable |
| Intersection Usage | Major road AADT | Vehicles/day |
| | Minor road AADT | Vehicles/day |
| | AM and PM peak-hour factor | Number |
| | Truck percentages on major and minor roads | % |
| Human Factors | Gender of driver (male/female) | Dummy variable |
| | Age of driver (young/middle/old) | Dummy variable |
| | Physical condition of driver (normal/impair/ill/ asleep/other) | Dummy variable |
| | Purpose of trip (personal/emergency/bus/ commercial/other) | Dummy variable |
| Climate Features | Weather condition (clear/rainy/snowy/foggy/other) | Dummy variable |
| Demographic and Socioeconomic | Population per census tract | Persons/tract |
| | Income per capita | $/person |

*Consideration of Model Types*. A model is the representation of a system in which a set of explanatory variables come together to provide a certain response. In the case of statistical experiments, a model is a mathematical relationship that describes the value of an observation in terms of a set of factors and an error component over which there is no control. A statistical model, if correctly specified, is capable of explaining random observation responses for all experimental conditions. The three important considerations for good models are a good set of explanatory variables, an appropriate response variable, and a model type that adequately explains the relationship between the selected explanatory variables and the response variable. Explanatory variables and response variables have already been discussed in previous sections. As summarized in the literature review, vehicle crash prediction models were developed using loglinear (Poisson and negative binomial) and lognormal models. These models will form the basis in the choice of model types in current study. Figure 1 illustrates the procedure for model calibration.

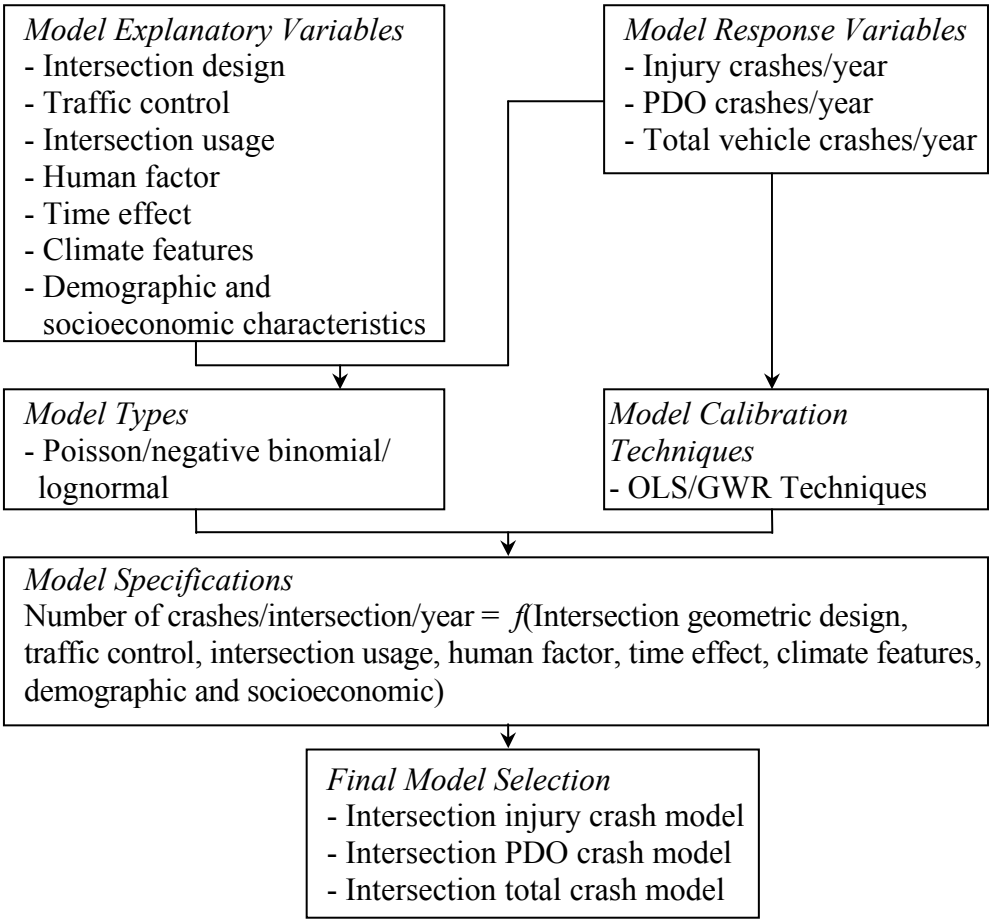



Figure 1. Procedure for calibrating intersection vehicle crash models.

**Data Collection and Processing**

*Intersection Crash Data*. The data collection effort is focused on City of Chicago area. Detailed data items listed in Table 1 are collected from the Chicago Department of Transportation (CDOT) for 245 signalized intersections for period 2001-2008. Not all intersections maintain 8-year data for the analysis period. This provides an unbalanced panel dataset. In total, there are 632 data points in the dataset. The data were entered into a database using MS Excel and a unique ID number was assigned to each intersection.

Table 2 presents summary statistics of injury, PDO, and total crashes in the dataset. As shown in the table, injury, PDO, and total crashes follow an overdispersion trend. This potentially eliminates choice of Poisson models for model calibration that leaves considering the candidacy of negative binomial or lognormal models.

Table 2. Summary Statistics of Injury, PDO, and Total Crashes in the Dataset

| Crashes | Mean | Variance | Minimum | Maximum |
|---|---|---|---|---|
| Injury | 5 | 7 | 0 | 11 |
| PDO | 25 | 36 | 9 | 40 |
| Total | 29 | 39 | 13 | 45 |

*Daily Traffic Data*. Annual average daily traffic (AADT) on each intersection approach was obtained from CDOT.

*Truck Percentage Data*. The truck percentage data were obtained from CDOT's traffic count sheets.

*Intersection Signal Timing Data*. All intersections are signalized and the relevant data were acquired from CDOT.

*Socioeconomic Data*. The socioeconomic data were obtained from U.S. Census Bureau. The Bureau maintains a database at census tract level. Income per capita and population by census tract in Cook County, covering City of Chicago area, were downloaded from the Bureau's website and was loaded to MS Excel spreadsheet. The data were for periods 1990 and 2000, respectively. The income and population data for each census tract were characterized by a unique ID number, STFID.

*Topologically Integrated Geographic Encoding and Referencing Files*. Topologically Integrated Geographic Encoding and Referencing (TIGER) files are used by reprocessing software such as ArcView, ArcGIS and Accuglobe to create a geo-referenced line or polygon files known as the shape files. These shape files can be projected in mapping software. TIGER shape files for the census tract, polygon shape file for City of Chicago and line shape file for the road network in Chicago region were downloaded from the geo-processed Environmental Systems Research Institute (ESRI) database.

*Intersection Geometry and Geospatial Coordinates*. Detailed addresses for individual intersections were first located using Google maps. The latitude and longitude coordinates of these intersections were then obtained running info functions on the Microsoft Terraserver web site. The info showed a grid of coordinates. The aerial imagery was zoomed in to the best resolution. The latitude and longitude coordinates of the intersections were then acquired from the image and entered into the vehicle crash database. The number of through-movement lanes and the number and presence of left-turn or right-turn exclusive lanes were also examined from the aerial image and then entered into the database. For each intersection, the aerial image was compared with that obtained from Google map to ensure consistency.

**Model Calibration Procedure**

The GWR technique uses a weighting function that deals with spatial position of an intersection in terms of its latitude and longitude coordinates to derive model coefficients. A higher weight is assigned to the proximal intersections, while the weight diminishes according to increase in the distance between intersections. Two types of statistical tests were performed for model selection: i) analyses of variance (ANOVA) to determine improved predictability of GWR models over conventional global regression models; and ii) Monte Carlo tests to verify localized non-stationary effect of significant explanatory variables in the GWR models. Once the non-stationary effect of any significant variable in a GWR model is rejected, the mixed GWR model was developed using the following procedure:

Step 0: Assigning Y as the response variable and splitting the significant explanatory variables into two groups. Group A contains variables with global stationary coefficients, while Group B includes variables with localized non-stationary coefficients.

Step 1: Regressing each variable in Group A against variables in Group B using GWR technique and computing residuals from the above regression.

Step 2: Regressing Y against variables in Group B using GWR technique and computing residuals from the above regression.

Step 3: Regressing residuals for Y against residuals for variables in Group A. This establishes the coefficients of global stationary variables.

Step 4: Subtracting the portion of Y explained by the global stationary variables using the calibrated coefficients. The remaining portion of Y is then used to regress against variables in Group B using GWR technique. This gives localized coefficients for variables in Group B.

Step 5: Combining the global and localized coefficients calibrated in Steps 3 and 4 will establish the coefficients of variables in the mixed GWR model.

The GWR 3.0 software was used in modeling calibration (Charlton, 2004). The input file for GWR3.0 software is a spreadsheet or *.csv* extension file containing geographically referenced file of the explanatory variables. The output file is a set of spatially varying coefficients of statistically significant explanatory variables that can be used for GIS mapping. The NLOGIT3.0 software (Econometric Software, 2004) was used to calibrate spatially varying model parameters in GWR models. The NLOGIT3.0 software was also used in mixed GWR model calibration. Prior to model calibration, the correlation matrix of all explanatory variables considered was prepared. This helped identify variables that are highly correlated and thus would not be added in the models simultaneously. Logarithmic transformations were made on the response and some explanatory variables.

## MODEL CALIBRATION RESULTS AND DISCUSSION

### Calibrated Intersection Injury Crash Models

Table 3 lists coefficients of the calibrated global regression and GWR models for predicting intersection injury crashes. Between the choice of negative binomial and lognormal distributions, the later model type has performed better. The significant explanatory variables are natural logarithm of major road AADT, number of left-turn lanes on major road, and income per capita. In general, increase of major road traffic volume will increase the likelihood of injury crashes, while adding number of left-turn lanes on major road will reduce the occurrences of injury crashes. Drivers from households with higher income levels tend to drive better vehicles and less aggressive. This could be reflected by reduction in injury-related crashes.

Table 3. Calibrated Global Regression and GWR Models for Intersection Injury Crashes

Global Regression

$$\ln(INJURY) = 1.23 + 0.13\ln(AADTMAJ) - 0.08\ln(AADTMIN) + 0.00001(INCOME)$$
*(2.89) (1.58) (-3.00) (3.28)*

$$- 0.0000000001(INCOME^2) - 0.013(LNMAJ) + 0.05(NLTMAJ)$$
*(-5.19) (-2.84) (3.28)*

Adjusted $R^2 = 0.10$

GWR Model

$$\ln(INJURY) = b_0(u,v) + b_1(u,v)\ln(AADTMAJ) + b_2(u,v)\ln(AADTMIN) + b_3(u,v)(INCOME)$$
*(1.95) (3.06) (-2.55) (3.42)*

$$+ b_4(u,v)(INCOME2) + b_5(u,v)(LNMAJ) + b_6(u,v)(LTMAJ)$$
*(-5.91) (-1.50) (2.69)*

$b_0(u,v) = - 3497.17 + 6.89\ LONGITUDE + 97.81\ LATITUDE$
*(-12.84) (2.08) (36.03)*

$b_1(u,v) = 141.00 - 0.19\ LONGITUDE - 3.75\ LATITUDE$
*(8.51) (-0.92) (-22.94)*

$b_2(u,v) = 77.26 - 0.35\ LONGITUDE - 2.58\ LATITUDE$
*(14.81) (-5.60) (-49.61)*

$b_3(u,v) = - 0.02 + 0.0002\ LONGITUDE + 0.0005\ LATITUDE$
*(-16.47) (1.51) (43.12)*

$b_4(u,v) = 0.000001 - 0.0000000006\ LONGITUDE - 0.00000003\ LATITUDE$
*(19.90) (-0.85) (-59.09)*

$b_5(u,v) = - 28.18 + 0.24\ LONGITUDE + 1.17\ LATITUDE$
*(-5.23) (3.62) (21.77)*

$b_6(u,v) = 62.94 + 0.20\ LONGITUDE - 1.08\ LATITUDE$
*(16.31) (4.31) (-27.89)*

Adjusted $R^2 = 0.16$

where

| | |
|---|---|
| INJURY | = Injury crashes per intersection per year |
| AADTMAJ | = Major road AADT, vehicles per day |
| AADTMIN | = Minor road AADT, vehicles per day |
| INCOME | = Annual average income per capita, dollars per year |
| LNMAJ | = Number of through movement lanes on major roads |
| LTMAJ | = Number of left-turn lanes on major roads |
| ln() | = Natural logarithm |

Note: *t*-statistic values are listed in brackets.

Significant improvement of model predictability is achieved from the GWR model as compared to the global regression model by an *F*-statistic of 4.47. This value is much greater than the critical value at 10 percent significant, $F_{29.62,\ 595.38,\ 0.10} = 1.36$, see Table 4. The significance of spatial variability in localized model coefficient estimates was examined by conducting a Monte Carlo test. As summarized in Table 5, the local coefficient estimates are all non-stationary for intercept, major and minor road AADT, average income per capita, squared term of average income, major and minor road through lanes, and major road left-turn lanes at 10 percent significance level. Guided by Monte Carlo test results, GWR model was developed.

Table 4. ANOVA Tests for Improved Predictability of GWR Models over the Global Regression Models for Injury, PDO, and Total Crash Predictions (10 Percent Significance Level)

| Model | *F*-Statistic | Critical Value | Significant |
|---|---|---|---|
| Injury | 4.47 | $F_{29.62, 595.38, 0.10} = 1.36$ | Yes |
| PDO | 24.48 | $F_{27.65, 596.35, 0.10} = 1.37$ | Yes |
| Total | 16.38 | $F_{15.59, 609.41, 0.10} = 1.50$ | Yes |

Table 5. Monte Carlo Tests for Spatial Variability in Calibrating Injury, PDO, and Total Crash Models (10 Percent Significance Level)

| Significant Variables | Injury | | PDO | | Total | |
|---|---|---|---|---|---|---|
| | *p*-Value | Significant | *p*-Value | Significant | *p*-Value | Significant |
| Intercept | 0.00 | Yes | 0.60 | No | 0.47 | No |
| Major road AADT | 0.00 | Yes | 0.09 | Yes | - | - |
| Minor road AADT | 0.00 | Yes | 0.04 | Yes | 0.01 | Yes |
| Average income per capita | 0.00 | Yes | 0.03 | Yes | 0.98 | No |
| Average income per capita$^2$ | 0.00 | Yes | 0.01 | Yes | 0.94 | No |
| Major road through lanes | 0.00 | Yes | 0.76 | No | 0.32 | No |
| Minor road through lanes | - | - | 0.01 | Yes | 0.00 | Yes |
| Major road left-turn lanes | 0.00 | Yes | 0.61 | No | 0.87 | No |

**Calibrated Intersection PDO Crash Models**

Table 6 lists coefficients of the calibrated global regression and GWR models for PDO crashes. Again, the lognormal model type has outperformed the negative binomial model type. As seen in the table, the significant explanatory variables are identical to those identified for injury crashes, including natural logarithm of major and minor road AADT, number of major and minor road through movement lanes, number of left-turn lanes on major road, and income per capita. In general, increase of major road traffic volume will increase the likelihood of PDO vehicle crashes, while adding number of left-turn lanes on major road will reduce the occurrences of injury vehicle crashes.

Given an *F*-statistic value of 24.28 shown in Table 4, which is much larger than the critical value at 10 percent significant, $F_{27.65, 596.35, 0.10} = 1.37$, the ANOVA test provides strong evidence to support improvement of GWR model over the global regression model. The significance of spatial variability in localized model coefficient estimates was examined by conducting a Monte Carlo test. As summarized in Table 5, the local coefficient estimates are stationary for intercept, and major road through and left-turn lanes at 10 percent significance level. Guided by Monte Carlo test results, mixed GWR model was calibrated.

Table 6. Calibrated Global Regression and Mixed GWR Models for Intersection PDO Crashes

| |
|---|
| Global Regression |
| $\ln(PDO) = 0.427 + 0.340\ln(ADTMAJ) - 0.084\ln(ADTMIN) - 0.000004INCOME$ |
| *(1.14) (43.32) (-12.93) (-4.00)* |
| $+ 0.000001(INCOME^2) - 0.020LNMAJ + 0.161LNMIN + 0.066LTMAJ$ |
| *(1.65) (-3.16) (8.95) (4.26)* |
| Adjusted $R^2 = 0.99$ |
| Mixed GWR Model |
| $\ln(PDO) = - 0.881 + b_1(u,v)\ln(ADTMAJ) + b_2(u,v)\ln(ADTMIN) + b_3(u,v)INCOME$ |
| *(-6.64) (21.76) (9.15) (5.28)* |
| $+ b_4(u,v)(INCOME^2) - 0.006LNMAJ + b_5(u,v)LNMIN - 0.009LTMAJ$ |
| *(-4.71) (-4.63) (0.83) (-4.41)* |
| $b_1(u,v) = -39.92 + 0.045\ LONGITUDE + 1.05\ LATITUDE$ |
| *(-10.30) (0.95) (27.47)* |
| $b_2(u,v) = -19.60 - 0.14\ LONGITUDE + 0.18\ LATITUDE$ |
| *(-4.73) (-2.75) (4.41)* |
| $b_3(u,v) = -0.005 - 0.000035\ LONGITUDE + 0.000054\ LATITUDE$ |
| *(-29.95) (-15.82) (32.51)* |
| $b_4(u,v) = 0.000000066 + 0.00000000049\ LONGITUDE - 0.00000000054\ LATITUDE$ |
| *(33.24) (20.18) (-28.73)* |
| $b_5(u,v) = 173.259280 + 0.05360786\ LONGITUDE - 4.01892899\ LATITUDE$ |
| *(13.66) (0.34) (-32.08)* |
| Adjusted $R^2 = 0.99$ |

| where | |
|---|---|
| PDO | = PDO crashes per intersection per year |
| AADTMAJ | = Major road AADT, vehicles per day |
| AADTMIN | = Minor road AADT, vehicles per day |
| INCOME | = Annual average income per capita, dollars per year |
| LNMAJ | = Total number of through movement lanes on major roads |
| LNMIN | = Total number of through movement lanes on minor roads |
| LTMAJ | = Total number of left-turn movement lanes on major roads |
| ln() | = Natural logarithm |

Note: *t*-statistic values are listed in brackets.

**Calibrated Intersection Total Crash Models**

Table 7 shows coefficients of the calibrated global regression and mixed GWR models for total vehicle crashes. The lognormal model type has also performed better than the negative binomial model type. As seen in the table, the significant explanatory variables include natural logarithm of minor road AADT, number of major and minor road through movement lanes, number of left-turn lanes on major road, and income per capita. In general, increase of minor road traffic volume will increase the likelihood of total vehicle crashes, while adding number of left-turn lanes on major road will reduce the occurrences of total vehicle crashes.

Given an *F*-statistic value of 16.38 shown in Table 4, which is much larger than the critical value at 10 percent significant, $F_{15.59,\ 609.41,\ 0.10} = 1.50$, the ANOVA test supports the assumption of improved predictability of the GWR model over the global regression model. The significance of spatial variability in localized model

coefficient estimates was examined by conducting a Monte Carlo test. As summarized in Table 5, the local coefficient estimates are non-stationary for minor road traffic volume and minor road left-turn lanes at 10 percent significance level. Mixed GWR model for total crashes was calibrated accordingly.

Table 7. Calibrated Global Regression and Mixed GWR Models for Intersection Total Crashes

| |
|---|
| Global Regression |
| $\ln(TOTAL) = 0.61 + 0.36\ln(ADTMIN) + 0.000002\ INCOME - 0.00000000003(INCOME^2)$ |
| *(5.13) (40.42) (1.99) (-3.62)* |
| $- 0.0035LNMAJ - 0.095LNMIN + 0.016LTMAJ$ |
| *(-1.00) (-7.03) (1.87)* |
| Adjusted $R^2 = 0.98$ |
| Mixed GWR Model |
| $\ln(TOTAL) = 0.79 + b_1(u,v)\ln(ADTMIN) - 0.000019\ INCOME + 0.0000000001(INCOME^2)$ |
| *(9.17) (28.23) (-5.78) (2.00)* |
| $- 0.024LNMAJ + b_2(u,v)LNMIN - 0.043LTMAJ$ |
| *(-1.58) (2.27) (-1.52)* |
| $b_1(u,v) = -56.96 - 0.26LONGITUDE + 0.81LATITUDE$ |
| *(-15.40) (-5.88) (22.13)* |
| $b_2(u,v) = 182.93 + 0.37LONGITUDE - 3.59LATITUDE$ |
| *(14.39) (2.38) (-28.77)* |
| Adjusted $R^2 = 0.99$ |
| where |
| TOTAL = Total crashes per intersection per year |
| AADTMAJ = Major road AADT, vehicles per day |
| AADTMIN = Minor road AADT, vehicles per day |
| INCOME = Annual average income per capita, dollars per year |
| LNMAJ = Total number of through movement lanes on major roads |
| LNMIN = Total number of through movement lanes on minor roads |
| LTMAJ = Total number of left-turn movement lanes on major roads |
| ln() = Natural logarithm |

Note: *t*-statistic values are listed in brackets.

## SUMMARY AND CONCLUSION

This paper introduced the use of the GWR technique to calibrate statistical models for predicting intersection injury, PDO, and total crashes. The current dataset does not support the Poisson assumption as overdispersion effects were observed for injury, PDO, and total crashes, respectively. Also, lognormal distribution appeared to be more appropriate than the negative binomial assumption. Factors contributing to intersection injury, PDO, and total crashes identified include major and minor road daily traffic, number of major and minor road through and left-turn lanes, and household income level. The ANOVA test revealed that improved model predictability is achieved from all crash models developed using the GWR technique compared with global regression models developed using the OLS technique. The Monte Carlo test identified significance of spatial variability of explanatory variables in the GWR models that helped model refinements in developing mixed GWR models as appropriate.

The best models calibrated for the prediction of injury, PDO, and total crashes at densely populated urban intersections are listed below:

- Intersection Injury Crash GWR Model (6)
  Injury crashes/intersection/year = exp[(- 3497.17 + 6.89Longitude + 97.81Latitude) + (141.00 - 0.19Longitude - 3.75Latitude)ln(Major road AADT)+ (77.26 - 0.35Longitude - 2.58Latitude)ln(Minor road AADT)+ (- 0.02 + 0.0002Longitude + 0.0005Latitude) (Income per capita) + (0.000001 - 0.0000000006Longitude -0.00000003Latitude)( Income per capita$^2$)+ (- 28.18 + 0.24Longitude + 1.17Latitude)(Number of major road through movement lanes)+ (62.94 + 0.20Longitude - 1.08Latitude)(Number of major road left-turn lanes)]

- Intersection PDO Crash Mixed GWR Model (7)
  PDO crashes/intersection/year = exp[- 0.881 + (-39.92 + 0.045Longitude + 1.05Latitude)ln(Major road AADT) + (-19.60 - 0.14Longitude + 0.18Latitude)ln(Minor road AADT) + (-0.005 - 0.000035Longitude + 0.000054Latitude) (Income per capita) + (0.000000066 + 0.00000000049Longitude -0.00000000054Latitude)(Income per capita$^2$) - 0.006(Number of major road through movement lanes) + (173.259280 + 0.05360786Longitude - 4.01892899Latitude)(Number of minor road through movement lanes)- 0.009(Number of major road left-turn lanes)]

- Intersection Total Crash Mixed GWR Model (8)
  Total crashes/intersection/year = exp[0.79 + (-56.96 - 0.26Longitude + 0.81Latitude)ln(Minor road AADT) - 0.000019(Income per capita) + 0.0000000001(Income per capita$^2$) - 0.024(Number of major road through movement lanes) + (182.93 + 0.37Longitude - 3.59Latitude) (Number of minor road through movement lanes) - 0.043(Number of major road left-turn lanes)]

The application of GWR technique in this study revealed that it is appropriate for calibrating the non-stationary effect associated with spatial variability. In light of the discussion of results, it is imperative that any meaningful and reliable intersection vehicle crash modeling be preceded by first determining the existence of spatial variability and then conducting model development using appropriate model calibration techniques. This should be carried out with respect to the type of vehicle crashes at intersections in the highway network in question and should utilize, as such as possible, current data on intersection geometric design, traffic control, usage, human factor, time period, climate features, and demographic and socioeconomic characteristics. The study approach can be adopted by state and local transportation agencies to develop enhanced models for intersection vehicle crash prediction, which is expected to serve the need for assessing the effectiveness of intersection safety improvements.

## REFERENCES

AASHTO. (2010). Highway Safety Manual, 1st Edition. American Association of State Highway and Transportation Officials, Washington, D.C.

Ang, A.H-S., and Tang, W.H (2006). Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering, 2nd Edition. John Wiley & Sons, New York, NY.

Bauer, K.M., and Harwood, D.W. (1996). Statistical Models of At-Grade Intersection Accidents. Report No. FHWA-RD-96-125. Federal Highway Administration, U.S. Department of Transportation, Washington, D.C.

Bauer, K.M., and Harwood, D.W. (1999). Statistical Models of At-Grade Intersection Accidents-Addendum. Report No. FHWA-RD-99-094. Federal Highway Administration, U.S. Department of Transportation, Washington, D.C.

Charlton, M. (2004). Geographically Weighted Regression Version 3 Software. National University of Ireland, Maynooth, County Kildare, Ireland.

Econometric Software, Inc. (2004). NLOGIT3.0 Software. Plainview, NY.

Fotheringham, A.S., Brunsdon, C., and Charlton, M. (2002). Geographically Weighted Regression- The Analysis of Spatially Varying Relationships. John Wiley & Sons, Chichester, West Sussex, UK.

Hadayeghi, A., Shalaby, A.S., and Persaud, B. (2003). "Macro-level Accident Prediction Models for Evaluating Safety of Urban Transportation Systems". *Transportation Research Record*, 1840, 87-95.

Harwood, D.W., Bauer, K.M., Potts, I.B., Torbic, D.J., Richard, K.R., Kohlman-Rabbani, E.R., Hauer, E., and Elefteriadou, L. (2002). Safety Effectiveness of Intersection Left- and Right-Turn Lanes. Report No. FHWA-RD-02-089. Federal Highway Administration, U.S. Department of Transportation, Washington, D.C.

Hauer, E. (2002). Observational Before-After Studies in Road Safety- Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety. Elsevier Science Ltd., Kidlington, Oxford, UK.

USDOT (2009). 2008 U.S. Transportation Statistics. U.S. Department of Transportation, Washington, D.C.

Wooldridge, J.M. (2003). Introductory Econometrics: A Modern Approach, 2nd Edition. South-Western, Mason, OH.

# Comparison of Contributory Causes of Fatal Truck and Non-Truck Crashes Using Bayesian Statistical Analysis

Nishitha Bezwada[1] E.I.T. and Sunanda Dissanayake[2] Ph.D., P.E

## ABSTRACT

Of the 41,059 motor vehicle fatalities in 2007, 12 % or 4,808 died in crashes that involved a large truck, but, only 17% of those fatalities were occupants of large trucks. Although large trucks contributed to only 3% of registered vehicles and 7% of vehicle miles travelled in the last 5 years, their impact in terms of severity proves to be a major concern especially for non-truck vehicle occupants. Fatal crash data from the Fatality Analysis Reporting System (FARS) was used to study this issue and to identify factors contributing to truck-involved crashes. Driver, vehicle, and crash-related contributory factors were identified, and the likelihood of occurrence of these factors in truck-involved crashes with respect to non-truck crashes was evaluated using the Bayesian Statistical approach. The most critical factors showed that likelihood ratios for factors like stopped or unattended vehicles, improper following, defective brake system, and vehicle being set in motion by a non-driver have greater probability of occurrence in truck crashes than in non-truck crashes Several other factors such as cellular usage, inattentiveness, and failure to obey traffic rules have been identified as more likely contributory factors for fatal truck crashes. By addressing these factors through the implementation of appropriate remedial measures, the overall truck crash rate can be reduced, which can help in improving overall safety of the transportation system.

## INTRODUCTION

Large trucks have led to high-risk crashes resulting in severe injuries and fatalities in the United States. As 85% of fatalities in truck crashes are non-truck occupants, these crashes, in particular are more devastating for occupants of other vehicles. Though contributing to only 8% of vehicles involved in fatal crashes for the last five years, their impact in terms of severity proves to be a major concern. Large trucks have different performance characteristics than other smaller vehicles. The physical dimension of the vehicle makes it difficult for drivers to maneuver large trucks smoothly on roadways. They can be 40 or more times heavier than the other vehicles in the traffic stream, and have a slower initial pickup and longer deceleration time.

[1,] Graduate Student, Kansas State University, Civil Engineering, 2118 Fielder Hall, Manhattan, KS 66502, Tel: 785- 217-8370,Email: nbezwad@ksu.edu

[2,] Associate Professor, Kansas State University, Civil Engineering, 2118 Fiedler Hall, Manhattan, KS 66502, Tel: 785.532.1540, Fax: 785.532.7717, Email: sunanda@ksu.edu

Truck drivers face many challenges while traversing interstate or state highways at high speeds, or at intersections while taking turns, to have control over the vehicle. Also, the element of blind spots makes it even more challenging for the truck driver and the surrounding vehicle drivers to avoid the heavy crash risk.

Crash statistics observed from the previous years, as seen in Table 1, show significant consistency in the frequencies of the different categories of large truck involved crashes. These trends reflect the need for a more effective analysis, which would provide characteristic facts pertaining to these crashes and help generate productive remedial measures. Achieving effective safety goals to downsize the intensity of the issue will require approaching truck safety aspects from a variety of parameters.

**Table 1. Large Truck Involved Crash Statistics**

| Year | Injury Crashes | Property Damage Only(PDO) Crashes | Fatal Crashes | Single Vehicle Fatalities | Multi Vehicle Crash Fatalities | Total Fatalities |
|---|---|---|---|---|---|---|
| 2003 | 85,000 | 347,000 | 4,335 | 457 | 4,579 | 5,036 |
| 2004 | 83,000 | 312,000 | 4,478 | 469 | 4,766 | 5,235 |
| 2005 | 78,000 | 341,000 | 4,551 | 478 | 4,762 | 5,240 |
| 2006 | 77,000 | 287,000 | 4,321 | 500 | 4,496 | 5,027 |
| 2007 | 76,000 | 333,000 | 4,584 | 502 | 4,306 | 4,808 |

Source: Large Truck Safety Facts 2007

Previous research had shown that many driver-related factors significantly contributed to fatalities in truck crashes (Krishnaswami 2005, Williams and Shabanova 2003, Crum et al. 2005, and Massie 1997). Long and strenuous work hours for truck drivers make them fatigued and reduce their attention levels, which will dramatically increase the crash risk factor. As the amount of truck travel is dramatically increasing with the growing rate of freight transport, which in turn requires continued attention in order to find ways of reducing truck crash risk, the Federal Motor Carrier Safety Administration (FMCSA) has set primary goals for the reduction in commercial truck-related fatalities. Accordingly, it is important for the safety community to identify the characteristics and contributory factors related to large truck involved fatal crashes.

To attenuate the fatal truck crash frequency in the country and achieve the sustainability of this trend seems difficult with the growing rate of movement of people and goods throughout the country. Hence, it is essential to analyze the situations under which fatal truck crashes are occurring. This study deals with identification of these characteristics and contributory causes for all fatal crashes in the country for the period of 2003-2007.

## LITERATURE REVIEW

Extensive research was conducted for nearly a decade in establishing a unique truck crash characteristic data base, and several parameters related to truck crashes

were analyzed (Krishnaswami 2005). Data was acquired from a number of sources including FARS, TIFA (Trucks Involved in Fatal Accidents), and GES (General Estimates System). The study looked at causes of heavy truck driver aggressiveness and their impact on two-vehicle truck/light vehicle crashes, and also developed detailed models which helped propose countermeasures to mitigate collision severity.

Another analysis, conducted using the same data sources on rear-end fatal truck crashes (Craft 2002), showed that though trucks initiate a collision by striking the other vehicle, in fatal crashes, trucks are struck by other vehicles more often. In a study about motor vehicle crash rate comparisons made with respect to truck or non-truck drivers, the at-fault status was observed as the chief criteria (Williams and Shabanova 2003). Data from FARS was used for the period of 1996-2000. Deaths in crashes involving one or more passenger vehicles, for which drivers of various ages were likely to be responsible, per 100,000 licensed drivers by occupant type and many other categories were calculated (Williams and Shabanova 2003).

Many other projects, based on the analysis of driver parameters like age and gender, were used to generate models using driver behavior factors to have a precise understanding of driver issues in crashes. One study investigated the influence of carrier scheduling practices on truck driver fatigue by developing and empirically testing a truck driver fatigue model (Crum et al. 2005). Earlier than this, a research team had developed another model with the four predictor variables of driver age, gender, time of the day, and average annual mileage (Massie 1997). The effect of these four variables on crash involvement rate was studied and their level of significance was obtained. Additionally, the U.S. Department of Transportation report noted three separate studies and found driver fatigue and loss of alertness to be primary or probable causes for crashes with high severity (Wylie et al. 1996).

In order to identify unsafe driver actions that lead to fatal car-truck crashes, a study analyzed two-vehicle crashes in the 1995–98 Fatality Analysis Reporting System (FARS) database to compare car-car crashes with car-truck crashes (Kostyniuk et al. 2002). A key finding of this study was that most of the 94 unsafe driver acts were about as likely in fatal car-truck crashes as in fatal car-car crashes. Therefore, general safe driving practices are also relevant around large trucks.

The United States General Accounting Office issued a report on the "Share the Road Safely" program whose goal is to educate the public about driving safely around large trucks (GAO-03-680 2003). This report analyzed crash risk factors that predominantly arise while driving around large trucks. The program elaborated the necessity of having specific roadway educative measures for the public to mitigate this issue and lower the truck crash rate in general.

## METHODOLOGY

### Crash Data

Data for the study was procured from the National Highway Traffic Safety Administration's Fatality Analysis Reporting System (FARS) for the period 2003-2007. The database documents descriptive data on vehicles, drivers, roadways, and environmental conditions collected from police reports, emergency medical service reports, hospital records, and coroner's reports of all fatal crashes in the country.

From this database, truck and non-truck crashes were the two categories examined in the comparitive study. In this study, a truck crash was a crash which involved at least one truck whose gross body weight was greater than 10,000 pounds. A non-truck crash was defined as a crash which did not involve a truck. Trucks considered for this study were vehicles with body type codes 61 to 64, 66, 67, 71, 72, 78, and 79 in the FARS database. These specific body types were considered as they included trucks which had a gross body weight greater than 10,000 pounds. All other motor vehicles, other than those body types and ones which had a body weight less than 10,000 pounds, were considered as non-truck vehicles.

Files from the database were merged using unique crash, person, and vehicle identification codes employing SAS computing software. Various crash characteristics were recorded using filtering techniques in Microsoft Excel and Access. After suitably merging and filtering accident, person, and vehicle files, fatal truck crash data for the five year time period of 2003-2007 was combined and truck and non-truck crash cases were separated to obtain more consolidated results with respect to several parameters and their frequencies.

Further, the values obtained were compared at various levels to analyze trends and patterns of specific crash parameters with respect to time or type of crash, or the extent of fault of the drivers involved. Eventually the driver, crash, and vehicle-related factors were extracted to compare the predominance of these factors in both truck and non-truck crashes.

**Bayesian Statistical Approach**

The statistical Bayesian approach is an effective tool in recognizing the predominance of crash-related factors while comparing truck and non-truck crashes in the given data set. The computation of likelihood ratios, using Bayesian posterior probabilities, is valid and useful. It makes good logical sense, while producing significant results from projected analysis of crash factors.

Equation (1) describes the conditional probability of the occurrence of a driver, vehicle, or crash-related contributory cause (CC), given that it is a truck crash.

$$P(CC/Truck) = \frac{P(Truck/CC) \ * \ P(CC)}{P(Truck)} \qquad (1)$$

where,

P(Truck/CC) = Probability that the crash was a truck crash, given that a specific contributory cause was also reported. As shown in Equation (2), this value is estimated from the data by considering total number of crashes and those in which a truck crash and its contributory factor were coded together.

P(CC) = Overall probability of the specific driver, vehicle, or crash-related cause being reported as a contributing factor, and as shown in Equation (4), is estimated from the numbers of cases in which the CC was reported in the dataset.

P(Truck) = Overall probability that a crash was a truck crash and was estimated from the data as shown in Equation (3).

$$P(Truck/CC) = \frac{No.of\ Truck\ Crashes\ with\ that\ Contributory\ Factor}{No.of\ All\ (Truck\ and\ Non-Truck)\ Crashes\ with\ that\ Factor} \quad (2)$$

$$P(Truck) = \frac{Number\ of\ Truck\ Crashes}{Number\ of\ All\ (Truck\ and\ Non-Truck)\ Crashes} \quad (3)$$

$$P(CC) = \frac{Number\ of\ Crashes\ with\ that\ Contributory\ Factor}{Number\ of\ All\ (Truck\ and\ Non-Truck)\ Crashes} \quad (4)$$

Similarly, the conditional probability of a contributory cause for a given non-truck crash is estimated, and the ratio of these probabilities generates the likelihood ratio of that contributory factor.

$$Likelihood\ Ratio = \frac{P(CC/Truck\ Crash)}{P(CC/Non-Truck\ Crash)} \quad (5)$$

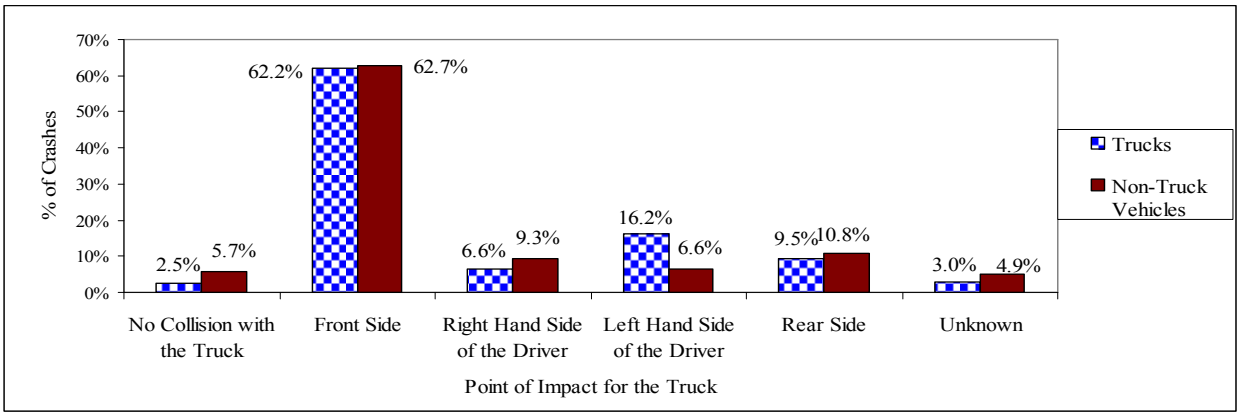
The likelihood ratio of a given contributory factor being recorded in a truck crash as compared with a non-truck crash was assessed from crash records. This likelihood ratio is the probability of a crash being a truck crash when the contributory factor was recorded, as compared with the probability of a crash being a non-truck crash when the same contributory factor was identified. The larger the likelihood ratio, the greater the association between the contributory factor and truck crashes relative to non-truck crashes.

## RESULTS

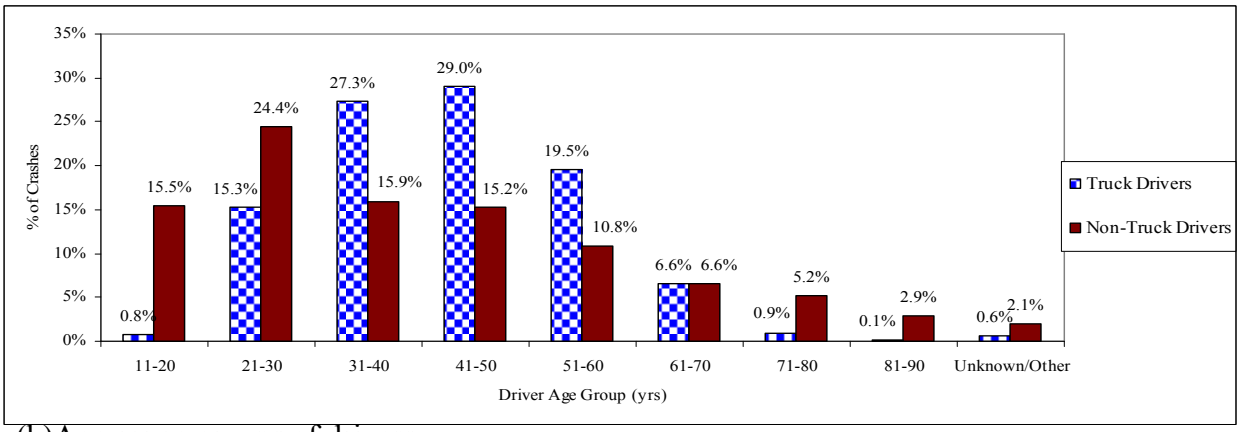
### Characteristic Comparison of Truck and Non-Truck Crashes

For this analysis the entire fatal crash data for the period of 2003-2007 was divided into crashes which involved trucks and those which did not involve trucks, or non-truck crashes. Different characteristic factors such as initial point of impact, driver age ranges, posted speed limits, manner of collision, level of deformation, rural/urban split and types of traffic flow ways, and roadway categories were compared by taking the percentage with respect to total number of truck or non-truck crashes as the base value.

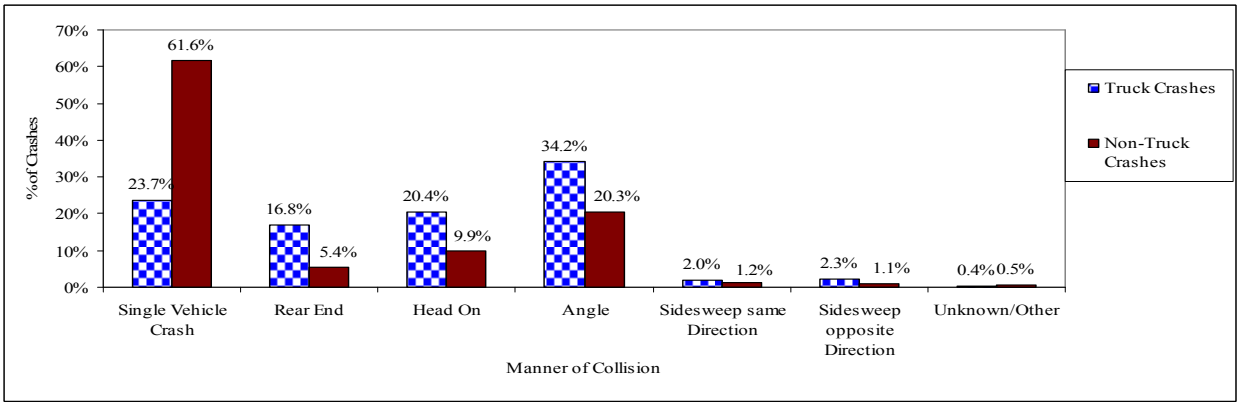
The initial impact point for vehicles in truck and non-truck crashes was observed to be mostly on the front side from Figure 1(a). Although all other categories had lower proportions in both truck and non-truck crashes, left-hand side of the driver as the impact point had a comparatively larger proportion of fatal truck crashes. Also, a larger proportion of truck drivers involved in fatal crashes seemed to be of the age group 41-50 yrs, whereas the non-truck drivers were mostly in the 21-30 yrs age group. Figure 1(b) shows that starting from the age group of 31-40 yrs, truck drivers had larger involvement than non-truck drivers in fatal crashes. Also, it was observed from Figure 1(c) that majority of the fatal non-truck crashes were single-vehicle crashes but most of the fatal truck crashes were angle crashes.




(a) Impact point on the vehicle in fatal crashes.



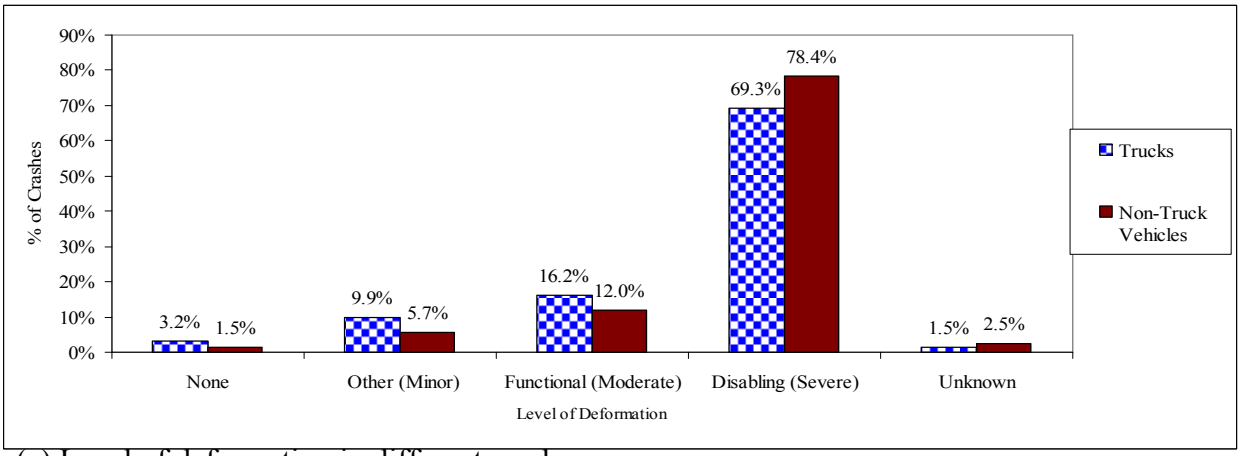

(b)Age-range groups of drivers.



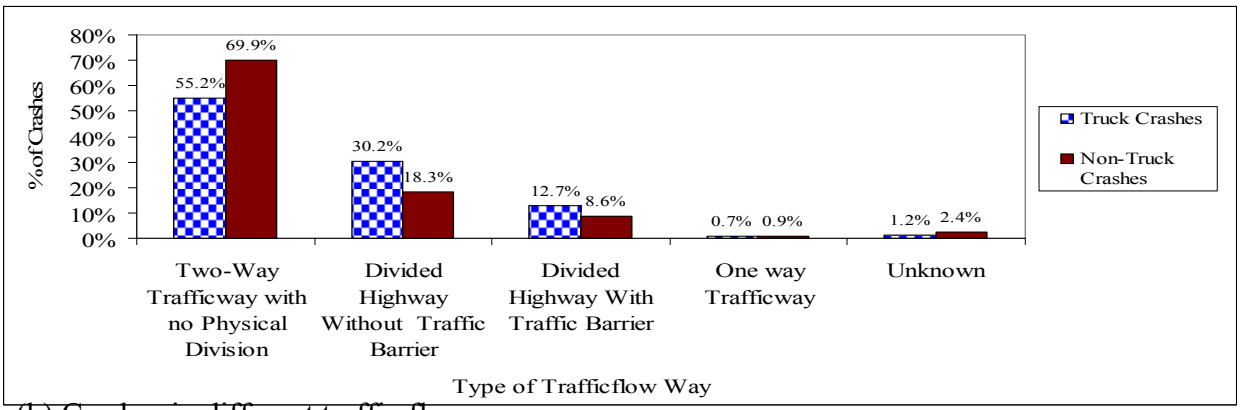

(c) Manner of collision for vehicles.

**FIGURE 1 Distribution of crashes in truck and non-truck collisions.**

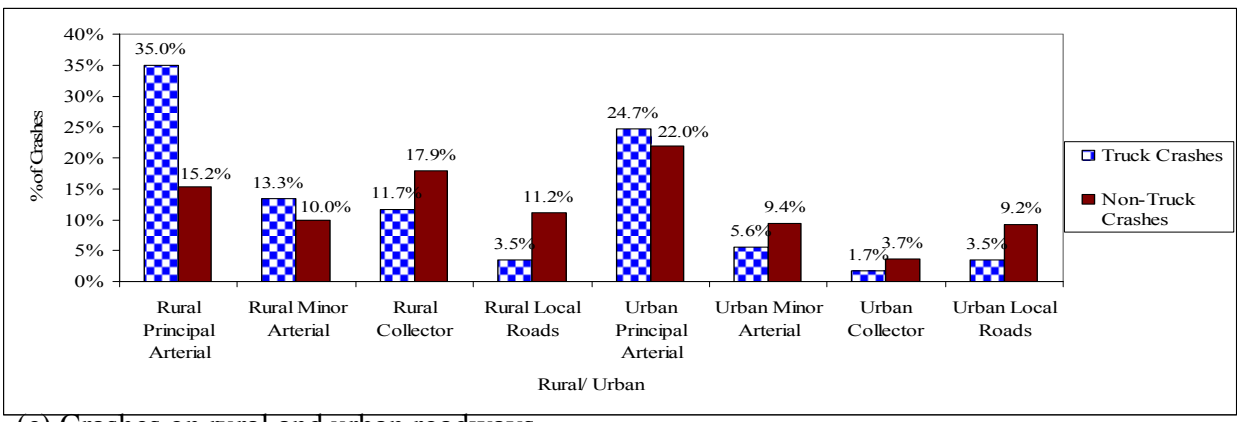
Truck and non-truck fatal crashes most commonly resulted in disabling vehicle deformation as shown in Figure 2(a) but it is observed from Figure 2(b) that more than half of the crashes in trucks and non-trucks occurred on two-way trafficways with no physical division. From Figure 2(c), it is observed that arterial roadways in both urban and rural sectors had a higher predominance of fatal truck crashes, whereas collector and local roads had a higher predominance of non-truck crashes.




(a) Level of deformation in different crashes.




(b) Crashes in different traffic flow ways.




(c) Crashes on rural and urban roadways.

**FIGURE 2 Comparison of truck and non-truck collisions using different factors.**

## Contributory Causes for Truck and Non-Truck Crashes

The following tables show the likelihood of occurrence of contributory factors in truck crashes when compared to non-truck crashes. If probability of the factor is greater than one, it indicates that the factor was more predominant in truck crashes than in the non-truck crashes. The factors in the tables belong to the categories of driver-related, vehicle-related, or crash-related issues. The likelihood ratios are recorded in the

descending order of predominance in each category. Crashes might have more than one contributory factor leading to the crash, as FARS records up to four driver-related, three crash-related, and two vehicle-related factors per crash. Hence, the sum of the number of crashes in truck and non-truck categories will not be equal to the number of crashes that occurred in that period.

Vehicle-related factors contributing to a fatal truck crashes are listed in Table 2. As most of the utility vehicles are trucks rather than other motor vehicles, this cannot be considered as a contributory cause. However defective brake systems had the second highest likelihood ratio, which seems to be more predominant in fatal truck crashes rather than in other vehicle crashes. Defective lights, mirrors, and engines also appear to have more likelihood because of the severe wear and tear trucks undergo as a result of long miles traveled.

These factors recorded as vehicle-related factors are subjective with respect to police officers present at crash sites. As officers are not professional vehicle inspectors these records might not be precise to the maximum extent.

**TABLE 2 Conditional Probabilities and Likelihood Ratio for Vehicle-Related Factors**

| Vehicle-Related Factor (VF) | Number of Truck Crashes | Number of Non-Truck Crashes | Conditional Probability of This VF Given a Fatal Truck Crash | Conditional Probability of This VF Given a Fatal Non-Truck Crash | Likelihood Ratio |
|---|---|---|---|---|---|
| Vehicle Identified as Utility/Emergency/Other Working Vehicle | 188 | 80 | 0.00859 | 0.00047 | 18.27 |
| Defect in Brake System | 445 | 421 | 0.02033 | 0.00247 | 8.22 |
| Defects in Lights/Horn/Mirror/Wiper | 89 | 260 | 0.00407 | 0.00153 | 2.66 |
| Defects in Steering/Suspension/Engine/ Exhaust System | 77 | 263 | 0.00352 | 0.00155 | 2.27 |
| Other Vehicle Defects(Wheels/Doors/Safety Belts/Air Bags) | 124 | 499 | 0.00567 | 0.00293 | 1.93 |
| Defective Tires | 358 | 2501 | 0.01636 | 0.01470 | 1.11 |
| Identified Vehicle Registration as Handicapped | 65 | 581 | 0.00297 | 0.00341 | 0.87 |
| Identified as a Hit-and-Run Vehicle | 306 | 7727 | 0.01398 | 0.04540 | 0.30 |

FARS records driver-related factors into 94 different categories which include mental, psychological, vision obscured, environmental, and other miscellaneous factors. Of these 94 cases, only those which reasonably reflect the truck driver contributing to the occurrence of the crash were included. As shown in Table 3, the conditional probability of each driver factor in truck and non-truck crashes and their likelihood ratios were estimated. When factors having considerable number of frequencies were selected, they were listed in descending order of their likelihood ratios. Stopped or unattended vehicles,

improper following, and starting and backing the vehicle improperly are the factors with the highest likelihood ratios, which show they may contribute to fatal truck crashes more often than fatal non-truck crashes.

**Table 3 Likelihood Ratios for Driver and Crash Related Factors**

| Driver-Related Factor(DF) | Truck Crashes | Non-Truck Crashes | Conditional Probability of This DF Given a Fatal Truck Crash | Conditional Probability of This DF Given a Fatal Non-Truck Crash | Likelihood Ratio |
|---|---|---|---|---|---|
| Stopped or Unattended Vehicle | 501 | 1019 | 0.02289 | 0.00599 | 3.82 |
| Following Improperly | 903 | 1902 | 0.04126 | 0.01118 | 3.69 |
| Starting or Backing Improperly | 147 | 349 | 0.00672 | 0.00205 | 3.27 |
| Overloading or Improper Loading of the Vehicle | 111 | 309 | 0.00507 | 0.00182 | 2.79 |
| Making Improper Exit or Entry | 76 | 287 | 0.00347 | 0.00169 | 2.05 |
| Erratic Lane Change | 525 | 2129 | 0.02399 | 0.01251 | 1.91 |
| Cellular Telephone in Use in Driving | 765 | 3488 | 0.03496 | 0.02049 | 1.70 |
| Signal Inattention/Unfamiliar Roadway | 128 | 643 | 0.00585 | 0.00378 | 1.54 |
| Driving on Wrong Side of the Road | 557 | 3379 | 0.02545 | 0.01985 | 1.28 |
| Failure to Yield Right of Way | 2968 | 18801 | 0.13562 | 0.11047 | 1.22 |
| Failure to Obey Traffic Rules | 1688 | 10899 | 0.07713 | 0.06404 | 1.20 |
| Drowsy ,Sleepy, Fatigued | 683 | 4499 | 0.03121 | 0.02644 | 1.18 |
| Tire Blow Out or Flat Tire | 134 | 887 | 0.00612 | 0.00521 | 1.17 |
| Inattentive(Talking, Eating) | 2569 | 17407 | 0.11739 | 0.10228 | 1.14 |
| Driving/Passing in Prohibited or Wrong Direction | 83 | 701 | 0.00379 | 0.00412 | 0.92 |
| Passing Where Prohibited by Posted Signs | 104 | 900 | 0.00475 | 0.00529 | 0.89 |
| Failing to Dim Lights or Have Lights When Required | 39 | 338 | 0.00178 | 0.00199 | 0.89 |
| Crash-Related Factor(CF) | Number of Truck Crashes | Number of Non-Truck Crashes | Conditional Probability of This CF Given a Fatal Truck Crash | Conditional Probability of This CF Given a Fatal Non-Truck Crash | Likelihood Ratio |
| Recent Previous Crash Nearby/ Vehicle Set in Motion by a Non-Driver | 416 | 1025 | 0.01901 | 0.00602 | 3.15 |
| Vehicle Struck by Falling Cargo | 558 | 1496 | 0.02550 | 0.00879 | 2.90 |
| Other Construction/ Work Area Condition | 122 | 342 | 0.00557 | 0.00201 | 2.77 |
| Inadequate Warning of Exits, etc. | 15 | 57 | 0.00069 | 0.00033 | 2.04 |
| Aggressive Driving or Road Rage of Non-Contact Vehicle Driver | 102 | 391 | 0.00466 | 0.00230 | 2.02 |

Also, Table 3 shows the top five contributing crash-related factors as defined by the database. Recent previous crash nearby/ vehicle set in motion by a non-driver, work

area conditions, poor shoulder conditions, and inadequate warning signs are the topmost factors which have more likelihood in truck crashes than non-truck crashes.

In addition, erratic lane change, cellular usage, and signal inattention are also factors significantly contributing to fatal crashes. Truck drivers appear to be more fatigued, drowsy, and inattentive when compared to other vehicle drivers, having a likelihood ratio of greater than one.

## CONCLUSIONS

Certain significant characteristics of fatal truck crashes have been observed from this analysis. Fatal crash frequency was observed to be greater with the initial impact point for the vehicle in the front side of the truck rather than anywhere else. Fatigue, drowsiness, and inattention were observed to be more predominant in truck drivers than in other motor vehicle drivers. Two-way two-lane traffic flow ways with no physical divisions are a leading characteristic to higher crash risk and fatalities. Such roadways could be altered by providing the necessary changes in the roadway design. Improper driving and non-compliance to traffic regulations were observed to be the main driver-related contributory factors in cases of fatal truck crashes.

From the likelihood ratios, stopped or unattended vehicles or improper following had greater probabilities of occurrence in truck crashes than in non-truck crashes. Other factors like cellular usage, failure to yield right of way, inattentiveness, and failure to obey traffic rules also had greater probabilities in truck crashes.

## REFERENCES

1. Krishnaswami, V., D. Blower, L. Schneider, and D. Putcha. (2005). "Heavy Truck Aggressiveness Reduction Statistics, Analysis, and Countermeasures". Publication University of Michigan Transportation Research Institute (UMTRI) -2005-38, UMTRI, U.S.
2. Craft, R. (2002). MC-RIA. Contact no. (202) 366-0324. "Rear-End Large Truck Crashes". Federal Motor Carrier Safety Administration (FMCSA).
3. Williams., A.F. and I.V. Shabanova. (2003) Responsibility of Drivers, by Age and Gender, for Motor Vehicle Crash Deaths. *Journal of Safety Research*, Vol. 34.
4. Crum, M.R., P.C Morrow, and R. Paula. (2005). "The Influence of Carrier Scheduling Practices on Truck Fatigue". *Transportation Journal*, pg.20-41.
5. Massie, D.L., D. Blower, and K.L. Campbell. (1997). *Short-Haul Trucks and Driver Fatigue*. Office of Motor Carriers, Washington, D.C., U.S. Department of Transportation, Federal Highway Administration.
6. Wylie, C.D., T. Shultz, J.C. Miller, M.M. Mitler, and R.R. Mackie. (1996). "Commercial Motor Vehicle Driver Fatigue and Alertness Study". Publication FHWA-MC-97-002. FHWA, U.S. Department of Transportation.
7. Kostyniuk P. L., Fredrick M. Streff, and Jennifer Zakrajsek. (2002). "Identifying Unsafe Driver Actions that Lead to Fatal Car-Truck Crashes". *AAA Foundation for Traffic Safety*.
8. United States General Accounting Office. (2003). "Truck Safety: Share the Report Safely Program Needs Better Evaluation of Its Initiatives". Publication GAO-03-680. GAO, U.S.