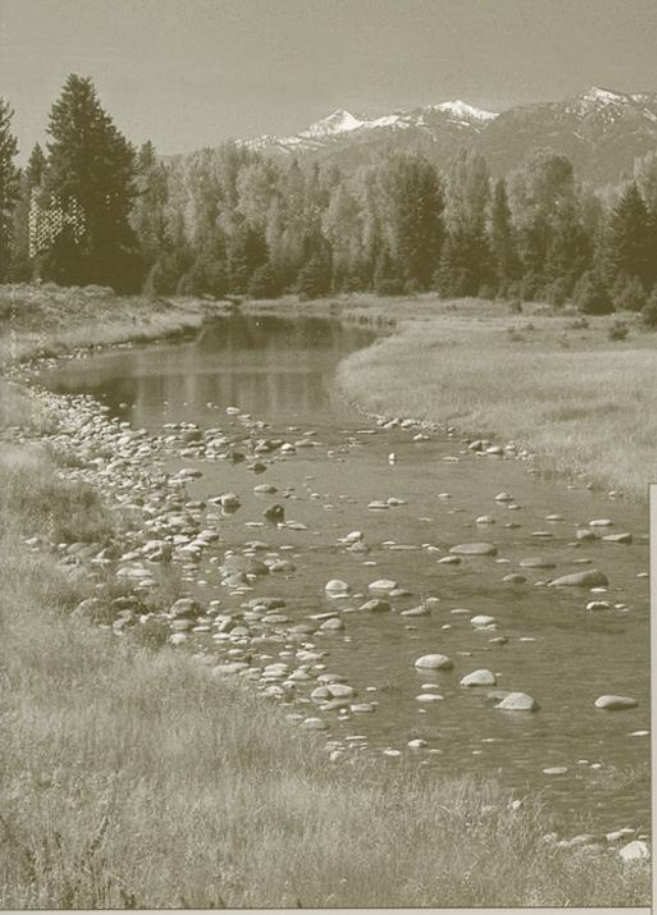
Spatial methods for solution of

environmental and hydrologic problems

science, policy, and standardization



Editors:
Vern H. Singhroy,
David T. Hansen,
Robert R. Pierce,
and A. Ivan Johnson

STP 1420

ASTM INTERNATIONAL

STP 1420

Spatial Methods for Solution of Environmental and Hydrologic Problems—Science, Policy, and Standardization

Vernon Singhroy, David T. Hansen, Robert R. Pierce, and A. Ivan Johnson, editors

ASTM Stock Number: STP1420

ASTM International
100 Barr Harbor Drive
PO Box C700
West Conshohocken, PA 19428-2959

Printed in the U.S.A.

Library of Congress Cataloging-in-Publication Data
ISBN:

Spatial methods for solution of environmental scales using remote sensing and GIS / Vern Signhroy ... [et al.], editors.
p. cm.
"ASTM stock number: STP1420."
Proceedings of a symposium held on 25 January 2001 in Reno, Nevada.
ISBN 0-8031-3455-X
1. Hydrology--Remote sensing--Congresses. 2. Hydrology--Mathematical models--Congresses. 3. Geographic information systems--Congresses. I. Singhroy, Vernon.

GB656.2.R44S625 2003
628.1--dc21

2002043885

Peer Review Policy

Each paper published in this volume was evaluated by two peer reviewers and at least one editor. The authors addressed all of the reviewers' comments to the satisfaction of both the technical editor(s) and the ASTM International Committee on Publications.

To make technical information available as quickly as possible, the peer-reviewed papers in this publication were prepared "camera-ready" as submitted by the authors.

The quality of the papers in this publication reflects not only the obvious efforts of the authors and the technical editor(s), but also the work of the peer reviewers. In keeping with long-standing publication practices, ASTM International maintains the anonymity of the peer reviewers. The ASTM International Committee on Publications acknowledges with appreciation their dedication and contribution of time and effort on behalf of ASTM International.

Printed in Bridgeport, NJ
January 2003

Foreword

The Symposium on Spatial Methods for the Solution of Environmental and Hydrologic Problems: Science, Policy, and Standardization—Implications for Environmental Decisions was held on 25 January 2001 in Reno, Nevada. ASTM International Committee D-18 on Soil and Rock, in cooperation with ASTM committees D-34 on Waste Management, E-47 on Biological Effects and Environmental Fate, and E-50 on Environmental Assessment served as its sponsors. The symposium chairmen of this publication were Vern Singhroy, David T. Hansen, Robert R. Pierce, and A. Ivan Johnson.

Contents

SESSION V: DEVELOPMENT OF STANDARD DATA SETS

SESSION VI: NATIONAL DATA

Overview

The Symposium on Spatial Methods for the Solution of Environmental and Hydrologic Problems; Science, Policy, and Standardization was held in Reno Nevada on January 25 and 26, 2001 as part of the D-18 scheduled meetings. The symposium was sponsored by ASTM Committee D-18 on Soil and Rock in cooperation with ASTM Committee D-34 on Waste Management, E-47 on Biological Effects and Environmental Fate, and E-50 on Environmental Assessment. Cooperating organizations in this symposium are the International Commission on Remote Sensing of the International Association of Hydrologic Sciences, the Canada Centre for Remote Sensing, the U.S. Geological Survey, and the U.S. Bureau of Reclamation. Over the past two decades, the simple graphic display of environmental data with hydrologic or cultural features of interest has progressed rapidly to modeling and analysis of environmental data with other spatially represented data. New tools such as global positioning systems (GPS) have developed to rapidly and accurately collect the position of data locations. Computer system component architecture has progressed to where data from one application can be incorporated with other applications. This includes the linkage and integration of surface water and groundwater modeling programs with geographic information systems (GIS). Geostatistical and statistical software packages have been developed and integrated with GIS and other spatial modeling software. Standards in computer systems and in the definition of spatial data have progressed to the point where geospatial data in a variety of formats and from different sources can be displayed and manipulated on common computing platforms and across the Internet.

Considering these developments, this symposium focused on issues related to spatial analysis of environmental or hydrologic problems. These issues include methods of spatial analysis, accuracy in the location and spatial representation of data and real world features, and emerging standards for digital spatial methods. Major session topics for the symposium included:

- Modeling and Spatial Analysis of Environmental and Hydrologic Systems
- Accuracy and Uncertainty in Spatial Data and Analysis
- Standardization and Standard Digital Data

This overview covers papers presented at the symposium and additional papers contained in this volume related to these topics.

Accuracy and Uncertainty in Spatial Data and Analysis

Underlying all spatial data is the coordinate control for features represented which are carried into some common coordinate system for manipulation and analysis. This may be standard survey control with measured bearings and distances from marked points or it may be established geodetic control with measured latitude and longitude values of established points. These established points, which in the United States are maintained and reported by the National Geodetic Survey, serve as the underlying control for the national map series and for other data that is compiled or registered to these base maps. GPS has rapidly developed as a tool to accurately capture coordinate values for both standard survey control and for geodetic coordinates. GPS is also commonly used for identifying sample locations and mapping features on the ground. This session discussed issues in the use and application

of GPS. This includes the characteristics of GPS and the various modes of operation and factors affecting the accuracy of values collected and reported by GPS receivers. This discussion included techniques for improving the values reported by post processing and the use of differential GPS. Ikehara discusses the application of GPS for developing highly accurate network for elevation control survey and factors affecting the reported values.

The national map series developed and maintained by mapping organizations in various countries form the underlying accuracy level for much environmental and hydrologic data. In this session, a variety of data products were presented by mapping organizations in Canada and the United States. Singhroy discusses the development of a standard merged product of satellite imagery and elevation data for resource mapping in Canada. In this session, the development and management of high-resolution elevation data and the stream network data for the United States was discussed.

Statistics and geostatistics applied to data represented in GIS or captured via remote sensing are important tools for environmental and hydrologic analysis. This session discussed the application of krieging and other geostatistical techniques. It included a session on fractal analysis for spatial applications. Other topics discussed in this session included the difficulty in defining the level of accuracy for environmental and hydrologic data used in spatial analysis including the variability in spatial accuracy of multiple data sets. Often, it is easier to discuss the uncertainty associated with the data or within the analysis. Knowlton, Peterson, and Zhang model uncertainty in spatial variability for risk assessment in a decision support system. Hansen discusses the uncertainty associated with habitat labels assigned to spectrally defined polygons. Knowlton and others describe spatial variability, uncertainty, and risk for use in decision support systems.

Modeling and Spatial Analysis of Environmental and Hydrologic Systems in Spatial Data Environments

This topic covered the use of spatial techniques to model environmental systems and the development of object models for hydrologic systems. This discussion included the linkage between detailed digital elevation models at a scale of 1:24,000 or better with object models of the hydrologic or stream network. García describes the development of a flood warning system for watersheds in Spain using spatiotemporal techniques to model flood events. Starks, Heathman, and Ross discuss modeling the distribution of soil moisture with remote sensing and GIS. Rich discusses data integration with GIS as a management tool for decision support. Paillet describes the integration of surface and borehole geophysical measurements to model subsurface geology and ground-water systems. Owe and De Jeu describe efforts to model surface soil moisture from satellite microwave observations. Flügel reports on the use of response units to assess erosion processes in semiarid areas in southern Africa.

Standardization and Standard Digital Data Sets

Interspersed throughout the symposium were discussions and presentations on standardization at national and international levels. This includes standards on methods, descriptions, and digital data products such as the watershed boundary standards for the United States. Hansen reviews the status of standards in use by the U.S. government related to GIS data and the role of other organizations in the development of standards for GIS. Recently, active development of standard data sets has been taking place. Singhroy reports on the development of standard merged products of satellite imagery and elevation data for natural resource mapping in Canada. In the United States, the U.S. Geological Survey has been particularly active in the development of a series of standard digital databases.

David T. Hansen
U.S. Bureau of Reclamation
Sacramento, CA

Geospatial Data Development and Integration

David. W. Rich[1]

Integration of Data Management, GIS, and Other Data Uses

Reference: Rich, D. W., **"Integration of Data Management, GIS, and Other Data Uses,"** *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy and Standardization, ASTM STP 1420*, D. T. Hansen, V. H. Singhroy, R. R. Pierce, and A. I. Johnson, Eds., ASTM International, West Conshohocken, PA, 2002.

Abstract: Efficient data management is becoming increasingly important in managing site environmental projects. Key decisions in designing a data management system include where the data will reside, how data will be moved between interested parties, and how connections will be built between applications for managing, interpreting, and displaying the data. Options for database locations include stand-alone, client-server, and increasingly, Web-based. Formats and protocols for data transfer or data access can be a challenge, and the advantages of direct connections (as opposed to export-import) must be weighed against the effort required to implement the connections. The benefits to be gained by overcoming data management and communication obstacles can, in many cases, greatly exceed the effort expended. If the end justifies the means, the displays that can be generated using GIS and other technologies can provide a much greater understanding of site technical and administrative issues.

Keywords: data management, geographic information systems, integration, data formats, protocols, stand-alone, client-server, web-based, export, import, laboratory data, statistics, mapping.

Nomenclature

ASCII	American standard code for information interchange
CAD	Computer-aided drafting
COM	Component object model
CORBA	Common object request broker architecture
DCOM	Distributed component object model
EDD	Electronic data deliverable
EJB	Enterprise Java beans
GIS	Geographic information system
GML	Geography markup language
GTTP	Geographic text transfer protocol

---

[1]President, Geotech Computer Systems, Inc., 6535 S. Dayton St., Suite 2100, Englewood, CO 80111 USA drdave@geotech.com

HTML	Hypertext markup language
IMS	Internet mapping server
LAN	Local area network
LIMS	Laboratory information management system
ODBC	Open database connectivity
SOAP	Simple online access protocol
SQL	Structured query language
TCP-IP	Transmission control protocol - internetworking protocol
WAN	Wide area network
XML	Extensible markup language
XSL	Extensible stylesheet language

Background

Management of sites with environmental issues is maturing. Many facilities have been quite thoroughly investigated, and in some cases remediated at least to some degree. These facilities are moving into an ongoing monitoring situation that can be expected to last for many years. Management of these projects can benefit greatly by efficient data management from the time samples are taken in the field, through laboratory analysis and validation, to the generation of final output. Streamlining the movement of data can save a significant percentage of the time and cost for these activities, providing a good return on investment on technology purchases and integration efforts [*1,2,3,4,5,6*]. In one example, a database user reported a decrease in time to import a laboratory EDD (Electronic Data Deliverable) from thirty minutes to five minutes each after implementation of more efficient software and a more structured process. This will add up to significant savings over the life of the project, far more than the implementation cost.

A key decision in designing a data management system is where the data will reside. Related to this is the issue of how data will be moved between interested parties, and between applications for managing and interpreting the data. Often in the planning stage for integration projects not enough attention is paid to the movement of the spatial and non-spatial data prior to its use in the GIS and other applications [*7,8,9*]. Likewise, the storage architecture of the spatial and non-spatial data can have a great impact on system performance.

Location Options

The location of applications, data, and users can be presented as three spectra (Figure 1). Looked at from this perspective, the ways that an application such as a database can store data range from stand-alone through shared files and client-server, to Web-enabled and Web-based. The ownership of data (who created it and who can use it) can range from proprietary through commercial, to public domain. The relationship between users extends from one user at a stationary desktop, to portable laptops and personal digital assistants, through public portals where anyone can walk up and access data and applications. The computing industry is migrating from left to right in this diagram, and the design of data management systems and the software that interacts

with them should take this into account. Environmental applications, data, and users are no exception.

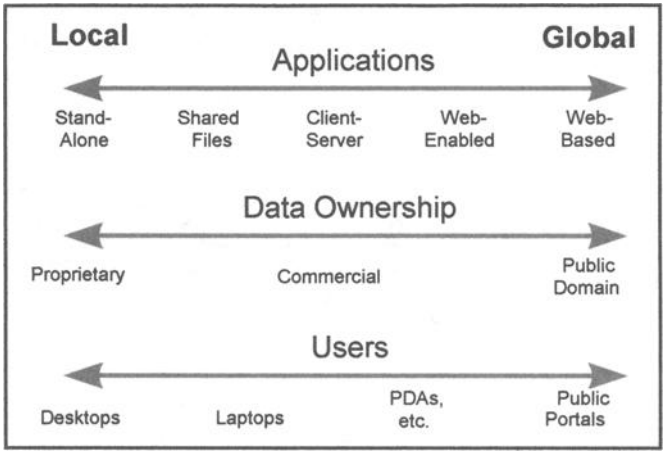


Figure 1 - *Local-Global Spectra.*

Stand-alone - The simplest database location is a stand-alone system. In this design, one user sitting at an isolated computer works with data and applications stored on that computer.

Shared files - Once the computers in an organization are networked, data management can progress to shared files. In this model, all of the applications such as the database and GIS execute on each person's local computer, but the applications connect to files either on a server or one of the workstations. The computer containing the data only shares the files. It doesn't provide any processing.

Client-server - The next step beyond a shared file system is a client-server system, usually based on a LAN (Local Area Network). In a client-server system, the server runs an application that provides data to applications running on each client computer. This provides significant performance benefits over a shared file system. A LAN design works for an organization with centralized data management. With environmental data, this is not always the case. Often the data for a particular facility must be available both at the facility and at the central office, as well as to other people, such as consultants and regulators. In situations where a full-time, high-speed communication link is or can be made available, a WAN (Wide Area Network) is often the best choice.

There are several situations where a client-server connection over a LAN or WAN is not the best solution, due to connection speed or data volume, and a distributed database system can make sense. This design often uses the same technology as the client-server system, but the data is replicated to multiple locations, which raises a suite of management issues.

Web-enabled - The Internet is causing changes in computing in a way as significant as those caused by the personal computer. Applications can be Web-enabled, or Web-based. A Web-enabled application is one where the Internet is used for communication, but the application still runs primarily on the local machine. An example would be a database that can download data from a Web server at a laboratory, and then store and processes it locally. Another example would be a GIS package that can download base map data from a Web site, and then display it using code running locally. This is referred to as a "thick client" system.

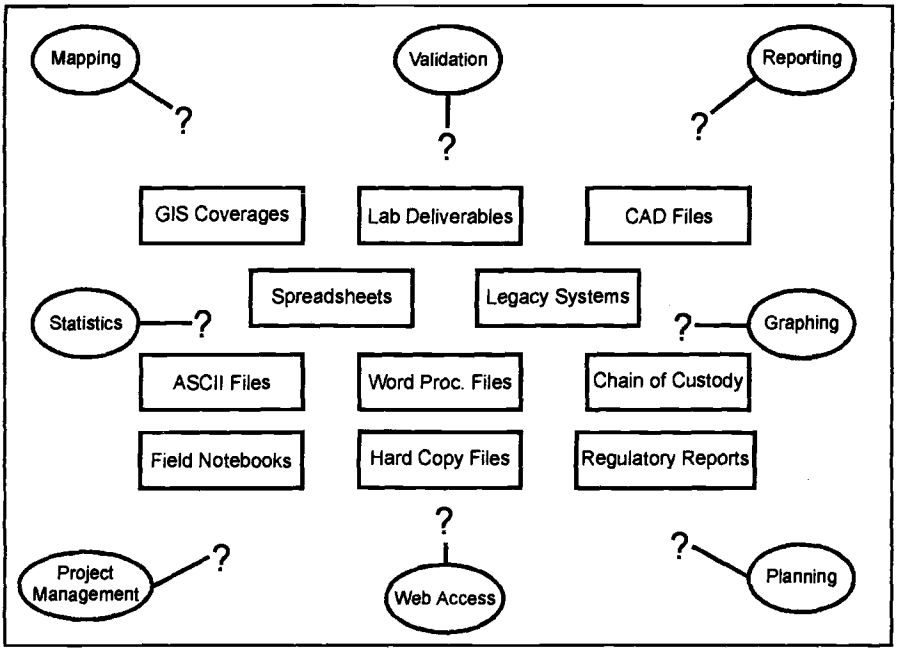


Figure 2 - *Application Connection Requirements.*

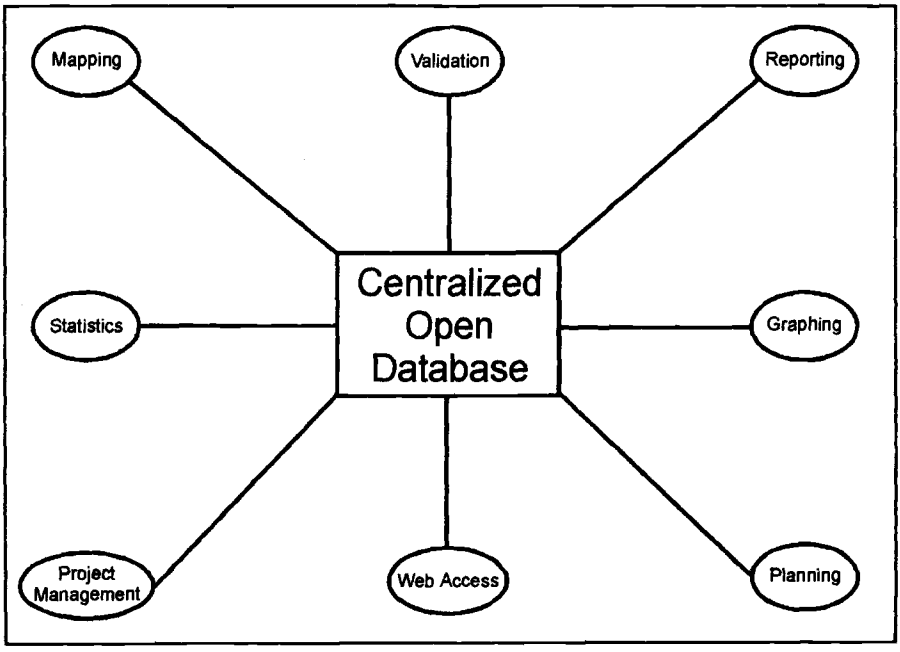


Figure 3 - *Connection to a Centralized Open Database.*

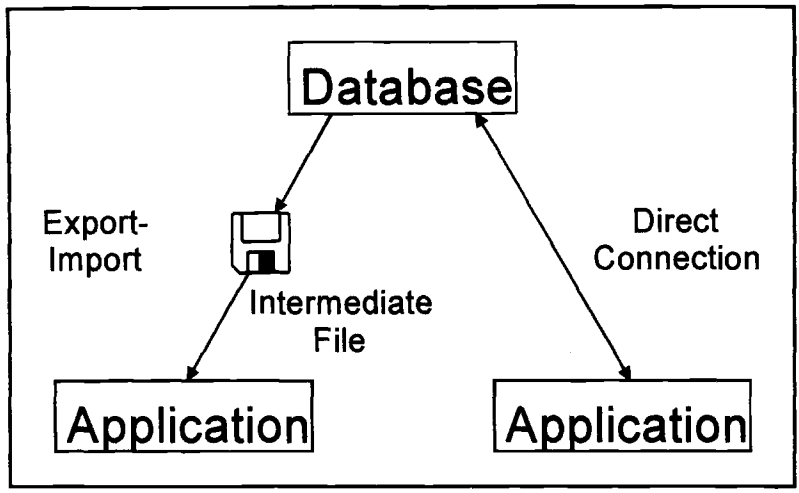


Figure 4 - *Connection Methods.*

Web-based - Tools are now available to provide access to data using a Web browser interface rather than a local application. This would be a Web-based system, also called a "thin client" system. For example, it is not difficult to provide a public or private page with environmental data for a facility. Tools like Dynamic HTML, Active Server Pages, and Java applets are making it much easier to provide an interactive user interface hosted in a Web browser. On the GIS side, a GIS application such as ArcView IMS[2] running on a Web server, can dish out maps displayed in a browser. Operations such as pan and zoom are performed by messages passed from the browser to the server, which regenerates the map image and sends it back to the browser.

Connecting Systems

Once the decisions have been made regarding database locations, the next step is to connect the systems and applications together. Often there are multiple connection requirements (Figure 2). This example shows a variety of possible data locations and connection needs for field and laboratory data for a typical environmental project. The multiple data locations and formats make it difficult to know where and how to connect. There are a number of benefits to moving as much of the data as possible into a single, centralized, open database (Figure 3). Some of the benefits of a common database include simplification of connections, elimination of problems due to multiple incompatible data formats, and minimization of redundancy. Other benefits include reduction in data entry, data entry validation, error checking, and so on. Once the data has been moved into a centralized location, then the various applications can be connected to the data.

[2] ArcView and Arc IMS are registered trademarks of Environmental Systems Research Institute, Inc.

Well	Elev	X	Y	SampDate	Sampler	Param	Value	Flag
B-1	725	1050	681	2/3/96	JLG	As	.05	not det
B-1	725	1050	681	2/3/96	JLG	pH	6.8	
B-1	725	1050	681	5/8/96	DWR	As	.05	not det
B-1	725	1050	681	5/8/96	DWR	Cl	.05	not det
B-1	725	1050	681	5/8/96	DWR	pH	6.7	
B-2	706	342	880	11/4/95	JAM	As	3.7	detected
B-2	706	342	880	11/4/95	JAM	Cl	9.1	detected
B-2	706	342	880	11/4/95	JAM	pH	5.2	
B-2	706	342	880	2/3/96	JLG	As	2.1	detected
B-2	706	342	880	2/3/96	JLG	Cl	8.4	detected
B-2	706	342	880	2/3/96	JLG	pH	5.3	
B-2	706	342	880	5/8/96	DWR	As	1.4	detected
B-2	706	342	880	5/8/96	DWR	Cl	7.2	detected
B-2	706	342	880	5/8/96	DWR	pH	5.8	
B-3	714	785	1101	2/3/96	JLG	As	.05	not det
B-3	714	785	1101	2/3/96	JLG	pH	8.1	
B-3	714	785	1101	5/8/96	CRS	As	.05	not det
B-3	714	785	1101	5/8/96	CRS	Cl	.05	not det
B-3	714	785	1101	5/8/96	CRS	pH	7.9	

Figure 5 - *ASCII File of Environmental Data.*

Stations

Well	Elev	X	Y
B-1	725	1050	681
B-2	706	342	880
B-3	714	785	1101

Samples

Well	SampDate	Sampler
B-1	2/3/96	JLG
B-1	5/8/96	DWR
B-2	11/4/95	JAM
B-2	2/3/96	JLG
B-2	5/8/96	DWR
B-3	2/3/96	JLG
B-3	5/8/96	CRS

Analyses

Well	SampDate	Param	Value	Flag
B-1	2/3/96	As	.05	not det
B-1	2/3/96	pH	6.8	
B-1	5/8/96	As	.05	not det
B-1	5/8/96	Cl	.05	not det
B-1	5/8/96	pH	6.7	
B-2	11/4/95	As	3.7	detected
B-2	11/4/95	Cl	9.1	detected
B-2	11/4/95	pH	5.2	
B-2	2/3/96	As	2.1	detected
B-2	2/3/96	Cl	8.4	detected
B-2	2/3/96	pH	5.3	
B-2	5/8/96	As	1.4	detected
B-2	5/8/96	Cl	7.2	detected
B-2	5/8/96	PH	5.8	
B-3	2/3/96	As	.05	not det
B-3	2/3/96	PH	8.1	
B-3	5/8/96	As	.05	not det
B-3	5/8/96	Cl	.05	not det
B-3	5/8/96	PH	7.9	

Figure 6 - *Normalized Environmental Data.*

Several issues are involved in the connection process. The first is the physical connection between the computers. This has been accomplished in most organizations using Ethernet-based connections (with twisted-pair or fiber-optic cable) running

standard protocols, usually TCP-IP (for the Internet), sometimes along with NetBEUI[3] (for Windows networking) and IPX/SPX[4] (for Netware).

A more challenging issue is integrating the applications needing the data (such as the GIS) with the database where the data is stored. This can be done either by moving data between applications using intermediate files, or by directly connecting the applications (Figure 4).

Export-Import Formats

In situations where data must be moved between applications or locations using the export-import approach, there are a variety of data access methods and file standards. Older formats like ASCII tab-delimited files, spreadsheets, and dBase files are giving way to new standards like XML (eXtensible Markup Language).

In an ASCII (American Standard Code for Information Interchange) file the data is represented with no formatting information (Figure 5). In the case of transferring laboratory data, the usual file structure has the disadvantage that the data is de-normalized, that is, the hierarchical (parent-child) relationships are not represented in the file. Breaking the data into separate tables to represent

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!--Hierarchical Environmental Data -->
<Stations>
  <StationName>MW-1</StationName>
    <Samples>
      <Sample SampleDate = "05/07/2000">
        <Analysis Parameter="Arsenic">
          <Value>23.5</Value>
          <Flag>V</Flag>
          <DetLim>10</DetLim>
        </Analysis>
        <Analysis Parameter="Lead">
          <Value>10</Value>
          <Flag>U</Flag>
          <DetLim>10</DetLim>
        </Analysis>
        <Analysis Parameter="Benzene">
          <Value>18.5</Value>
          <Flag>V</Flag>
          <DetLim>10</DetLim>
        </Analysis>
      </Sample>
      <Sample SampleDate = "08/04/2000">
        <Analysis Parameter="Arsenic">
          <Value>10</Value>
          <Flag>U</Flag>
          <DetLim>10</DetLim>
        </Analysis>
      </Sample>
    </Samples>
</Stations>
```

Figure 7 - *XML File of Environmental Data.*

Analyses

Station Name / Sample Date	Parameter	Value	Flag	Det.Lim.
MW-1				
05/07/2000				
	Arsenic	23.5	V	10
	Lead	10	U	10
	Benzene	18.5	V	10
05/07/2000				
	Arsenic	10	U	10

Figure 8 - *Rendered XML File of Environmental Data.*

[3] Microsoft, Windows, NetBEUI, Excel, Access, SQL Server and FoxPro are trademarks or registered trademarks of Microsoft Corporation.
[4] IPX, IPX/SPX, and Netware are trademarks or registered trademarks of Novell, Inc.

this structure is called "normalization" (Figure 6) [*10,11,12,13*].

The advantage of a newer format like XML is that it can communicate both the data itself and the hierarchical structure of the data in one file (Figure 7) [*14*]. In this file, "tags" are used to define data elements in the file, and the positioning of the tags and the data define the hierarchy. Using an advanced format like XML allows the data to be rendered (displayed) in a more efficient and understandable way (Figure 8). This rendering can be done in a flexible way using style sheets, which define how each data element in the file is to be displayed. Style sheets used to render XML data use a language called XSL (extensible style language). There are many benefits to separation of the data from how it is displayed, allowing the data to be displayed in different ways in different situations, such as a browser versus a portable phone.

The "extensible" part of XML and XSL is important. Because the language is extensible, features can be added to handle specific data requirements. For example, it is possible to add extensions to XML to handle the spatial aspects of the data. GML (Geography Markup Language) is an example of this approach. GML schema (data structures) define how geographic information is put into XML. The GML tags describe content such as the geographic coordinates and properties of a map element. A style sheet can then define the appearance of the map elements. The separation of the data from how it is displayed, in this case, might allow different scale-dependent displays depending on the resolution of the output device.

Two other common intermediate formats are spreadsheet files and database files. A benefit to these formats is that most programs can read and write one or more of these formats. In both cases, it is necessary that the application creating the file and the one accessing it agree on the program format and version of the file. Spreadsheet files, nowadays usually meaning Microsoft Excel, have the advantage that they can easily be edited. Since spreadsheet files are usually loaded into memory for processing, they usually have file size limits that can be a significant problem. For example, Excel 97 has a limit of 65535 rows, which, for an investigation of a site with 50 organic constituents and 100 wells, would limit the file to 13 quarters of data. This capacity problem severely limits spreadsheet storage of environmental data.

Database files don't have the file size limitation of spreadsheets. The two most common formats are dBase[5] (and the very similar FoxPro format) and Microsoft Access. dBase is an old format and may not be supported by applications in the future. Access is a modern, flexible, and widely supported format, but the file structure changes every two or three years, so versions can be a problem. Overall, versions of software and file structures are a common problem, and choosing the software and version is a design consideration that should be considered.

Export-Import Advantages and Disadvantages

All export-import approaches have inherent advantages and disadvantages. The primary advantage is in a situation where there is a clear "hand-off" of the data, such as when a laboratory delivers data to the client. Once the laboratory delivers the data, the responsibility for managing the data rests with the client, and any connection back to

[5] dBase is a registered trademark of Borland International, Incorporated.

the LIMS (Laboratory Information Management System) that created the electronic deliverable would be inappropriate. Transfer of the intermediate file breaks this connection, enforcing the hand-off. Another advantage is where the data must be made available even though no direct connection is available, such as at a remote location.

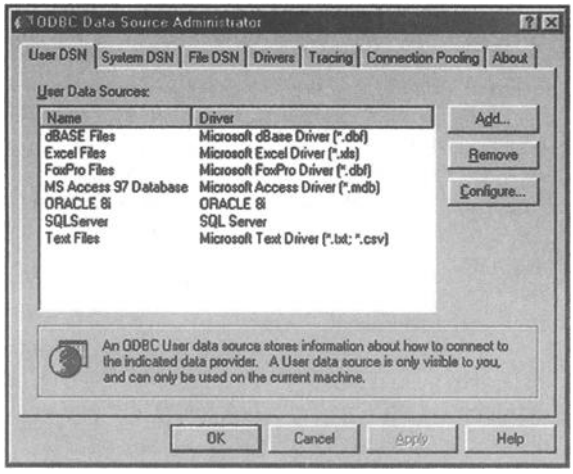


Figure 9 - *ODBC Data Source Configuration.*

There are several disadvantages to the export-import model. One is that it is necessary to define formats that are common to the programs on either side of the process, which can be difficult to do initially and to maintain. A more severe disadvantage is the proliferation of multiple copies of the data, which wastes space and can present a management problem. The biggest disadvantage, however, is lack of concurrency. Once the export has been performed, the data can become "stale". If the data in the original location is corrected, that correction is not automatically reflected in the intermediate file. The same is true of updates. With the export-import model, great care must be taken to minimize the chance of using incorrect data.

Direct Connection

In many cases, direct connection through ODBC[6] (Open DataBase Connectivity) or other protocols can eliminate the step of exporting and importing. These tools can help facilitate movement of data between the database and applications, such as a GIS, that can use the data.

ODBC is the most widely used method at this time for connecting applications to data, especially in the client-server setting. There are several parts to setting up ODBC communication. For a client-server system (as opposed to a stand-alone system, where ODBC can also be used), the first step is to set up the database on the server.

Typically this is a powerful back-end database such as Oracle[7] or SQL Server. The next step is to set up the ODBC connection on the client computer using the ODBC Data Source Administrator (Figure 9), which is part of Windows. Then the application needs to be configured to use the ODBC connection.

One example (Figure 10) shows a database client being connected to a server database in Microsoft SQL Server. This screen also shows the option to connect to an Access database.

---

[6] ODBC is a software standard developed by Microsoft Corp.

[7] Oracle is a registered trademark of Oracle Corporation.

Another example (Figure 11) shows a GIS program being connected to a database through ODBC. In this example the GIS program (Enviro Data[8]) understands the normalized environmental data model, and once the attachments are made, the SQL language in the GIS can join the tables to use the data. The Feature_Lines and Feature_Text tables, which contain base map data, will be local to the GIS, while Sites, Stations, Samples and Analyses are having their attachment changed from the EnvDHyb1 to the MyNuData data source.

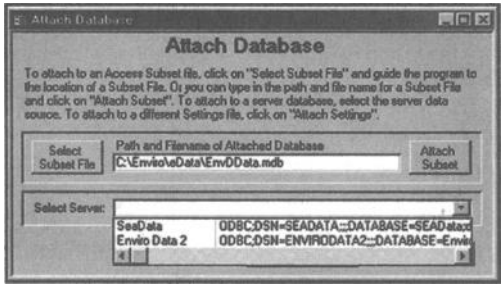


Figure 10 - *Database Attachment Screen.*

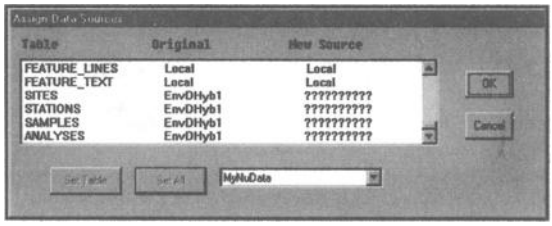


Figure 11 - *Enviro Data GIS Attachment Screen.*

In another example (Figure 12) a different GIS program, ArcView, is being attached to tables via an ODBC data source. Here the attachment will result in a denormalized table, which will be used for display.

Other connection protocols are also available, and appropriate for some situations. COM[9] (Component Object Model), DCOM (Distributed Component Object Model), CORBA[10] (Common Object Request Broker Architecture), EJB[11] (Enterprise Java Beans) and more recently SOAP[12] (Simple Online Access Protocol), are usually used

Figure 12 - *ArcView GIS Attachment Screen.*

[8] Enviro Data and Spase are registered trademarks of Geotech Computer Systems, Inc.
[9] COM and DCOM are standards developed by Microsoft, Inc.
[10] CORBA is a standard maintained by the Object Management Group.
[11] EJB and Java are trademarks of Sun Microsystems
[12] SOAP is a standard maintained by the World Wide Web Consortium.

with Internet data communication. COM, DCOM and SOAP are popular for use with Microsoft products, and CORBA and EJB are popular with the anti-Microsoft camp, especially with advocates of Java (a programming language from Sun Microsystems) and Linux[13] (an open source variant of the Unix[14] operating system).

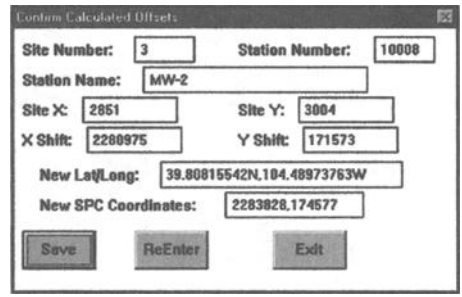


Figure 13 - *Multiple Coordinate Systems.*

A direct connection method specific to geographic data is GTTP (Geographic Text Transfer Protocol). This is a transfer protocol at the same level as HTTP (HyperText Transfer Protocol), the protocol on which the bulk of the World Wide Web is based. GTTP is a protocol to request and receive geographic data across the web. In order for it to work, there must be an application running on the Web server that can accept GTTP requests and then send out the requested geographic data. A thin client such as a Java applet running in a browser would send a request, receive the data, and create the map display. In the implementation of this protocol by Global Geomatics[15], the server application can combine raster and vector map data, convert it to a common projection, and send it to the client application for display.

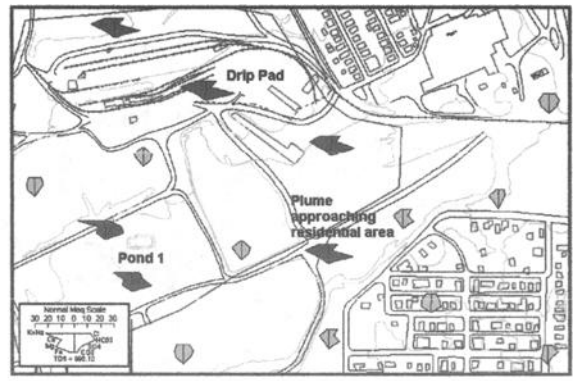
Data Integration

Integration of spatial and non-spatial data raises special issues. Data for different features or from different sources may be in different coordinate systems [*15,16*] or of different accuracies and resolutions. Displaying this data together in a common coordinate system requires adequate information for each data source, and software that can handle this information.

Environmental projects can be particularly troublesome in this area, since for some sites some of the data might be in a local coordinate system, some in a global Cartesian system in one or more projections and scales, and some in latitude-longitude. In some cases it is necessary to maintain three separate coordinate systems, but yet be able to display all of the data on one map (Figure 13). In this example, the user has chosen a station (well) for which the site coordinates are known, selected the location of the well from the map, and the software has calculated the offsets between real-world XY (state plane) and site coordinates. Tying non-spatial data to map locations can be a difficult process on many projects, and items like defining offsets for posting can be dependent on both the spatial and non-spatial data.

[13] Linux is a trademark of Linus Torvalds.
[14] UNIX is a registered trademark of The Open Group.
[15] Global Geomatics is a trademark of Global Geomatics, Inc.

Benefits of Integration



The benefits to be gained by overcoming data management and communication obstacles can, in many cases, greatly outweigh the effort expended. If the end justifies the means, the displays that are generated can provide a much greater understanding of site technical and administrative issues. The benefits to an integrated system fall into at least three areas: increased efficiency, improved quality, and enhanced output.

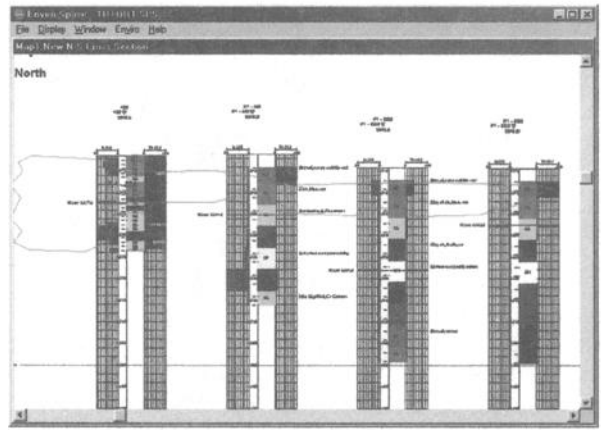


Figure 15 - *Cross Section from Relational Data.*

Increased Efficiency

The increased efficiency comes from several sources. One is elimination of redundant handling of the data. In the past it was not unusual to take data that was printed out from one computer system and type it into another to use it. In many cases this was done multiple times for the same data. A streamlined data management process eliminates this tedious work. Operating a centralized database where everyone knows where the data is and what level of review (verification, validation, etc.) has been applied to that data can be a great time saver. When someone needs the data, they can have access to it without wasting time. Finally, when new uses arise for the data, such as a new requirement for doing a historical comparison, no time is wasted organizing the data, and the effort can be expended on the analysis itself. Over the life of a complex project, these time savings can be in the thousands of hours.

Improved Quality

Re-typing data as described above not only wastes time, it allows errors to creep in. Managing a system where data is moved around in multiple intermediate files also provides many opportunities for choosing the wrong file, using stale data, and other

things that result in errors that can be very hard to notice. Centralizing the data, and setting up applications to directly access it, can reduce this type of error to an absolute minimum.

Another area where an integrated, centralized database can contribute to improved quality is in consistent presentation of data. Routines can be developed that properly format data with correct handling of significant digits, non-detected data, and flagged data, and these routines used for all data output and display. The result is data where the quality is maximized and fully understood.

Enhanced Output

It's not unusual for a large site to spend hundreds of thousands or even millions of dollars per year on sampling and analysis. Project managers and regulators are now realizing that the data gathered through this process can have value beyond simple reporting of sample events. Getting the maximum information out of the data makes good management sense. A big part of this is reports and displays generated by GIS and other software. Once it is easy to create sophisticated displays with reliable and current data, project personnel can take the time that in the past was wasted on inefficient data manipulation processes and use it to better understand the site so that it can be managed better.

Some examples highlight some displays that are facilitated by using an open, centralized database, and that can lead to a better understanding of site conditions.

In the first example (Figure 14), the GIS has been attached to the relational data model in the database. The user then selected a sampling event by date, and the GIS then drew Stiff water quality diagrams on the maps next to the wells. The user can then use the diagrams to better understand the configuration of the groundwater chemistry at the site.

In the second example (Figure 15), the user has used the GIS to select wells for a cross section from a map display, and defined what data is to be displayed and how it will look. Then the GIS retrieved the data from the centralized database to create the display. This process took just a few minutes, compared to hours or days to generate a cross section like this by manual drafting.

The third example (Figure 16) shows a combination of time sequence graphs and a map display. The user selected wells for display and date ranges for the graphs, and the GIS created the graphs, including control chart limits and outlier displays, in just a couple of minutes.

The above examples have one thing in common: adding dimensions to the displays to increase understanding. The graph display adds the time dimension to the usual X-Y display of the map. The cross section employs the vertical axis along with a horizontal dimension. And the Stiff diagram [*17*] display combines eight dimensions of chemical concentration to the two spatial dimensions.

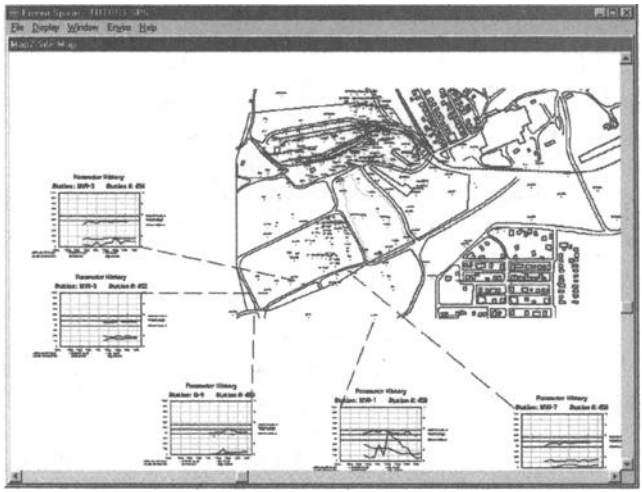


Figure 18 – *Graphs from Relational Data on a Map.*

The other point to be made from these examples involves efficiency and time savings. Tools were available in the past to create displays like those shown above. The problem was that it took too much time to use the older tools except on an occasional basis. By making the data easily available, and providing efficient tools to create information-rich displays, project staff can look at the data in enough different ways that they can gain a greater understanding of the site. And they can communicate that understanding. If the displays can make project staff and regulators confident that the remediation or monitoring program is working, then maybe some of the wells can be sampled at greater intervals. The savings of time and money that this represents over the next few years can easily cover the time and expense of implementing an integrated system. As the focus of site management moves towards more efficient operations, it's a win for everyone.

We live in a multi-dimensional world, and the answers to questions about our sites are often contained in data we already have. The purpose of a database-GIS combination is to make it easy to find and display the data to answer those questions.

References

[1] Feng, P. and Easley, M., 1999, "Managing environmental data," *Environmental Protection*, April, 1999.

[2] Edwards, P. and Mills, P., 2000, "The state of environmental data management," *Environmental Testing & Analysis*, September/October, 2000, pp. 22-30

[3] Giles, J. R. A., 1995, *Geological Data Management*, The Geological Society, London.

[4] Harmancioglu, N. B., Necdet Alpasian, M., Ozkul, S. D., and Singh, V. P., 1997, *Integrated Approach to Environmental Data Management Systems*, Kluwer Academic Publishers, Dordrecht. Proceedings of the NATO Advanced Research Workshop on Integrated Approach to Environmental Data Management Systems, Bornova, Izmir, Turkey, September, 1996.

[5] Rich, D. W., 2002, *Relational Management and Display of Site Environmental Data*, Lewis Publishers, Boca Raton, FL.

[6] Carter, G. C. and Diamondstone, B. I., 1990, *Directions for Internationally Compatible Environmental Data*, Hemisphere Publishing Corporation, New York.

[7] Barnwell, C. E., 2000, The USGS GeoExplorer Project: using new GIS technology for scientific data publishing, in *Geographic Information Systems in Petroleum Exploration and Development*, Coburn, T., C., and Yarus, J. M., Eds. AAPG Computer Applications in Geology, No. 4, pp. 249-260.

[8] URISA, 1998, *GIS Database Concepts*, The Urban and Regional Information Systems Association, Park Ridge, IL.

[9] Goodchild, M. F., Parks, B. O., and Steyaert, L. T., 1993, *Environmental Modeling with GIS*, Oxford University Press, New York.

[10] Codd, E., 1970, A Relational Model of data for large shared data banks, Communications in the ACM, June, 1970.

[11] Codd, E., 1972, Further normalization of the data base Relational Model; in Data Base Systems, Courant Computer Science Symposia Series, v 6, Prentice Hall, Englewood Cliffs, NJ.

[12] Stonebraker, M. and Hellerstein, J. L., 1998, Readings in Database Systems, Morgan Kaufmann Publishers, San Francisco, CA.

[13] Walls, M. D., 1999, Data Modeling, The Urban and Regional Information Systems Association, 1460 Renaissance Drive, Suite 305, Park Ridge, IL 60068.

[14] Dragan, R. V., 2001, XML your data, *PC Magazine*, June 26, 2001.

[15] Snyder, J. P., 1987, *Map Projections - A Working Manual*, U.S. Geological Survey Professional Paper 1395, U.S. Government Printing Office, Washington, DC.

[16] Tufte, E. R., 1983, The Visual Display of Quantitative Information, Graphics Press, Cheshire, CT.

[17] Hem, J. D., 1985, Study and Interpretation of the Chemical Characteristics of Natural Water, U.S. Geological Survey Water-Supply Paper 2254, 3rd ed.

Arthur F. Lange[1] and Rosalind Buick[2]

Differential GPS Update

Reference: Lange, A. F. and Buick, R., "**Differential GPS Update,**" *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy and Standardization, ASTM STP 1420*, D. T. Hansen, V. H. Singhroy, R. R. Pierce, and A. I. Johnson, Eds., ASTM International, West Conshohocken, PA, 2003.

Abstract: In May 2000 the U. S. Department of Defense eliminated the purposeful degradation of accuracy of the Global Positioning System (GPS) called Selective Availability (SA). This action resulted in the accuracy of GPS improving from around 50 m to better than 10 m. With this improvement in accuracy civilians are enjoying very good performance from low-cost handheld GPS receivers, although GPS is still not accurate enough for many professional uses. The limitation of accuracy of GPS is caused by a number of system errors. Some system-wide correctable errors of GPS are reduced through the use of differential GPS techniques. The elimination of SA will allow the optimization of Wide Area DGPS networks because of the change in characteristics of the remaining error sources.

Keywords: global positioning system, GPS, differential GPS, DGPS, wide area DGPS, WADGPS, differential corrections

GPS Error Sources

With the removal of Selective Availability (SA), the purposeful degradation of GPS in May 2000 by a presidential order, there was a sudden improvement in the accuracy of GPS. Civilians noticed an improvement from 50 m to 10 m with their low-cost handheld GPS receivers. After SA was stopped, you might wonder why a user might only obtain 10 m accuracy and might wonder what is necessary to obtain better accuracy. The answer is that GPS has a number of ranging errors. By understanding the

[1] Product Manager, Trimble Navigation Ltd., 645 North Mary Avenue, Sunnyvale CA 94086.

[2] Business Development, Agriculture, Trimble Navigation Ltd., 545 North Mary Avenue, Sunnyvale CA 94086.

magnitude and time scale of these ranging errors, we can devise methods to take advantage of the accuracy inherent in GPS and get position fixes to the maximum accuracy possible. Besides the ranging errors, the accuracy of GPS is affected by the satellite geometry which is represented by the geometric dilution of position (GDOP).

GPS Ranging Errors

There are six classes of ranging errors: ephemeris data, satellite clock, ionosphere, troposphere, multipath, and receiver (Parkinson et al. 1996). Table 1 summarizes the magnitude of these errors. All errors in this paper are expressed as one-sigma, unless specified otherwise.

Table 1 – *GPS Error Sources*

Error Source	One-sigma error, m	Time constant, s
Ephemeris data	2.1	1800
Satellite Clock no SA	2.1	8 h
Satellite Clock with SA	20.0	30
Ionosphere	4.0	6 h
Troposphere	0.7	
Multipath	1.4	
Receiver	0.5	

Ephemeris data - These errors occur when the GPS message for a satellite location is not correct. Ephemeris data are predictions for the satellite and errors grow slowly with time from the last upload. Ephemeris errors average 2.1 m.

Satellite clock - Clock errors are minimized by the use of atomic clocks in the satellites with a stability of 1 part in 10^{13} over a day. This stability results in average errors of 1-2 m for 12 hour updates between uploads, growing with time quadratically. The clock errors are shown for both with and without Selective Availability (SA.) The satellite clock errors, under SA were increased by clock dither to induce range errors on the order of 20 m and were the dominant source of GPS error.

Ionosphere - Free electrons in the ionosphere delay the GPS signals and are the source of ionosphere errors. GPS receivers use a simple model to correct for ionosphere errors. The parameters for the ionosphere model are transmitted in the GPS message and are in the order of 2-5 m in temperate zones. Near the magnetic equator and poles the errors are much larger.

Troposphere errors - The troposphere causes an additional delay to the GPS signals. Variations in temperature, pressure and humidity contribute to variations in the speed of light. The simple model used by GPS receivers is effective to about 1 m.

Multipath errors– An error occurs in a GPS receiver when a reflected signal arrives with a sight delay from the direct signal, masking the correlation peak in the GPS receiver. These errors are greatest near large reflecting surfaces, like buildings and can be 15 m or more in extreme cases. GPS receivers have special algorithms to reduce the impact of multipath errors. Special GPS antennas, often with ground planes or choke

rings are used to reduce the reception of low angle signals for high accuracy GPS measurements. There are some urban canyon environments that prevent accurate GPS positioning. The combination of large reflecting surfaces with the blocking of the direct path to the satellite from the buildings causes limited availability and large errors.

Receiver errors – GPS receivers can be built which contribute less than 0.2 m in noise and 0.5 m in bias. Since the designer of the GPS receiver may choose to make compromises to reduce cost at the expense of increased errors, not all GPS receivers are built for low noise.

Geometric Dilution

The geometric dilution can be calculated for any satellite configuration and is displayed in the user's GPS receiver as position dilution of precision (PDOP). Typical useful values of PDOP range from 1 to 8. The estimated position error is equal to the geometric dilution times the ranging error. PDOP is composed of two orthogonal components: the horizontal dilution of position (HDOP) and the vertical dilution of position (VDOP).

Dominant GPS Errors

The dominant SA induced error is the fast clock errors, the dominant non-SA error is caused by errors in the ionosphere model. At 4.0 m of range errors the resulting position error with a HDOP of 2.0 is 10.2 m and with a VDOP of 2.5 is 12.8 m. Greater errors are caused by unfavorable geometry of the satellite constellation used by the receiver. An unfavorable geometry induces errors roughly in proportion to the HDOP or VDOP. A poor constellation geometry can cause large errors, for example, if the HDOP were 20 because some of the satellites were blocked by an obstruction, then the HDOP induced position error would be approximately 100 m. If the HDOPs are low, for example less than 4, then a user with a high quality GPS receiver will see only a small position drift with time. This drift is caused by the slowly changing ionosphere and ephemeris errors.

Correctable GPS vs. Un-correctable Errors

To obtain sub-meter accuracy, which is required for many GIS data collection projects, Differential GPS (DGPS) is used to counteract the correctable errors. DGPS removes the Ranging errors that are correctable from a GPS position calculation by using differential correction messages. A differential correction message is created in a differential base station, and sent to users in the vicinity of the base station using a radio link.

GPS errors common to an area can be minimized with the use of a differential GPS reverence station. The errors correctable with differential GPS are the satellite ephemeris errors, satellite clock errors, and atmospheric errors (ionosphere and troposphere errors). Errors that are localized to a single receiver are not correctable; these errors are receiver errors, constellation geometry errors and multipath errors.

Differential GPS Base Station

A DGPS base station consists of three important parts: a GPS antenna at a known location, a specially constructed high-quality GPS reference receiver, and a method for getting the differential correction message to the user. For a real-time DGPS base station, a radio link is used, and for a post-processed DGPS base station, a computer file server is used with an internet or modem dial-up connection. The GPS reference stations must be constructed to minimize errors caused by receiver noise, multipath and interference, since these uncorrelated errors at the reference station are added to the broadcast differential correction message and affect all users.

The GPS reference receiver estimates the errors for each satellite, and encodes these errors and the rate of change of the errors into a message for each satellite. With no SA the rate of change of the errors is very small. At the user receiver, the satellite corrections are applied to the range measurements, which results in a more accurate position calculation.

Geographic De-correlation

As the distance increases from the GPS reference station, the accuracy of the differential correction decreases. This effect is called geographic de-correlation. This accuracy decrease is caused by differences in the ionosphere, troposphere and effect of ephemeris errors on the position calculation between the base station and the user. Satellite clock errors are not affected by the distance to the reference station, and have been almost completely eliminated by a high quality DGPS correction message with low latency. This is the main reason that users have only seen small increases in the before and after SA differential GPS accuracy from a single base station. The magnitude of the geographic de-correlation errors in DGPS is on the order of 1 m at 200 km.

Wide Area DGPS

Latency

To overcome the distance limitations of a single reference station, a network of reference stations is used to create a Wide Area DGPS (WADGPS) correction. The accuracy of WADGPS corrections is uniform over the monitored area and degrades gracefully at the perimeter. The WADGPS accuracy is dependent on the quality of the WADGPS reference stations and the algorithms used to create the WADGPS messages. With the elimination of SA, the errors related to the latency of the message through the processing at the network hub and through the distribution system have been greatly reduced.

When SA is present, an important consideration for a high accuracy DGPS correction signal provider is the minimization of the latency between the measurement of the signal errors at the reference station and the utilization of the corrections at the user receiver. To cope with the fairly large rate of change in the clock when SA is present, the latency was required to be less than 10 seconds to prevent a noticeable reduction in accuracy. (Parkinson et al.1996) High accuracy DGPS correction stations were designed to keep the latency around 5 seconds. With the elimination of the fast clock errors of SA, the requirement for a low latency DGPS message was eased somewhat. Non-SA residual errors change more slowly, on the order of minutes to hours. This non-SA change in the nature of the time course of errors has the effect of greatly reducing the bandwidth required for the same accuracy when SA was present. With an easing of the requirements for low latency correction messages, the algorithms used in a WADGPS network can be now be improved to take advantage of this change. With the same bandwidth for correction messages, improved WADGPS networks will begin to appear. The improved WADGPS networks will use improved ionosphere and ephemeris models to reduce the correctable errors and will increase the rate of these messages and reduce the clock correction message rate. Users will also see improved performance for WADGPS systems with dual frequency user receivers because of their ability to create more accurate model for the ionosphere.

Ground and Satellite Distribution of DGPS Correction

Ground based radio transmitters are normally used distributing DGPS corrections from a single reference station. The best example is the U. S. Coast Guard medium frequency (MF) reference stations in the 285 – 325 kHz band. The signals from most of these reference stations can be received over 200 km during daytime and much further at night. Approximately 60% of the co-terminus U. S. (CONUS) landmass is covered by at least one MF beacon. The U. S. Coast Guard has announced plans for a complete coverage by at two stations for the entire CONUS. A completion date for this ground system depends upon U. S. congressional budget allocations.

Satellites are generally used for distributing WADGPS corrections. Recently, some experiments have been started to use the internet for distribution of real-time DGPS and WADGPS corrections, with a short distance internet to mobile user radio link at the user location. The use of the Internet for primary distribution of corrections allows much lower power transmitters and higher bandwidths to be used. Some of these

Internet differential correction experiments are directed towards carrier phase corrections for survey grade accuracy and can provide accuracy in the cm range.

Commercial and Government Providers of DGPS Corrections

In the U. S. there are currently (January 2001) three commercial providers of WADGPS corrections and one government supplier. The commercial WADGPS providers are Thales-Landstar, Fugro-Omnistar and Deere Greenstar networks. Each of these providers use a satellite L-band transponder in a geo-stationary orbit and provide corrections over the entire CONUS and parts of Alaska. The price for a commercial one-year subscription in the U. S. is approximately $800. In addition, Landstar and Omnistar signals are available over much of the world's land masses.

The U. S. Federal Aviation Administration has created a system, which will be supplied at no cost to users, called Wide Area Augmentation System (WAAS), with an equivalent European (EGNOS) and Japanese (MCAC) system. The purpose of WAAS is to increase the integrity and availability of GPS for aviation users. WAAS has the capability to add additional ranging signals besides providing WADGPS corrections. These additional ranging signals may help provide an additional satellite range signal to help reduce time periods with high PDOP. Because of the vital nature of the FAA's mission, an extensive system test is being performed with an estimated fully operational system of September 2002. Currently (January 2001), the FAA WAAS differential correction broadcasts are used by ground based users in non-mission critical situations. The accuracy and integrity of the WAAS signals are not guaranteed during the test period. Preliminary measurements of WAAS DGPS accuracy with high-performance user GPS receivers is 1-2 m. The FAA's WAAS published 3-sigma accuracy goal is 7 m vertical and 5 m horizontal.

Carrier Phase, Survey Grade Corrections

The previous discussion of DGPS was based on code phase GPS receivers. There is another class of GPS receiver that uses the carrier phase of the GPS signal to create a much more accurate position solution. If the carrier phase GPS receiver is stationary, a static solution is created, and if the receiver is moving, a kinematic solution is created. If real-time corrections are used, the moving solution is called Real-Time Kinematic (RTK). Static accuracy, with data collection sessions on the order of several hours, is capable of accuracy in the range of 1 mm over 1000 km base lines. RTK accuracy with 10-km base lines is on the order of 10 to 20 mm. The limitation for a RTK base line appears to be 10 km. This troposphere-induced limitation is caused by the inability of the user receiver to resolve the integer number of wavelengths after a loss of lock has occurred. To overcome this serious distance limitation for a single RTK base station, a Network RTK solution is being developed. Over the next few years, commercial Network RTK solutions will become more available. Different solutions are being

proposed with accuracy in the range of 20 mm to 200 mm. Each of the several different proposed Network RTK systems have advantages and disadvantages. Some will use satellite transmissions, some mixed Internet and short-range radio.

Optimizing GPS Receiver Performance

The user of a GPS receiver can make a trade off between accuracy and availability. For example, a mask change might be made to improve accuracy by eliminating low elevation satellites which may cause a decrease of accuracy. This change may improve accuracy. The user, however, may find that the periods of time during which a constellation with a low PDOP has been reduced, limiting GPS availability. This interaction is explained as follows. Low elevation satellites are very useful in reducing the PDOP, however low elevation satellites suffer much more from errors caused by multipath, weak signals and troposphere variations. Increasing the elevation angle mask to exclude low elevation satellites, may increase the accuracy of the position calculation, however, this increase in elevation angle mask may increase the PDOP of the remaining constellation to such a value that the resultant position calculation is worse than it would be with the inclusion of a low elevation satellite. To eliminate the use of a constellation with a high PDOP, the PDOP mask is set to an appropriate level which may limit GPS availability.

One way to understand the complex interaction between satellite elevation angle mask and PDOP at your particular location is to use a satellite visibility program, available on several Internet web sites. (Trimble) By varying the elevation angle and plotting the PDOP over a 12-hour period, the user can determine a strategy which balances trade-off between accuracy and availability for his particular situation.

GPS Receiver Standards

Many U. S. Government agencies purchase a large number of GPS receivers for many different uses. A few of these agencies have published testing standards for certification of use for their particular users. These agencies include the Federal Aviation Administration who test and certify aviation receivers, The Radio Technical Commission for Aeronautics, the Federal Geodetic Control Subcommittee (FGCS) and National Geodetic Service (NGS). Although the Bureau of Land Management (BLM) and U. S. Forest Service (USFS) purchase many GPS receivers for GIS data collection, and have devised testing procedures for their own use, they have not formally published their testing procedures or certify receivers. The BLM and USFS have published guidelines for Cadasteral Surveys with GPS (Sumpter et al. 2000). The USFS has published the results of its testing of a GPS receiver designed for GIS data capture (USFS GPS Information Page). These documents are of great help for potential users in evaluating the specifications of a receiver for a particular use.

Standards are lacking for characterizing the accuracy of a GPS receiver that manufacturers use to describe their receiver specifications. This makes it difficult to compare products from different manufacturers. For instance, horizontal accuracy can be specified CEP, circular error probable, representing 50% of the positions within a circle of the specified radius; SEP, spherical error probable, 50% of the positions within a sphere of the specified radius; 1dRMS, one-standard deviation, 63% of the positions within a circle of the specified radius; and 2dRMS, 95% of the positions within a circle of the specified radius. It is important to make sure that the receiver chosen has the necessary accuracy when operating in the user's environment.

References

Parkinson, B. W., Spilker, J. J., Axelrad, P., and Enge, P., 1996, *Global Positioning System: Theory and Applications,* American Institute of Aeronautics and Astronautics, Washington, DC 20024-2518.

Sumpter, C., Londe, M., Chamberlain, K., and Bays, K., 2000, "Standards and Guidelines For *CADASTRAL SURVEYS,"* U. S. Forest Service, draft document, URL: http://www.fs.fed.us/database/gps/gpsguidelines/GPS_guidelines.htm

Trimble Navigation Ltd. "GPS Mission Planning: Satview," URL: http://www.trimble.com/satview/index.htm

USFS GPS Information Page & Receiver Performance Reports, URL: http://www.fs.fed.us/database/gps/gpsusfs.htm

David T. Hansen [1]

Defining Cooperative Geospatial Projects Between Organizations

Reference: Hansen, D. T., **"Defining Cooperative Geospatial Projects Between Organizations,"** *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy, and Standardization, ASTM STP 1420*, D. T. Hansen, V. H. Singhroy, R. R. Pierce, A. I. Johnson, Eds., ASTM International, West Conshohocken, PA, 2003.

Abstract: Cooperative development of data between organizations is increasingly common. This can reduce the data development costs for individual organizations. It also affects the data development process and the resulting products. The U.S. Bureau of Reclamation and the U.S. Fish and Wildlife Service are jointly developing a habitat monitoring program for the Central Valley, California. The program is described in a work plan between the U.S. Fish and Wildlife Service and the U.S. Bureau of Reclamation. The overall intent of this program is to evaluate changes in land cover and habitat on a periodic basis. There are three phases for the program. Phase one covers the development of a habitat base map. Phase two identifies spectral change between the base year and the year 2000. In phase three, the cause of spectral change will be identified. At that time, changes in land cover and habitat over the period will be evaluated. During these phases, close coordination between these agencies is required to see that data products meet the needs of both agencies.

The base year for the program is 1993. During phase one, a uniform habitat base map for 1993 and a habitat classification system for the Central Valley have been adopted. The 1993 base map has been developed from the best existing land use and land cover data for that time period. Surrounding the Central Valley and overlapping much of the project area is a statewide change detection program by the California Department of Forestry and Fire Protection and the U. S. Forest Service. The habitat legend will be based on a vegetation classification system that is represented in this change detection project. The Central Valley habitat program will use many of the methods and some data developed by the statewide change detection project.

Including data collected under the statewide change detection program and use of an established habitat classification system greatly accelerated this program. However, not all habitats of interest to the U.S. Fish and Wildlife Service could be initially represented in the habitat base map. Developing the base map from multiple

[1]Geospatial Scientist, Soil Scientist, MPGIS, U.S. Bureau of Reclamation, 2800 Cottage Way, Sacramento CA. 95825-1898.

sources increased the amount of uncertainty in habitat labels for some areas. Addressing these issues in the data development process requires the commitment of staff. Managers and staff need to recognize that cooperative efforts will require substantial additional time for coordination during all phases of the process. With vegetation, land cover, land use data, this coordination must extend to include the statewide change detection project and other state and national efforts at standardizing classification systems.

Keywords: Central Valley of California, wildlife habitat, vegetation classification, remote sensing, change detection, spatial resolution, temporal differences, uncertainty analysis, cooperative geospatial data development, geospatial standards, land cover, land use.

Introduction

Data development remains one of the most expensive aspects of geographic information systems (GIS). Increasingly, organizations are cooperatively developing data sets based on common interests or requirements. There are efforts at the regional, state, and national levels to standardize characteristics for basic or framework data. Cooperation and integration in data development can reduce costs to individual agencies. Coordination efforts affect the entire data development process. Coordinators must consider effects on the actual mapping area, on the mapping or digital capture process, on the database structure, and on the classification system. The organizations involved need a clear understanding of their expectations for the GIS data. They need to evaluate benefits based on the original intent for the data. For an effective data development, they need to understand staff and time requirements for coordination.

Habitat Monitoring Program

The U. S. Fish and Wildlife Service (USFWS) and the U.S. Bureau of Reclamation (USBR) have been involved since January 2000 in a joint program to map and monitor change in habitat for the Central Valley of California (Fig. 1). The overall project area covers approximately 12.5 million hectares (31 million acres). The overall objective for the organizations is to develop a tool to monitor habitat and land cover change. It is divided into three phases described in a work plan for the project (U. S. Bureau of Reclamation and U. S. Fish and Wildlife Service 2000). The first phase is to develop a land cover base map for 1993. This is a key year for the two organizations in evaluating changes in land cover and habitat. This is the period when the Central Valley Improvement Act came into effect. The program will meet part of the requirements for a biological assessment of the Central Valley Project. The second phase is the development of a change detection layer between 1993 and 2000.

This phase identifies change based on spectral differences of satellite scenes between these two dates. Change detected will be due to a variety of factors many of which are not related to change in habitat. The third phase consists of actual identification, evaluation, and labeling of change. During this phase, the initial habitat legend and 1993 base layer are evaluated and updated with additional information. While three distinct phases are recognized for processing purposes, these phases overlap. The overall objectives and phases are described in a jointly developed work plan (U. S. Bureau of Reclamation and U. S. Fish and Wildlife Service 2000). Figure 1 shows the project area within the state.

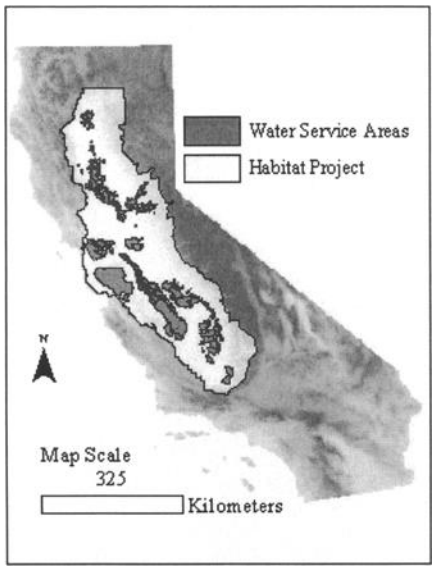


FIG. 1— *Project area for the 1993 habitat database.*

The USFWS is concerned with habitat change that may affect listed species within the entire program area. USBR is chiefly concerned with about 2.1 million hectares (5 million acres) scattered throughout the area (Fig. 2). These are water contract service areas and represent about 18 percent of the project area. Habitat mapping of these areas will meet requirements under the record of decision by the USFWS. The project area overlaps a long term monitoring program by the California Department of Forestry and Fire Protection (CDF) and the U.S. Forest Service (USFS). This CDF-USFS program monitors change in vegetation and habitat on a five-year cycle (Levien et al. 1998). To effectively map the entire area to meet the needs of USFWS, USBR will incorporate methods and data developed by CDF - USFS. For the 1993 habitat base map, about 10.1 million hectares (24.9 million acres) are being processed by USBR. Habitat for the remaining 20 percent will incorporate mapping by CDF and

USFS based on 1994 imagery. In phase 2 and phase 3, only about 4.8 million hectares (12 million acres) will be fully processed for habitat change. About 38 percent of the area will directly incorporate the CDF and USFS change detection data (Fig. 2).

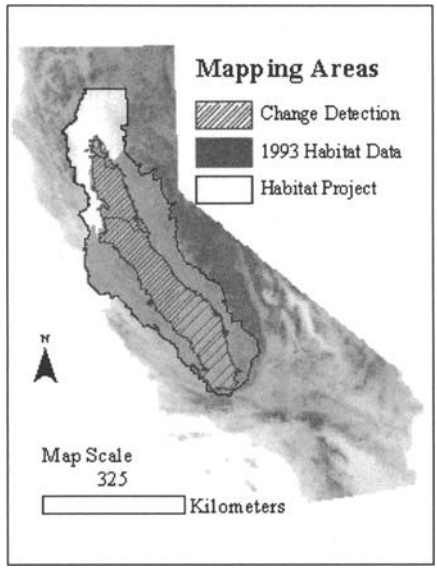


FIG. 2 — *Change detection area within the habitat project area.*

Linking with this program provides substantial savings in the time required to develop the initial base map and for monitoring change over the long term. Following established methods and merging with other agency products reduces the cost for data development over the entire area. This should improve the quality of the products developed. Data integration with the CDF and USFS change detection project also affects other aspects of this project. It places some constraints on the classification system used for habitat, the areas evaluated, and processes followed by USFWS and USBR. In addition, other State, and National efforts are underway to standardize classification systems for vegetation, land cover, and land use. Staff must be cognizant of and involved with these efforts. They must be able to communicate the effect of these combined efforts on areas that are of concern to each agency to their managers.

Phase I – Development of 1993 Habitat Base Map

Neither agency has a habitat base map for 1993 or a defined and active habitat classification system. The base map for 1993 has been constructed from existing

digital data sets. To integrate these separate data sets, a uniform base map was developed from spectral polygons of 30 meter satellite scenes for 1993. Spectral polygons were generated with a minimum size delineation of 1.0 hectares (2.5 acres) using public domain software developed by Boston University (Ryherd and Woodcock 1990). Figure 3 shows a group of spectral polygons for a processing area. The spectral polygons are out lined in white with a gray scale image of the satellite scene. These spectrally similar polygons are generally about 1 to 5 hectares in size. Groups of spectral polygons effectively block out similar natural features.

FIG. 3 — *Spectral polygons generated for a portion of the project area.*

Six different digital data layers were selected to provide labels for the spectral polygons. These data layers were constructed by a variety of methods to map vegetation, land cover, or land use. These data sets represent different types of land cover and land use mapping. Classification systems are similar but not identical between the data sets. They represent different time periods around 1993. These separate sources have other characteristics affecting their usefulness in identifying habitat for the project.

Primary Data Sources

DU - This is data developed by Ducks Unlimited and others for the California Department of Fish and Game and others (Ducks Unlimited 1997). It is based on image processing of satellite scenes for the winter - summer period of 1992 to 1993.

The minimum mapping unit for this data is 1.0–2.0 hectares (2.5 to 5 acres). It covers most of the area on the Central Valley floor that will be processed for spectral change. The chief focus of this mapping is identifying seasonal and non-seasonal wetlands. For the 1993 base map, it is relied on to identify water, wetland types and some riparian types.

DWR - This is crop and land use data developed by the California Department of Water Resources (DWR 1993). Mapping covers a range of years from 1989 to 1995 depending on the county involved. Source scale for this data is 1:24,000. It covers only portions of the project area. Data was developed from field mapping and aerial photography. Major crop types and residential, urban, and industrial areas are identified. It generally identifies natural vegetation areas but not specific habitats.

DOC - This is mapping of farmland by the California Department of Conservation (DOC 1994). Mapping is done every two years to identify changes in agriculture land area as part of the farmland mapping and monitoring program. It has a minimum mapping unit of 4.0 hectares (10 acres) with mapping done from aerial photography at a scale of 1:24,000. It covers only portions of the project area. For the 1993 base map, this data is relied on to identify agricultural and urban areas. It does not identify any habitat types.

HDWD - This data is native hardwood mapping by Pillsbury and others in the early 1980's and subsequently updated in 1990 from satellite imagery (Pillsbury 1991). Source scale varies from close to 1:24,000–1:58,000. Besides scale issues and the broad time frame for this data, it covers only portions of the project area. It only identifies hardwood types or mixed hardwood and savannah. Where woodland types occur in multiple sources, this data is used to identify the hardwood type for the 1993 base map.

GAP - This is mapping of major vegetation or land cover categories on a statewide basis by the University of California, Santa Barbara. It is part of the National GAP program prepared for the California Department of Fish and Game (California Department of Fish & Game 1998). It is largely based on 1990 satellite scenes with some 1:58,000 scale aerial photography. The target resolution of the data is 1:100,000 with a 40.0 hectare (100 acre) minimum mapping unit. Of the six data sets, it is the coarsest resolution (Figure 4) but it has the most complete and detailed habitat legend. The wildlife habitat relationships (WHR) system from this source is used as the basis for the legend of the 1993 base map (Mayer et al. 1988).

NLCD - This is National Land Cover Data (NLCD) recently released by the U. S. Geological Survey (USGS 2000). This data is based on 1993 satellite imagery for the area and has a minimum resolution of 30 meters on the ground. Each 30 meter pixel carries a label for land cover. The land cover classification of this data is more general then the other sources. Table 1 shows the major habitat categories, WHR names, and the sources matched to these categories.

Table 1 – *Habitat Legend for Base Map based on WHR.*

Major Category	WHR	Sources
Water	Open Water	DU, DWR, DOC, MLRC
	Riverine	GAP
	Lacustrine	GAP
	Estuarine	GAP
Urban - Developed	Urban	DWR, DOC, GAP, MLRC
Barren	Barren	DU, DWR, GAP, NLCD
Deciduous Forest		NLCD
	Blue Oak Woodland	DWR, HDWD, GAP
	Valley Oak Woodland	DU, HDWD, GAP
	Valley Foothill Riparian	GAP
	Montane Hardwood	HDWD, GAP
	Montane Riparian	GAP
Evergreen Forest		NLCD
	Closed Cone Pine Cyprus	GAP
	Juniper	GAP
	Pinyon – Juniper	GAP
	Douglas Fir	GAP
	Eastside Pine	GAP
	Jeffrey Pine	GAP
	Ponderosa Pine	GAP
	Redwood	GAP
	Red Fir	GAP
	Lodgepole Pine	GAP
	Subalpine Conifer	GAP
	White Fir	GAP
Mixed Forest		NLCD
	Interior Canyon Live Oak	HDWD, GAP
	Coastal Oak Woodland	HDWD, GAP
	Blue Oak - Foothill Pine	HDWD, GAP
	Klamath Mixed Conifer	GAP
	Sierran Mixed Conifer	GAP
	Montane Hardwood Conifer	GAP
Shrub Land		NLCD
	Coastal Scrub	GAP
	Alkali Desert Scrub	GAP
	Desert Scrub	GAP
	Sagebrush	GAP
	Low Sagebrush	GAP
	Chamise Redshank Chaparral	GAP
	Mixed Chaparral	GAP
	Montane Chaparral	GAP
	Alpine Dwarf Shrub	GAP
	Bitterbrush	GAP
Non-Natural Woody	Orchards / Vineyards / Other	DU, DWR, GAP, NLCD
Herbaceous Upland	Annual Grassland	DU, DWR, HDWD, GAP, NLCD
	Perennial Grassland	GAP
Planted / Cultivated	Cropland - Agriculture	DU, DWR, DOC, GAP, MLRC
Wetland Areas	Emergent Herbaceous	DWR, NLCD
	Woody Wetlands	NLCD
	Wet Meadows	DU, DWR, GAP
	Fresh Emergent Wetlands	DWR, GAP
	Saline Emergent Wetlands	DU, GAP

Phase One - Development of the Habitat Legend

GAP provides the initial legend for the 1993 base map based on the WHR system. There are several advantages to using this existing classification system. WHR identifies general vegetation communities for California and is in common use. There are cross walks with other vegetation and land cover classifications used in the state. WHR is a classification system carried in the CDF - USFS change detection mapping. It provides a basis for linking to efforts for a national vegetation classification system (FGDC 1997).

As can be seen in Table 1, GAP represents all habitat types. As part of the process of evaluating the source legends, source legend elements are ranked on their fit into a WHR category. Some source legend elements have a clear or near match to a WHR definition and are carried forward into the analysis. Some elements in source legends do not fit any category. Those elements are dropped in the analysis. NLCD generally has a broader classification system that matches only at broader categories.

WHR categories are broad vegetation or habitat categories. Some habitat types of interest to staff biologists do not fit into the existing WHR system. It is known that two different vegetation communities are not represented in this legend. Vernal pools are one type with unique flora and fauna. They occur as complexes of seasonally wet grassland types. These complexes are often small in size and are not expected to be well represented with 30 meter satellite imagery. Riparian vegetation communities are only identified to a limited extent in WHR. They are typically long linear features, which also tend to be masked out with 30 meter resolution satellite imagery. Available digital data for both communities were not included in phase one of this project. They probably have a limited representation in the spectral polygon base. Available mapping is either limited to a few years or not available in the 1993 time period. Additional work is needed by staff to incorporate these communities into a common classification system. They will be addressed in phase three.

Phase One – Capturing Information from Source Data

The spectral polygons generated from the 1993 satellite scenes provide a new map base integrating the separate sources. Each polygon captures the most common legend value of each source. Each polygon has the code combination from six sources. DU, DWR, and DOC data represents the most detailed mapping of the six sources. GAP is of coarser resolution. This can be seen in Fig. 4 where the sources are symbolized by source resolution. The narrow white lines are from the 1:24,000 scale sources with a display width of about 20 meters on the ground. The thicker black lines are from GAP. The display width is about 100 meters on the ground. Display dimensions were set by the GIS application based on known resolution of the sources (Hansen, 1999). The spectral polygons effectively integrate general delineations of the GAP data with the finer resolution data. Spectral polygons also

eliminate differences caused by different methods used in the construction of the source map.

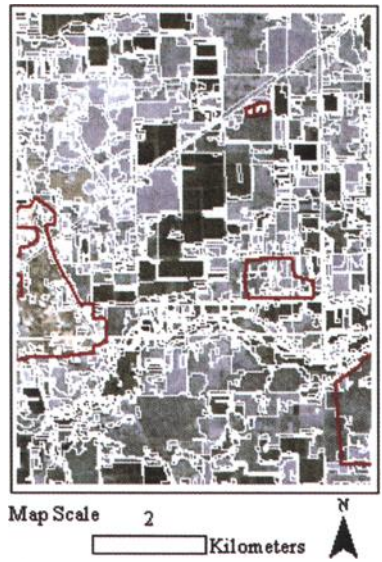


FIG. 4—*Display of detailed source data and GAP data on Orthophotoquad.*

GAP and NLCD cover the entire project area. DU data and combinations of the other data sources cover most of the area where the change detection process will be focused (Fig. 2). Initially, the 1993 base map will have an extensive area of spectral polygons depending on only GAP and NLCD. Much of this area will be replaced by change detection data from CDF - USFS in phase three. This will improve the quality of the labeling for this area. It will leave a relatively narrow band where labels will only be from GAP and NLCD.

Phase One – Assignment of Legend Values to Spectral Polygons

Characteristics of the source data provide the basis for development of general processing rules for label assignment to the spectral polygons. These rules take into account the combination of legend symbols captured from the sources and the characteristics of each source. Source characteristics include the uncertainty in assigning the source units into a habitat legend category, the time period represented by the source, and the resolution of that source. These general rules are applied and evaluated for randomly selected sets of processing areas.

Initial legend assignment is based on spectral polygons having a similar code combination. Where there is a match between three or more sources are considered to

be relatively certain. The spectral polygon carries that habitat legend code and a high certainty value. This applies to a narrow set of legend codes. Typically, these are open water, urban, wetland types, cropland - agriculture, orchard - vineyard, perennial grassland, and some woodland types.

Where only two sources agree on a habitat legend value, a variety of assignments may be made based on the sources involved and the uncertainty associated with those sources. For example, the DU source is considered to be excellent in identifying wetland types. Agreement with another source provides the basis for assigning that wetland type to the code combination. DWR and DOC are considered excellent sources for identifying agricultural and urban types. Where they agree, the code combination will carry that type. Since DOC only identifies agricultural types and not Orchard - Vineyard, the spectral polygon will be labeled with Orchard-Vineyard where the DWR source carries that value. Such code combinations carry decreased values of certainty or increasing uncertainty.

Where there is not common agreement between two sources on a specific habitat legend value but there is agreement within a broader category, the same type of rule is used. For example, where HDWD identifies a hardwood type and other sources identify a woodland type, the HDWD label is used. Where GAP identifies a specific shrub or woodland unit, and NLCD identifies shrubs or woodland species, the spectral polygon will carry the GAP label for that code combination. These code combinations will also carry a lower certainty value or higher uncertainty.

This set of rules effectively assigns an initial habitat legend value with an associated value for certainty to code combinations representing a majority of the area. These combinations are randomly selected for review and evaluation with aerial photography for 1993. It also leaves a significant number of code combinations where there is not agreement even at higher levels of the habitat legend between two or more sources. These are being reviewed for further evaluation and label assignment.

Construction of the 1993 habitat base map is a complex process. The 1993 habitat legend, evaluation of the separate sources, and the process of assigning labels need to be understood by the managers and staff of both agencies. This includes the legend value assigned to the spectral polygons, code combinations associated with that legend value, and the uncertainty in the label assignment. These items are all carried as part of version one of the 1993 habitat base data. Figure 5 is displays this information. The upper map shows the habitat label assignment. This is an area of agriculture and urban land use. The lower map, displays uncertainty in the label assignment as a gray scale. Light gray is highly certain. Dark gray is highly uncertain. For this area, the label assignment for urban and some of the agricultural areas are very certain. Generally for the overall area, natural habitat communities have higher uncertainty. The highest uncertainty values tend to occur outside the change detection area. These are areas that will be relying on the CDF-USFS data.

Underlying the 1993 base map are the six separate sources. These carry the original legend from that source, the assigned habitat legend value, and a value identifying the degree of fit into the habitat legend description. The process of evaluating the 1993 habitat legend and the label assignment is an ongoing process going forward into phase 2 and phase 3 of the overall monitoring program.

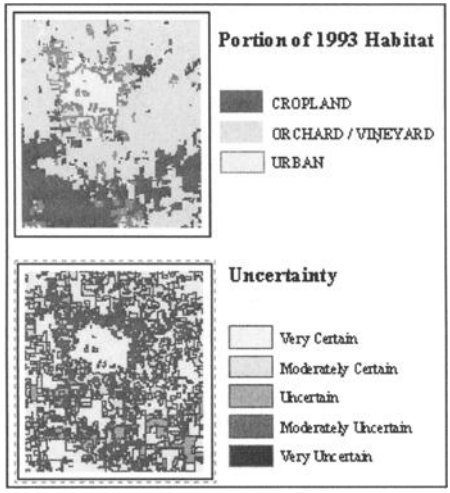


FIG. 5 — *Portion of the 1993 habitat data displayed with uncertainty in label assignment.*

Phase Two - Development of Spectral Change Polygons

Phase two follows procedures used by the CDF - USFS for the statewide change detection project (Levien et al. 1998). It requires registration of satellite scenes for the 1993 and 2000 dates. This phase is independent of phase one. The process in this phase identifies spectral change between scenes after the image normalization to correct for radiometric differences between image dates. A pixel level change map is developed for the two images based on a multi-temporal Kauth-Thomas (MKT) transformation. This produces values for brightness, greenness, and wetness to identify changes in vegetation between the two dates. The segmentation algorithms are then run to produce segments of spectrally contiguous pixels. At this stage, they represent spectral change in the combined values of brightness, greenness, and wetness. They are classified into levels of change. Coordination of phase two is required to ensure integration of the spectral change data between the Central Valley effort with the statewide effort.

Phase Three - Identify Cause of Spectral Change and Change Analysis

At this stage, habitat, land cover, or land use will be determined for change areas for the year 2000 and a cause of change will be identified. It will begin in the spring of 2001. For USFWS and USBR, most of this is focused on the area identified for change analysis in Fig. 2. This phase will include field and aerial photo interpretation of randomly selected areas to identify the cause of change and current habitat, land

cover, or land use. In this phase, other data sources will be incorporated. The major block of additional data will be the change detection work by CDF and USFS. Other ancillary data will be incorporated. This includes detailed mapping of riparian system along some waterways and vernal pool mapping for portions of the Central Valley. The major product produced in this phase will be a habitat change analysis between 1993 and 2000. A major component of this phase is the design and implementation of an accuracy assessment for the change analysis.

Also assessed at this stage is the habitat legend used in the initial development of the 1993 base map. This legend can be expected to change. The legend should be consistent with USFWS standards for both wetland classifications (Cowardin et al. 1979) and riparian classifications (Dall et al. 1997). Several groups in California are evaluating the existing classification systems and how those systems will nest into the National effort for a classification system. These standardization efforts will affect the legends used for this change detection work. USFWS and USBR need to ensure that the legend carried from the 1993 base map will nest within the statewide effort. This requires a commitment of staff resources particularly by USFWS. Their staff will need to spend time in identifying the characteristics for these habitats and coordinating with other organizations to include these types into a common classification scheme. This also carries forward into ensuring that the mapping and legend is consistent with ongoing National efforts for vegetation classification, and land cover, land use mapping.

Accuracy assessment will cover both the effective base map area for 1993 as well as the focus area for change analysis (Figure 2). This will evaluate the identification of the cause of change for a randomly selected set of spectral change polygons. It will also evaluate the habitat label assignment for 1993 based on aerial photography for that time period. Assessment will be split between staff of USFWS and USBR based on the areas of primary interest between the organizations. During this phase, the staff of both organizations will to be working closely together to coordinate sampling and evaluation to avoid duplicating efforts. Methods for this assessment need to be consistent with those used by CDF and USFS for the statewide change detection project.

Summary

The effort by USFWS and USBR to jointly develop a habitat monitor program provides an opportunity to link into statewide change detection efforts. This coordination permits the use of other active mapping data for major portions of the project area. It reduces the amount of time required to prepare a habitat legend. There represents significant cost savings for this project. There are tradeoffs. To successfully incorporate the statewide change detection work, similar procedures need to be followed. Some habitats of interest to USFWS are not represented in the initial legend. The use of satellite imagery does not adequately represent some habitat types. To address these issues, close coordination and cooperation of staff from the different organizations are required. Staff must understand the information contained

and not contained in data products, and associated accuracy or uncertainty. They need to convey this information to their managers to refine the process to identify information required by the organizations. This may require the identification of additional sources of information or methodologies. This requires a commitment by managers to provide significant staff time for this coordination.

This coordination represents part of the standardization process. It began with the development of a common work plan or statement of work between USBR and USFWS. This document identified major objectives and requirements for the project. To meet those objective, the cooperative effort has extended beyond these organizations to other state efforts. It has evolved into identifying common definitions, classifications, and methodologies among several organizations. The goal of this process is fit within other regional state and National efforts.

Reference

California Department of Conservation (DOC), Division of Land Resource Protection, 1994, *A Guide to the Farmland Mapping and Monitoring Program*, California Department of Conservation, Sacramento, CA, http://www.consrv.ca.gov/dlrp/FMMP/pubs/guide.html.

California Department of Fish and Game, University of California Santa Barbara, November, 1998, *Gap Analysis of Mainland California, an Interactive Atlas of Terrestrial Biodiversity*, CD-Rom, California Department of Fish and Game, Sacramento, CA.

California Department of Water Resources (DWR), Statewide Planning Branch, Division of Planning, July 1993, *Standard Land Use Legend*, California Department of Water Resources, Sacramento, CA.

Cowardin, L.M., V. Carter, F. C. Golet, and E. T. LaRoe, 1979, *Classification of Wetlands and Deepwater Habitats of the United States*, U.S. Fish and Wildlife Service, Biological Services Division, FWS/OBS 79/31, Washington, DC.

Dall, D., Elliot, C., and Peters, D., December 1997, *A System for Mapping Riparian Areas in the Western United States*, U.S. Fish and Wildlife Service, National Wetlands Inventory, Washington, DC.

Ducks Unlimited, January 17, 1997, *California Wetland and Riparian Geographic Information System Project, Final Report*, California Department of Fish and Game, Natural Heritage Division, Sacramento, CA.

Federal Geographic Data Committee (FGDC), Vegetation Subcommittee, June 1997, *National Vegetation Classification Standard*, U.S. Department of Interior, Federal Geographic Data Committee, Washington, DC.

Hansen, D. T., July 1999, *Using Accuracy or Uncertainty in the Spatial Characteristics of Themes in ArcView GIS*, ESRI 1999 User Conference Proceedings, San Diego, CA.

Holland, R. F., June 1998, *No Net Loss? Changes in Great Valley Vernal Pool Distribution from 1989 to 1997*. California Department of Fish and Game, Sacramento, CA.

Levien, L. M., Fischer, C. S., Roffers, P. D., and Maurizi, B. A., April 1998, *Statewide Change Detection Using Multitemporal Remote Sensing Data*, Seventh Forest Service Remote Sensing Applications Conference, Nassau Bay, TX.

Mayer, K. E., and Laudenslayer, W. F., Jr. Eds., October 1988, *A Guide to Wildlife Habitat of California*, California Department of Forestry and Fire Protection, Sacramento CA.

Multi-Resolution Land Characteristics Consortium, NLCD Classification System (NLCD), September 2000, URL: http://www.epa.gov/NLCD/classes.html, U.S. Environmental Protection Agency, September 7, 2000.

Pillsbury, N. H., et al., 1991, Mapping and GIS Database Development for California's Hardwood Resource, California Department of Forestry and Fire Protection, Sacramento, CA.

Ryherd, S. L. and C.E. Woodcock, April 1990, *The Use of Texture in Image Segmentation for the Definition of Forest Stand Boundaries*, in 23rd International Symposium on Remote Sensing of Environment, Bangkok, Thailand.

U. S. Bureau of Reclamation and U. S. Fish and Wildlife Service, August 28 2000, *Central Valley Habitat Monitoring Program General Work Plan*, Sacramento, CA.

Modeling Environmental and Hydrologic Systems

Sandra G. García[1]

On the Use of Spatiotemporal Techniques for the Assessment of Flash Flood Warning

Reference: García, S. G. **"On the Use of Spatiotemporal Techniques for the Assessment of Flash Flood Warning,"** *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy and Standardization, ASTM STP 1420*, D. T. Hansen, V. H. Singhroy, R. R. Pierce, and A. I. Johnson, Eds., ASTM International, West Conshohocken, PA, 2003.

Abstract: In this paper spatial analysis tools are presented which allow the simulation and prediction of flash flooding in semiarid basins. Real-time data from automatic hydrological information systems, the SAIH system, in Spain are integrated in a Geographical Information System (GIS) environment. Making intensive use of the Digital Elevation Model (DEM), the parameters of the developed distributed hydrological models are estimated. The proposed operational environment, *Shyska*, is applied to the management of flash floods in real time in the basins of the southeast of Spain. This system constitutes a support tool for decision taking, based on spatial analysis, when a flood event occurs.

Keywords: Spain, flash flood, distributed hydrological model, flood hazard, GIS, DEM

Introduction

The semiarid southeast of Spain is widely affected by flash flooding. In general, the Mediterranean area is characterized by high-intensity convective storms. Small and medium-sized basins (with steep gradients) usually present flash floods, especially in autumn. These riverbeds, generally dry, are called *ramblas*.

These localized storms can translate into great flow increases with high flow velocities. These sudden events can cause landslides with high levels of material damage and even loss of life.

[1]Professor, Department of Thermic Engineering and Fluids, Hydraulic Engineering Area, Technical University of Cartagena, Paseo de Alfonso XIII, 52, Cartagena (30203), Murcia, Spain.

Due to the short period of time between the peak of the flash hydrograph and the associated rainfall, warning of events of this kind must be based on hydrological forecasting. As reported by Lanza et al. (1994), in Mediterranean basins the community's response time is often greater than the basin's own response time.
The development of hydrological models for arid and semiarid areas, such as those which concern us here, presents certain difficulties. The small number of storm events added to the spatial heterogeneity of the rainfall makes it difficult, or even impossible to calibrate the parameters of traditional models, generally aggregated. In these cases it is advisable to develop and apply simple distributed models with skill to represent the spatial distribution of the characteristics of the basin and the precipitation. The analysis of information in digital form allows the estimation of runoff using spatial techniques.
In these cases, Geographical Information System (GIS) play a fundamental part, allowing the acquisition, management, processing and analysis of spatial information. At present, the Digital Elevation Model (DEM) are becoming more widely available all over the world. The development of hydrological models based on the geomorphology of the basin, by means of the DEM is becoming a possibility.
In the present work, tools integrated in the *Shyska* system (García 1997, 2000) are presented within a GIS environment, which allow the application of spatial techniques as a means to our end.
In semiarid basins, the state of soil moisture has a decisive influence on the response of the basin when a flood event occurs, and may be substantially different between basins and year seasons. This type of basin undergoes cycles of drought which should be considered when applying real time forecasting models. To define the degree of severity of the aridity of the basin makes it easier to select in real time the most suitable values of model parameters and thus to obtain more reliable forecasts.
Operational flood management methodologies in real time are presented for flash flood situations. Finally, the hydrological simulations of a series of storm events are presented for basins found within Spain.

Alternatives for the Follow-up, Forecasting and Warning of Flooding

The concept of flood defence usually consists of (Yevjevich 1994) flood forecasting, warning, evacuation and physical defence. The operational flood management can be considered as a sequence of following activities (Catelli et al. 1998):

- Detection of the likelihood of a flood event.
- Forecasting of future river flows conditions for the hydrometeorological observations.
- Warning issued to the appropriate authorities on the severity and timing of the flood.
- Response by the public and authorities.

Based on different works (Feldman 1994, Tucci 1998, García 2000) the following alternatives of monitoring and forecasting flooding, have been defined.

a- Real time measurement of flow rates and rainfall by a SAIH network.

b- Real time measurement of flow-rates by a SAIH network and input to a river routing model.
c- Real time measurement of rainfall (and flow-rate) by means of a SAIH network and a distributed precipitation-runoff (P-R) model oriented to forecasting.
d- Real time measurement of rainfall (and flow-rate) based on a SAIH network and radar (or radar-satellite) data and a distributed P-R model oriented to forecasting.
e- Integration of quantitative precipitation forecasting (QPF) in a distributed P-R model oriented to forecasting. Also considering flow-rates registered by a SAIH network.
f- Hydrograph forecasting based on levels or flow rates upstream or in tributaries.
g- Hydrograph forecasting based on levels or flow-rates upstream and one of the alternatives c, d, or e for the lateral contribution basin.

At present, in the *Shyska* environment, hydrological forecasting options c, d and e are available. The most suitable alternative should be defined according to the objectives of the forecast, the data available in the study basin and the lead time required to minimize the damage.

Methodologies

Hydrologic Models

In the present work, distributed hydrologic models oriented to simulation and forecasting are used, having been developed with functions embedded in a GIS environment. When faced with a flood episode in semiarid basins, the hydrologic processes implied may be limited to the production and runoff routing at the basin outlet. The models developed, described in detail by García (2000), may be divided into three components:

- Estimation of precipitation field.
- Runoff generation.
- Runoff routing.

The precipitation fields were estimated by applying an interpolation scheme to the pluviometric data, supplied for short periods of time by telemetric networks such as the SAIH systems in Spain.

Runoff Generation Component — In real time forecasting applications, it is best to give priority to relatively simple trustworthy models, where the number of parameters is minimal compatible with the basic information available and with the required precision.

Based on these premises, we chose to use for excess precipitation estimation the Curve Number Method (*CN*) of the Soil Conservation Service (SCS) modified for its application in a spatially determined way. Spatially, the *CN* parameter is a function of other layers in a GIS environment.

This model offers the possibility of adjusting the term of initial retentions in accordance with a moisture balance at storm level when necessary.

Runoff Routing — Distributed Unit Hydrograph (UH) models applied at cell level have been integrated into the *Shyska* environment. Each cell (or group of cells) dividing the basin is considered as an independent unit, with a reply function at the basin outlet. Its most significant parameters are derived from the DEM. Here the methods proposed by Maidment et al. (1996) are used:

- Pure Translation model, which applies the area-time curve extracted from the DEM. The area-time curve indicates the distribution of the partial draining areas which contribute runoff at the basin outlet as a function of travel time.
- Translation-Storage model, allows the identification at cell level of the flow translation time of a linear channel plus a residence time in a reservoir. This methodology may be considered as a variant of Clark's model modified for its application in a distributed manner.

The environment considers the combination of the Geomorphological Instantaneous UH model (Rodriguez-Iturbe and Valdés 1979) and Nash's model, given by Rosso (1984), but it will not be applied in this work.

Study Basins

The catchments studied are the Mula River basin (169 Km^2) and the Rambla Salada basin (112 Km^2). These basins, with torrential regimes, belong to the Segura River basin (18 870 Km^2) in the southeast of Spain (Figure 1).
An average annual temperature of 16-18 °C characterizes the study region. The average annual rainfall and potential evapotranspiration over the region are about 300-500 mm and 850 mm, respectively.
The Segura River basin presents a real-time hydrometeorological data collection system, known as SAIH-SEGURA (Figure 1). SAIH network is equipped with automated rain gauges, level detectors for reservoirs and channels, and flowmeters for pipes.
The Mula River basin is regulated by the La Cierva Reservoir. Water levels in La Cierva Reservoir are recorded digitally every 5 min by SAIH system.

Derivation of the DEM and Topographical Attributes

For the study regions, TIN (Triangular Irregular Network) models were derived from contour lines digitized from topographical maps of scale 1:25 000, also included are scattered marked points which increases its precision. From the TIN model, working DEMs raster were extracted with a cell size of 50 m by 50 m.
From the idealized DEM raster (whose depressions have been completed numerically) applying spatial analysis commands by *Shyska* environment, hydrologic features of the terrain (i.e., flow direction and drainage accumulation) were obtained.

From the flow direction maps, the contributing basins to the flow-measuring SAIH stations were defined (stream gauge and reservoir). These were found to match well the limits of the basins drawn by hand from printed topographic maps (black lines). (Figure 2 (a) and (b)).

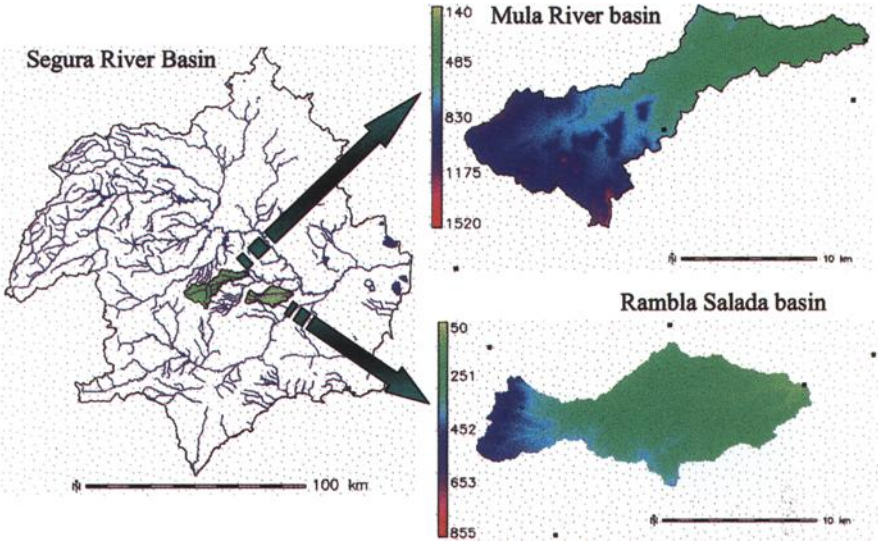
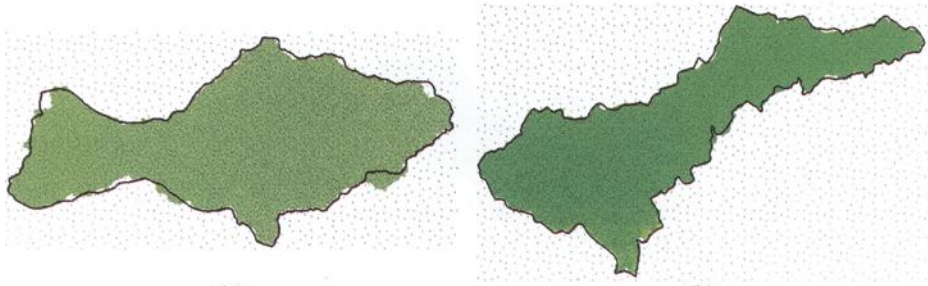


Figure 1 – *Study basins. DEMs (m). Geographical location of SAIH sensors.*



(a) (b)

Figure 2 - *Contrast of basins limits: (b) Rambla Salada basin and (b) Mula River basin.*

For basins defined from the DEM, spatial distributions at cell level of flow velocity, length path to outlet and travel time were calculated.
The flowpath length (L_j) from each cell to the basin edge, is defined as the sum of the flow distances through each cell (l_i):

$$L_j = \sum_{i=1}^{n_j} l_i \qquad (1)$$

where n_j is the number of cells that compound the flowpath j.
The flow velocity (V_i) is assessed spatially distributed but time invariant, as a function of a mean velocity (V_{mean}) for the whole basin, local slope, and drainage area for each cell. This method was proposed by Maidment et al. (1996). For modelling, V_{mean}=20 m/min per calibration. A velocity variation range [V_{min}, V_{max}]= [1, 240] m/min was considered in order to avoid flow velocities which were anomalous or without physical sense.
The flow travel time (T_j) from each cell to the basin outlet, is estimated as the sum of the partial flow times through the cells that make up the flowpath, as follows:

$$T_j = \sum_{i=1}^{n_j} \frac{l_i}{V_i} \tag{2}$$

The spatial distributions of these attributes, for the River Mula basin, are presented in Figure 3.

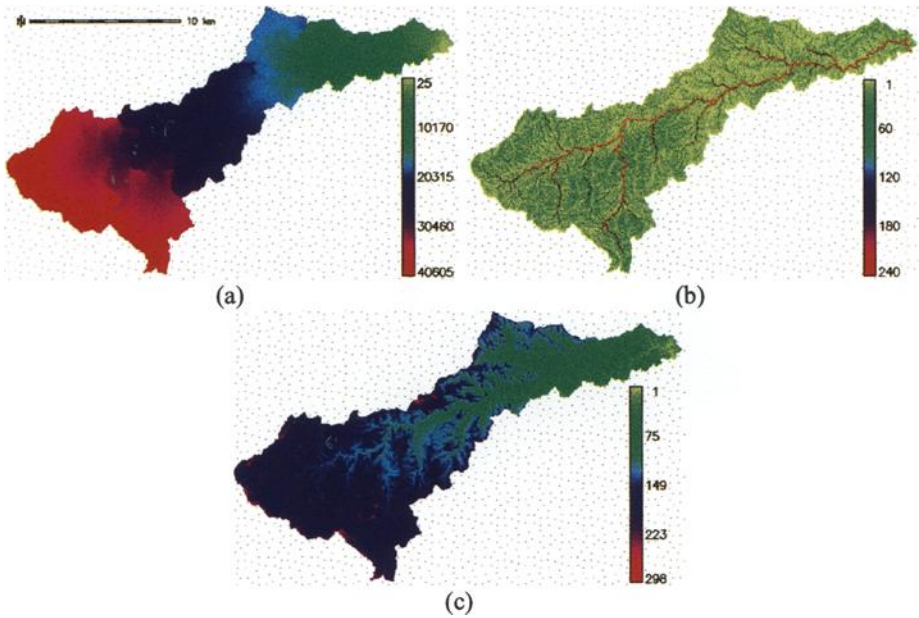


Figure 3 - *Spatial distributions of: (a) flowpath length, m, (b) flow velocity (m/min) and (c) flow travel time (min). Cell size 50 m. Mula River basin.*

From the flow travel time, the time-area histograph is estimated. Figure 4 represents the *Shyska* widget developed for the extraction of these spatially distributed parameters, the time-area histograph and the drainage direction map for the Mula River basin.

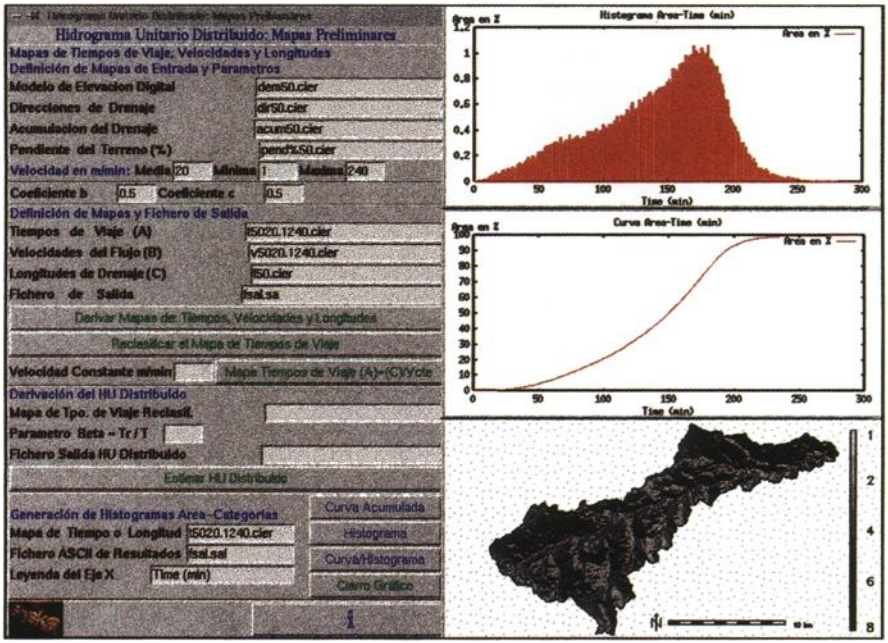


Figure 4 - *Shyska widget for the extraction of topographical attributes and category-area histographs*

Hydrometeorological Analysis: Contrast between Episodes

In order to interpret the response of the basin during storm events, the incidence of different episodes registered by the SAIH-SEGURA telemetric network were contrasted. An analysis was carried out from the pluviometric and hydrometric points of view, estimating the runoff coefficients. The hydrometeorological spatial distributions have been estimated for the basins extracted from the DEM (cell size 50 m).

Rainfall and Runoff Coefficient

The episodes registered on the study basins for the period 1997-2000. They are named according to date, 0997 (29/09/1997 21:00 hs - 30/09/1997 21:00 hs), 0299 (27/02/1999 08:00 hs - 28/02/1999 19:00 hs), 0999 (06/09/1999 15:00 hs - 07/09/1999 21:30 hs) and 1000 (23/10/2000 06:00 hs - 24/10/2000 20:00 hs). The spatial distribution of the total precipitation, which fell in these episodes, in the Rambla Salada Basin can be seen in Figure 5.

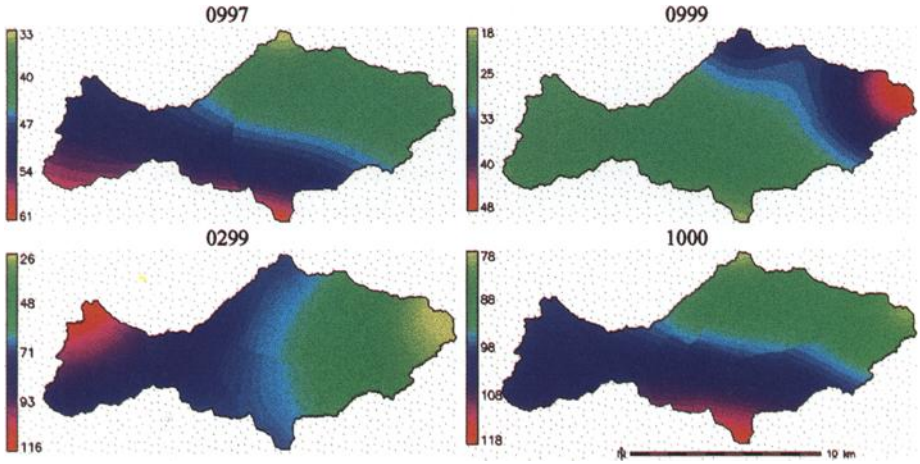


Figure 5 – *Total precipitation maps in several storms. Rambla Salada basin.*

The spatial distribution of the total precipitation in the Mula River basin, corresponding to episodes 0997, 0299 and 1000 can be seen in Figure 6. Episode 0999 was not of great importance in this basin, so that it has not been taken into consideration.

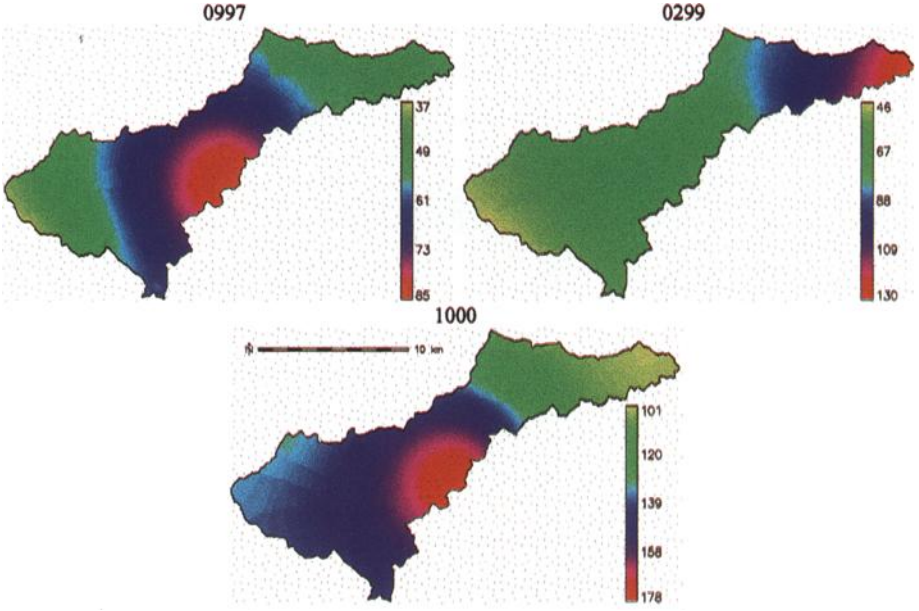


Figure 6 – *Total precipitation maps in several storms. Mula River basin.*

If the analysis concentrates on episodes 0997 and 1000 in both basins, a similar spatial distribution of total rainfall (Figures 5 and 6) can be seen. Although the values found were comparatively higher in episode 1000.
The incidence on the terrain of both events, summed up in (Table 1), shows that in episode 1000 the total volume of rainfall was more than 50 % greater than that registered in 0997. However, the volume of runoff in 1000 was only 40 % of that observed in 0997. The runoff coefficients in both basins for episode 1000 were very low (3-4 %), while in 0997 they reached high values (16-25 %).

Table 1 – *Analysis of the Episodes 0997 and 1000*

Description	Rambla Salada basin		Mula River basin	
	0997	1000	0997	1000
Total Precipitation Volume, Hm^3	5.87	12.40	10.51	23.68
Runoff Volume, Hm^3	0.95	0.42	2.65	0.96
Runoff Coefficient, %	16	3	25	4
Peak Discharge, m^3/s	100	44.75	166.79	41.48

In the absence of anthropic acts (changes in land use, urban development, forestry, etc), the reasons why comparatively greater precipitations produce less flow are to be found in differences in the previous state of soil moisture of the basin at the time of the events mentioned and in the intensity and length of the storms analyzed.
As Martín-Vide et al. (1999) show when analyzing a series of events in a typical ephimeral Mediterranean basin, the storms which go on longer and have a greater amount of associated rainfall produce moderate to small flows. However, short heavy storms with small amounts of associated rainfall generated greater flows.
The antecedent rainfall over the 5 days prior to the storms analyzed, in both basins, show a significant state of drought prior to episode 1000 which did not occur in 0997. However, when a rainfall event occurs, it is necessary to know the state of soil moisture which is to be found in the basin in real time. In semiarid basins, it is vital to have an idea of the severity of drought rather than an exact quantification.
In the literature different techniques for estimating soil moisture based on information from remote sensing are to be found (Gomer et al., 2000). In the present work, we use estimates of potential evapotranspiration and water deficit, based on rainfall and temperature.

Potential Evapotranspiration and Potential Water Deficit

The spatial distributions of certain hydrometeorological variables (precipitation, potential evapotranspiration and potential water deficit) have been analyzed for the hydrological years 1996-1997 and 1999-2000. These include periods of 90 days and 20 days prior to the episodes mentioned (0997 and 1000). The aim was to evaluate the degree of aridity of the basins analyzed.

The potential evapotranspiration map (*ETP*) was estimated as a function of the mean monthly temperature (*TEMP*) applying the Thornthwaite method. The spatial distribution of the potential water deficit (*DEFP*) corresponds to the difference in *PREC-ETP*, where *PREC* is the spatial distribution of precipitation accumulated over the period.

The spatial differences for the mean annual *TEMP* and the annual *ETP* for the hydrological year 1999-2000 can be seen in Figure 7 for the Mula River basin. High values are to be found for temperature (18 °C) and ETP (750 mm), even for the hydrological year 1996-1997. This behavior is similar in the Rambla Salada basin.

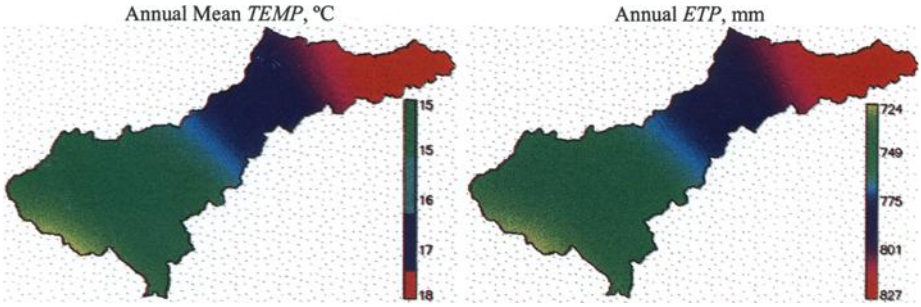
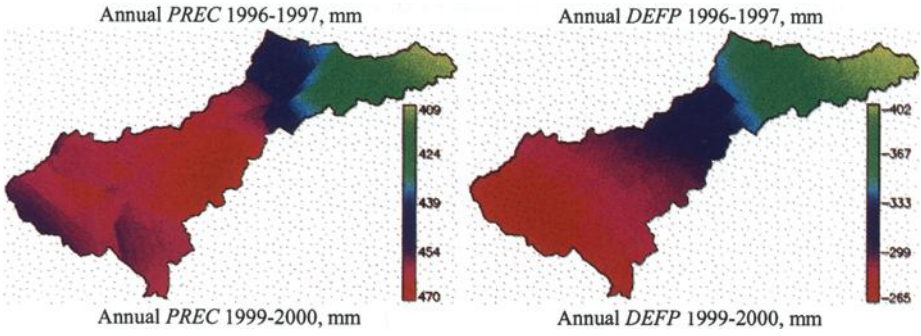


Figure 7 – *Hydrometeorologic Spatial Distributions. Hydrologic year 1999-2000. Mula River basin.*

In Figure 8 the spatial distributions of *PREC* and *DEFP* estimated for hydrological years 1996-1997 and 1999-2000 are presented for the Mula River basin. It can be seen that hydrological year 1999-2000 presented areal precipitations (245 mm) which are 54 % of those registered in 1996-1997 (450 mm). This fact contributed to areal annual water deficits in the year 1999-2000 higher by 40 % than those to be found in 1996-1997. This behavior is similar in the Rambla Salada basin. It can be seen that the degree of humidity was low when episode 1000 occurred. In the analysis of precipitation, evapotranspiration and potential water deficit for the periods of 90 and 20 days prior to events 0997 and 1000, the differences in the potential water deficit between the two periods are not substantial.



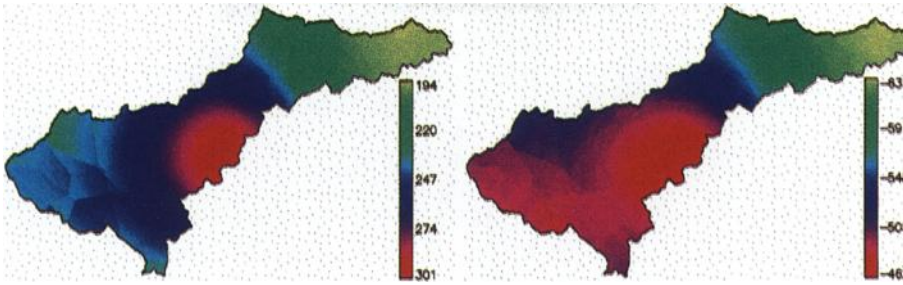


Figure 8 – *Hydrometeorological Spatial Distributions. Hydrologic years 1996-1997 and 1999-2000. Mula River basin*

Hydrological Simulation

The models developed above were applied to one of the study basins (Rambla Salada) for four episodes registered by the SAIH- SEGURA system over the 1997-1999 period. The parameters of the models were calibrated and validated, obtaining satisfactory results working with mean conditions of the state of soil moisture of the basin.

The calibrated models for previous episodes with different soil moisture conditions of the basin, were applied in the simulation of episode 1000. The hydrographs obtained were over-estimated for the same basin. This fact underlined the importance which the consideration of the spatial distribution of the Antecedent Moisture Condition (*AMC*) of the basin presents. As it has been shown, hydrological year 1999-2000 is characterized by its aridity. This year had lower precipitations than the yearly average.

When using the SCS Method of runoff generation, the *AMC* value for the basin should be defined. In episode 1000, on using *AMC*=1 instead of *AMC*=2, flows more similar to the real ones were obtained although still somewhat over-estimated. If the term of initial retentions for the model is increased, by means of a parameter, the fit of the simulated- observed hydrographs for both basins is improved.

To estimate the flow velocity field, the same mean velocity value was used for the whole basin, in both basins and episodes under study, obtaining good results in the hydrograph timing. In the Rambla Salada basin the pure translation model (T.P.) was used, while in the Mula River basin the translation-storage model (T.A.) was adjusted. In Figure 9 (a) the observed-simulated hydrograph fit and the hyetograph obtained for episode 0997 (*AMC*=2) is presented for the Rambla Salada basin. While in Figure 9 (b) the corresponding contrast for episode 2000 selecting *AMC*=1 and adjusting the initial retentions term of the model is presented.

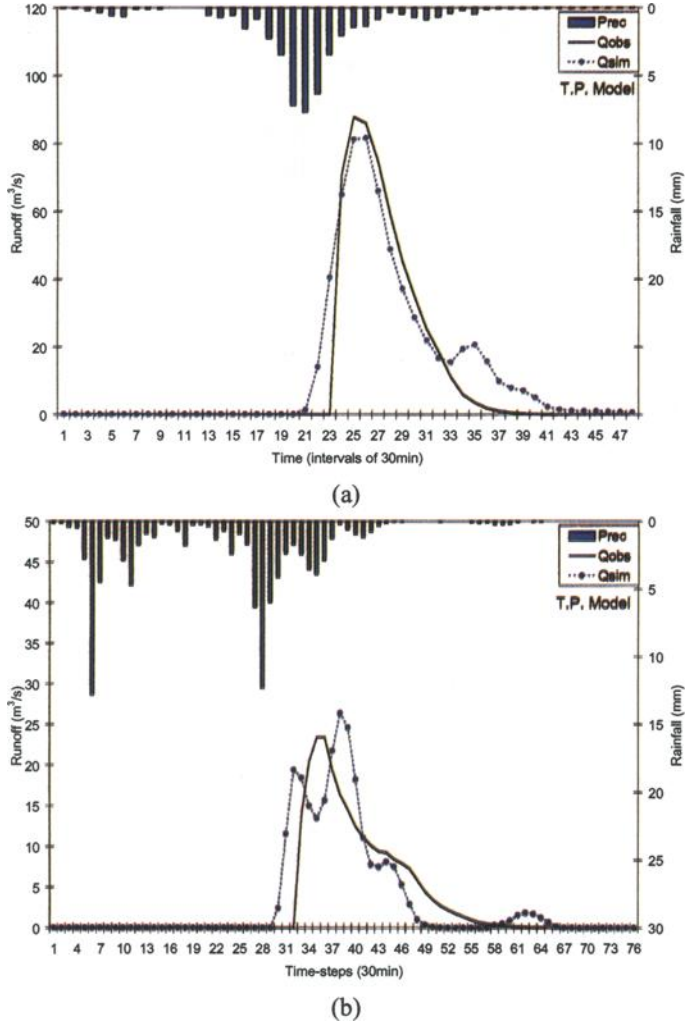


Figure 9 – *Simulation results for Rambla Salada basin: (a) Episode 0997 and (b) Episode 1000.*

An example of *Shyska* session is presented in Figure 10 for an episode simulation (0997) in the Rambla Salada basin. This figure shows the *Shyska*'s main menu, widget for basin and simulation time selection, and simulation results (graphical and numerical results).

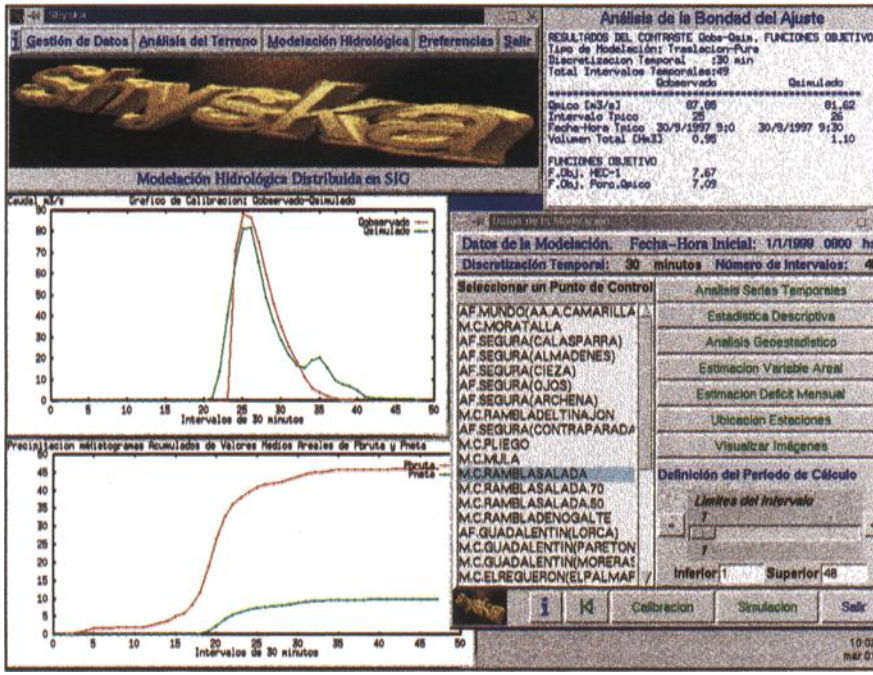


Figure 10 – *Shyska session example. Episode 0997. Rambla Salada basin.*

The Figure 11 presents the simulation results (hydrographs and hyetograph) for Mula River basin in the episode 0997.

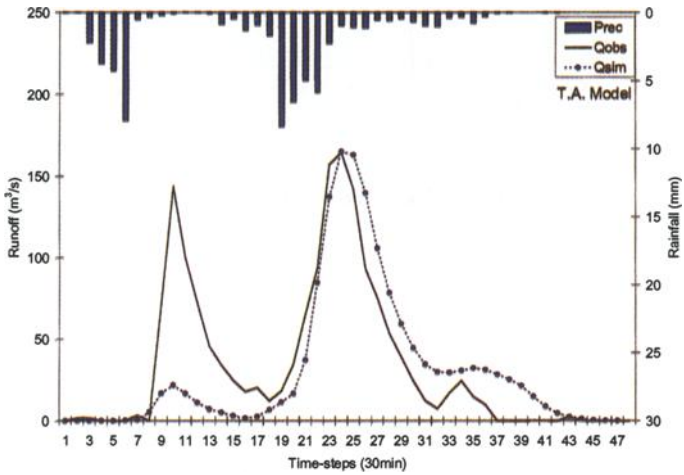


Figure 11 - *Predicted and observed flow for Mula River basin. Episode 0997.*

Conclusions

In this paper the application of a computer environment, *Shyska*, which allows the real-time operational management of a flood event, is presented. It is as a tool to assist in decision taking, for the use of local and regional authorities to the broadcast of warnings of possible flash floods and to mitigate the effects of flooding.
The precipitation grids were estimated by applying an interpolation scheme to the pluviometric data supplied by telemetric networks. These maps may not be representative of the episode studied due to the low density of automatic rain gauge networks in some basins and the space-time discontinuity of convective storms (very localized) which cause the main episodes in arid and semiarid areas.
Over the last few years, techniques have been developed to estimate precipitation using radar or a combination of radar and satellite data. The use of radar information and the development of a compatible distributed hydrologic methodology are looked upon as the best option when applying deterministic P-R forecasting models in real time.
This work supports a conclusion, that particularly for semiarid basins, antecedent soil moisture is a critical parameter. At present, under the *Shyska* environment, work is being done on the integration of the methodologies which consider in a detailed manner the soil moisture balance and the subsurface flow.
In the case of flash flooding in a small basin, the time which exists between the runoff hydrograph peak and the associated rainfall varies between tens of minutes and several hours. Therefore, quantitative rainfall forecasting is considered crucial in Mediterranean basins such as those with wich we are concerned.
Actually, the *Shyska* environment is at the implantation phase in the different sub-basins of the River Segura basin under the auspices of the Confederación Hidrográfica del Segura in its role as authority over the basin.

Acknowledgments

The author thanks the Confederación Hidrográfica del Segura (Spain) for providing data for this study. Opinions given throughout the work belong to the author, and do not represent any institution.

References

Catelli, C., Pani, G., and Todini, E., 1998, "FLOODSS: Flood Operational Decision Support System," *Hydroinformatics '98*, (Eds. Babovic, V., and Larsen, L. C.). AA. Balkema, Rotterdam, ISBN 90 5410 983 1, Vol. 1, pp. 255-262.

Feldman, A., 1994, "Assesment of forecast technology for flood control operation," *Coping with Floods*, (Eds. Rossi, G., Harmancioğlu, N., and Yevjevich, V.), Kluwer Academic Publishers, Netherlands, pp. 445-458.

García, S. G., 1997, "Shyska: A System Oriented to Hydrological Analysis and Distributed Modelling in GIS," *Proceedings of GIS AM/FM ASIA '97 & Geoinformatics '97*, Taipei, Taiwan, Vol. 1, pp. 143-152.

García, S. G., 2000, *A Real-Time Flood Forecasting and Simulation System Based on GIS and DEM: Analysis of Sensitivity to Scale Factors*, Doctoral Thesis, Department of Hydraulic Engineering and Environment, Technical University of Valencia, Valencia, Spain (in spanish).

Gomer, D., Vogt, T., Belz, S., and Vogt, H., 2000, "Landsat TM and GIS as a tool for monitoring the spatial distribution of soil moisture," *Remote Sensing and Hydrology 2000 Symposium*. Proceedings of a symposium held at Santa Fe, New Mexico, April 2000, (Eds. Owe, M., Brubaker, K., Richtie, J., and Rango, A.), IAHS Publ. No. 267 (2001). (in press)

Lanza, L., La Barbera, P., and Siccardi, F., 1994, "Early warnings and quantitative precipitation forecasting," *Coping with Floods*, (Eds. Rossi, G. Harmancioğlu, N., and Yevjevich, V.), Kluwer Academic Publishers, Netherlands, pp. 413-435.

Maidment, D. R., Olivera, F., Calver, A., Eatherall, A., and Fraczeck, W., 1996, "Unit hydrograph derived from a spatially distributed velocity field," *Hydrological Processes*, No. 10, pp. 831-844.

Martín-Vide, J. P., Niñerola, D., Bateman, A., Navarro, A., and Velasco, E., 1999, "Runoff and sediment transport in a torrential ephemeral stream of the Mediterranean coast," *Journal of Hydrology*, Vol. 225, pp. 118-129.

Rodriguez-Iturbe, I., and Valdés, J.B., 1979, "The Geomorphologic Structure of the Hydrologic Response," *Water Resources Research*, Vol. 15, No. 6, pp. 1409-1420.

Rosso, R., 1984, "Nash model relation to Horton order ratios," *Water Resources Research*, Vol. 20, No. 7, pp. 914-920.

Tucci, C. E. M., 1998, *Modelos Hidrológicos*, Editora da Universidade, UFRGS, ABRH, Porto Alegre-RS, Brasil.

Yevjevich, V., 1994, "Technology for coping with floods in the 21st century," *Coping with Floods*, (Eds. Rossi, G., Harmancioğlu, N., and Yevjevich, V.), Kluwer Academic Publishers, Netherlands, pp. 35-43.

Patrick J. Starks,[1] John D. Ross,[1] and Gary C. Heathman[1]

Modeling the Spatial and Temporal Distribution of Soil Moisture at Watershed Scales Using Remote Sensing and GIS

REFERENCE: Starks, P. J., Ross, J. D., and Heathman, G. C., **"Modeling the Spatial and Temporal Distribution of Soil Moisture at Watershed Scales Using Remote Sensing and GIS,"** *Spatial Methods for Solution of Environmental and Hydrologic Problems -- Science, Policy, and Standardization, ASTM STP 1420,* D. T. Hansen, V. H. Singhroy, R. R. Pierce, and A. I. Johnson, Eds., ASTM International, West Conshohocken, PA, 2003.

Abstract: Soil water content (2_v) is of fundamental importance in meteorology, agriculture, and hydrology, among other scientific disciplines. In hydrology, 2_v partitions rainfall into runoff and infiltration, thus impacting surface and groundwater recharge, flood forecasting, and flow routing modeling. Measurement of 2_v at a point is straightforward, but point measurements are inadequate for watershed hydrology due to variability of soil properties, land cover, and meteorological inputs over space. Passive microwave remote sensing systems have been successfully used to provide regional estimates of surface 2_v (0-5 cm surface layer) at the spatial resolution of the sensor. To extend these data to other depths and scales, a two-layer soil water budget model was used to combine remotely sensed estimates of 2_v and spatial information on land cover, soil type and meteorological inputs to predict root zone 2_v over a 611 km^2 watershed. A GIS was used to pre-process and geo-register spatial data sets for input into the soil water budget model, and analyze the results.

Keywords: soil water content, spatial data, remote sensing, hydrology, modeling, watershed, geographic information systems (GIS)

[1]United States Department of Agriculture, Agricultural Research Service, Grazinglands Research Laboratory, El Reno, Oklahoma 73036.

Volumetric soil water content (θ_v) is of fundamental importance in several areas of the Earth Sciences, including meteorology, agriculture, ecosystem dynamics, bio-geochemical cycling and hydrology. In hydrology, θ_v partitions rainfall into runoff and infiltration, thus impacting surface and groundwater recharge, flood forecasting, and flow routing modeling. Measurement of θ_v at a point is straightforward and can be achieved by a number of different techniques. However, point sampling of θ_v over large watersheds is impractical because it is labor intensive, expensive and is generally not acquired on a timely basis.

Microwave remote sensing can provide surface measurements of θ_v to about 10 cm deep, depending upon sensor type and wavelength used (Engman and Chauhan 1995). In a recent study, Jackson et al. (1999) used passive microwave remote sensing to produce regional estimates of surface (0-5 cm surface layer) θ_v over a 10 000 km^2 study area in central Oklahoma. One specific objective of their study, and the objective of this paper, is to examine the feasibility of inferring soil profile θ_v in time and space by combining remotely sensed surface estimates of θ_v, *in situ* field data and modeling techniques (SGP97 1997). The factors that influence estimations of θ_v at large scales are also examined. To address the research objective, data sets from remote sensing and other spatial data sources were processed using a geographical information system (GIS) to provide input to a one-dimensional, soil water budget model. The soil water budget model was modified to accept spatially distributed data and to output results that could be analyzed with GIS software.

Materials and Methods

The Model

Ragab (1995) developed a simple, one-dimensional, two-layer soil water budget model for application in remote sensing studies. The model is based upon the force-restore concept (Bhumralkar 1975, Deardorff 1977) and is divided into a surface layer which corresponds to the depth represented by remotely sensed estimates of θ_v (0-5 cm for this study), and a root zone layer (soil surface to the bottom of the root zone). The model runs on a daily time step and requires inputs of daily rainfall and potential evapotranspiration (ET_p). Required soil data include depth of the surface and root zone layers; and, for each layer, θ_v at "field capacity", "wilting point", and initial soil water content. Maximum and minimum model-allowed values of θ_v are used to prevent the model from becoming unrealistically wet or dry, respectively. The model also requires surface runoff percentage (set to 0 in this study), uptake ratio (surface layer contribution to actual ET) and a psuedo-diffusivity coefficient (governs the rate at which the surface layer dries or wets as a function of soil texture and θ_v of the root zone).

The original version of the model was developed for application at a single location, whether a point or a pixel representing an area. The model was revised to accommodate distributed, spatial data obtained from remote sensing and GIS data layers. In this study, initial soil moisture conditions required by the model were obtained from remotely sensed data. The model was then used to simulate θ_v of the surface and root zone based on daily meteorological inputs of rainfall and ET_p. Values for the pseudo-diffusivity

coefficient were obtained from the literature and adjusted according to the method given in Mehrez at al. (1992).

Study Site

The Little Washita River Experimental Watershed (LWREW) is located in south central Oklahoma (Fig. 1), and is approximately 611 km^2 in size. This site was one of three major ground truthing locations used during the Southern Great Plains 1997 (SGP97) Hydrology Experiment (SGP 1997, Jackson et al. 1999), which was conducted between June 18 and July 16, 1997.

Land use on the LWREW is approximately 60% grazing land and 20% cropland, with the remaining 20% divided between urban areas, streams and water bodies, riparian vegetation, forests, and oil waste land. Grazing lands are primarily tallgrass prairie dominated by big bluestem (*Andropogon gerardii* Vitman), little bluestem (*Schizachyrium scoparium* (Michx.) Nash), indiangrass (*Sorghastrum nutans* (L.) Nash) and switchgrass (*Panicum virgatum* L.). Large grazing areas of bermudagrass (*Cynodon dactylon* (L.) Pers.) are also abundant. Cropland is primarily in winter wheat with smaller areas of corn, cotton, peanuts, and alfalfa.

The LWREW is classified as sub-humid, receiving about 75 cm of precipitation annually, with most of the precipitation occurring in the spring and fall. Soils in the watershed have been categorized into several hydrologic groups on the basis of soil properties that influence infiltration and runoff. In general, most soils have moderate infiltration rates and cover approximately 70% of the watershed. Certain areas of shallow soils in the western portion, as well as a few soils in the eastern region, have high runoff potential and higher infiltration due to predominately sandy soils.

A meteorological network of 45 stations (known as the Micronet), placed on a 5 km spacing, is located within and just outside the watershed boundaries (Fig. 1). Every 5 minutes these stations measure air temperature, relative humidity, incoming solar radiation, soil temperature, and rainfall. Three of these stations also record barometric pressure and windspeed. The data are radio-telemetered every 15 minutes to a central archiving facility.

During SGP97, daily profile measurements of θ_v were made at selected Micronet sites using time domain reflectometery (TDR). Segmented profiling TDR rods (15 cm long segments) were used to sample the soil profile at 0-15, 15-30, 30-45, and 45-60 cm. Total profile θ_v from the TDR was calculated as an average of the readings from the four segments. Nine soil samples, representing the 0-5 cm layer, were also collected daily during the study period at these Micronet locations. Gravimetric soil water content was determined for each sample, averaged and mulitiplied by the respective soil's bulk density to obtain a representative surface value of θ_v for each site. Soil Heat and Water Measurement systems (SHAWMS) were also co-located at selected Micronet sites, and provided hourly measurements of soil water matric potential at five depths in the soil profile (5, 10, 15, 20, 25 and 60 cm) using soil heat dissipation sensors (HDSs). Matric potential data were converted to estimates of θ_v via site- and depth- specific soil water release curves. The 1200 hr CST soil moisture readings from the SHAWMS were used to represent daily θ_v. Surface θ_v from the SHAWMS was taken directly from the 5 cm HDS, while root zone θ_v was computed as a weighted average of all HDSs in the soil

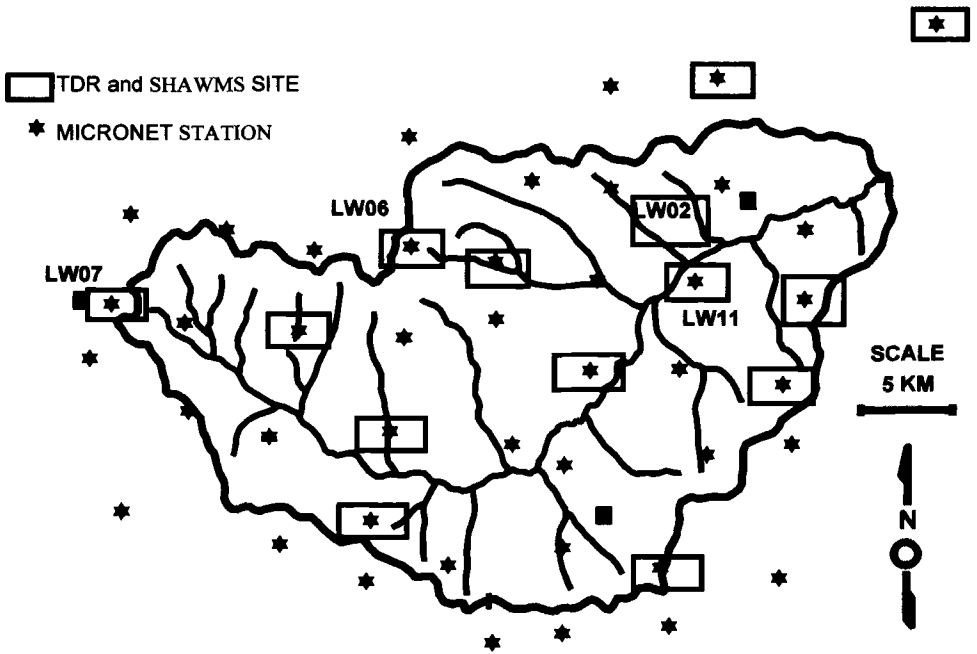


FIG. 1-- *Little Washita River Watershed and Micronet meteorological network. Profiles of soil water content are measured at sites enclosed by the rectangles using time domain reflectometry and/or soil heat dissipation sensors. Sites LW02, LW06, LW07 and LW11 were used as model validation sites.*

profile. Together, the TDR, gravimetric samples and the SHAWMS data provided measured θ_v data for comparison against model simulations.

Remotely Sensed Data

The Electronically Scanned Thinned Array (ESTAR) instrument was flown aboard a NASA P3B aircraft on 16 days during the 30-day study period. ESTAR is an L band, synthetic aperture, passive microwave radiometer operating at a center frequency of 1.413 GHz (21 cm wavelength) and a bandwidth of 20 MHz (Le Vine at al. 1994). The ESTAR was flown at an altitude of 7.5 km, producing surface soil moisture estimates at pixel spatial resolutions of 800 m by 800 m. These remotely sensed data were used directly in the model to provide initial surface θ_v conditions. Initial root zone θ_v was estimated by using the ESTAR θ_v surface data as the independent variable in a regression equation relating near-surface water content to root zone soil water storage (Jackson 1986, Jackson et al. 1987, Zotova and Geller 1985). The equation was developed by regressing measured surface θ_v against measurements of soil water stored in the root zone. The equation took into account a wide range of soil moisture conditions and soil textures. Root zone θ_v was then calculated as the product of the regression result and the depth of the root zone (in mm). The correlation coefficients from the regression equation were comparable to those reported by Ragab (1995) for two soils in southern England.

Zara and Doraiswamy (1998) produced a 30 m by 30 m land cover classification of the entire SGP97 study area. This classification was based upon Landsat TM images obtained on April 29, June 7, July 9, and July 25, and is fully described at http://daac.nasa.gov/CAMPAIGN_DOCS/ SGP97/oklarev.html. The land cover data, in conjunction with soil texture information, were used in this study to identify vegetated and non-vegetated areas in order to assign values of the uptake ratios required by the model. Uptake ratios for non-vegetated soil surfaces were assigned a value of 0.80, areas classified as urban were given a value of 0.05, while vegetated surfaces occurring on sand and sandy loam soils were assigned a value of 0.1. All other vegetated surfaces were given an uptake ratio of 0.35.

Other Spatial Data

Natural Resource Conservation Service State Soil Geographic (STATSGO) data was available for the LWREW, and provided texture classification and particle size fractions of the soil profile down to a depth of 250 cm. These data were available as a series of GIS raster data sets representing a particular location and depth increment in the soil profile. The spatial resolution of the STATSGO data was 1 km by 1 km. Documentation files allowed extraction of soil texture names and average values of percent of sand, silt, and clay for each pixel. The surface layer of the STATSGO data corresponded to the 0-5 cm layer. The texture name was used to extract θ_v at saturation (corresponding maximum allowed model value), field capacity (-33 kPa), wilting point (-1500 kPa), and the residual water content (corresponding to the minimum allowed model value from the tabular information presented in Rawls et al. (1982). The soil texture for the total root zone was determined by calculating the depth-weighted average of the particle fractions

down to a depth of 60 cm. The soil texture name was then determined from the average particle fraction value, and the needed model values selected from the Rawls et al. (1982) data set.

Distributed point values of rainfall were obtained from the LWREW Micronet. The point values were used in a kriging routine, together with their respective latitude and longitude coordinates, to produce spatially continuous rainfall fields for the study area. The rainfall point values were gridded to a 1 km by 1 km spatial resolution. A single value of ET_p was calculated for the study area for each day during the experimental period using data acquired at a meteorological station located near the center of LWREW.

Data Preprocessing

The data sets discussed above represent three different spatial resolutions. Since the model output cannot be any more spatially accurate than the coverage with the coarsest spatial resolution, all data sets were re-sampled to match the STATSGO pixel size. A windowing technique was used to re-sample the spatial data, whereby the majority value of the pixels within the sampling window was assigned as the value for the re-sized pixel. Re-sampling of the ESTAR data from 800 m to 1 km represented little change to the original data. However, re-sampling of the TM land cover data from 30 m spatial resolution to 1 km reduced the number of pixels in the image by a factor of about 800. This reduction could potentially create significant changes in the land cover classification. A histogram of the original TM land cover classification and the land cover after re-sampling is shown in Figure 2. As can be observed, the re-sampling of the original TM land cover theme did not significantly change the distribution of land cover from the original classification. All spatial data sets were geo-referenced and placed into the UTM coordinate system.

Validating Model Output

Validation of model output representing large pixels is problematic since one cannot measure every point within a pixel to verify the model results. Therefore, the reasonableness of the spatial model output is viewed from four approaches. First, the point version of the two-layer model is run using measured meteorological, vegetation and soil property data at four Micronet sites within the LWREW (Fig. 1). The model output is then compared with measured 2_v data to determine if the model is performing as expected. Second, the spatial version of the model is run and output representing the four Micronet sites (1 km by 1 km pixels) is compared to 2_v from the point model above. The purpose of this comparison is to elucidate the effects, if any, on model output caused by different input values used in these two models, although they are run at the same four sites. If both versions of the model used the same input values, then output values of 2_v would fall exactly on a 1:1 line. However, the spatial version of the model uses values for the soil parameters which are obtained from the literature and which may be different than that used in the point version of the model. Further, it is possible that the land cover category used in the spatial version is different than that specified in the point version of the model. Additionally, the point version of the model uses measured rainfall at the site,

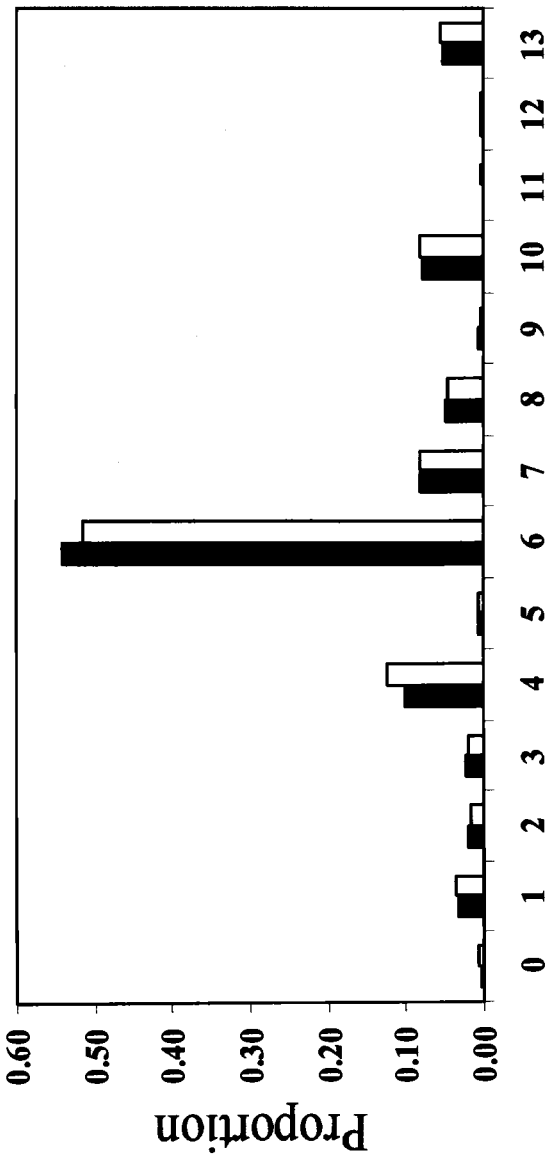


FIG. 2--*Histograms of Landsat TM-derived land cover categories before re-sampling (30 meter data) and after re-sampling (1 km data).*

whereas the spatial version of the model uses an interpolated rainfall value. Third, watershed averages of surface θ_v determined from the spatial version of the model are compared to ESTAR surface data. Since the model is only initialized with ESTAR data, this test provides a comparison of independent estimates of surface θ_v. Finally, time series of watershed-averaged surface and root zone θ_v from the spatial model are compared to time series of θ_v from the four Micronet sites.

Results

Measured and Modeled θ_v at a Point

Figures 3a and b are graphical comparisons of measured and modeled surface and root zone θ_v at the four Micronet sites. The surface layer shows considerable scatter overall, but data from sites LW02 and LW11 cluster about the 1:1 line more than the other two sites. Mean absolute differences between modeled and measured θ_v at sites LW02 and LW11 are quite good at 0.02 $m^3m^{-3}$and 0.03 m^3m^{-3}, respectively. The mean absolute differences at sites LW06 and LW07 are much larger, with values of 0.08 m^3m^{-3} and 0.07 m^3m^{-3}, respectively. Absolute differences averaged over the four sites is 0.05 m^3m^{-3}.

Modeled output of root zone θ_v compares well to measured data (Fig. 3b). Measured θ_v at sites LW06 and LW11 are slightly underestimated by the model but absolute differences are $\leq$ 0.04 m^3m^{-3}. An average absolute difference of about 0.02 m^3m^{-3} was calculated from the four sites.

Spatial Model vs. Point Model Output

Figures 4a and b graphically compare point model and spatial model estimates of θ_v for the surface and root zone layers, respectively, for the four study sites. As expected, the surface comparison exhibits considerable scatter about the 1:1 line. Absolute mean differences between the point and spatial versions of the model are 0.08, 0.09, 0.06, and 0.11 m^3m^{-3} for sites LW02, LW06, LW07 and LW11, respectively. Average mean difference between the modeling scenarios is 0.09 m^3m^{-3}.

Figure 4b indicates a strong linear relationship between θ_v estimates from the two versions of the model. However, except for site LW07 the spatial version of the model greatly underestimates θ_v relative to the point version. Averaged over the four sites, the spatial model underestimates the output from the point model by 0.08 m^3m^{-3}. Sites LW02, LW06, and LW11 exhibit differences of 0.16, 0.06, and 0.08 m^3m^{-3}, respectively. Site LW07 has a mean absolute difference of 0.01 m^3m^{-3}.

It is interesting to note that in Figures 3b and 4b that the data tend to plot into three distinct groups, possibly indicating the impact of soil texture on the model simulations. According to laboratory data, sites LW06 and LW11 are classified as sandy loam and sandy clay loam soils, while site LW02 is a loam and site LW07 is loamy sand.

Average Surface θ_v: ESTAR vs. Watershed

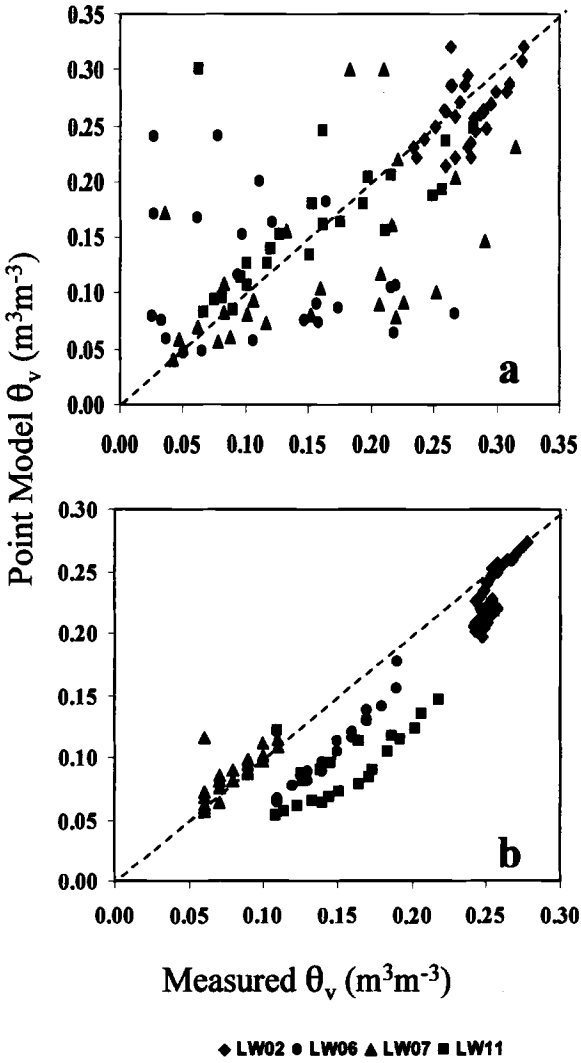


FIG. 3--*Plot of measured vs. modeled volumetric water content at the surface (a) and in total root zone (b) at four Micronet sites.*

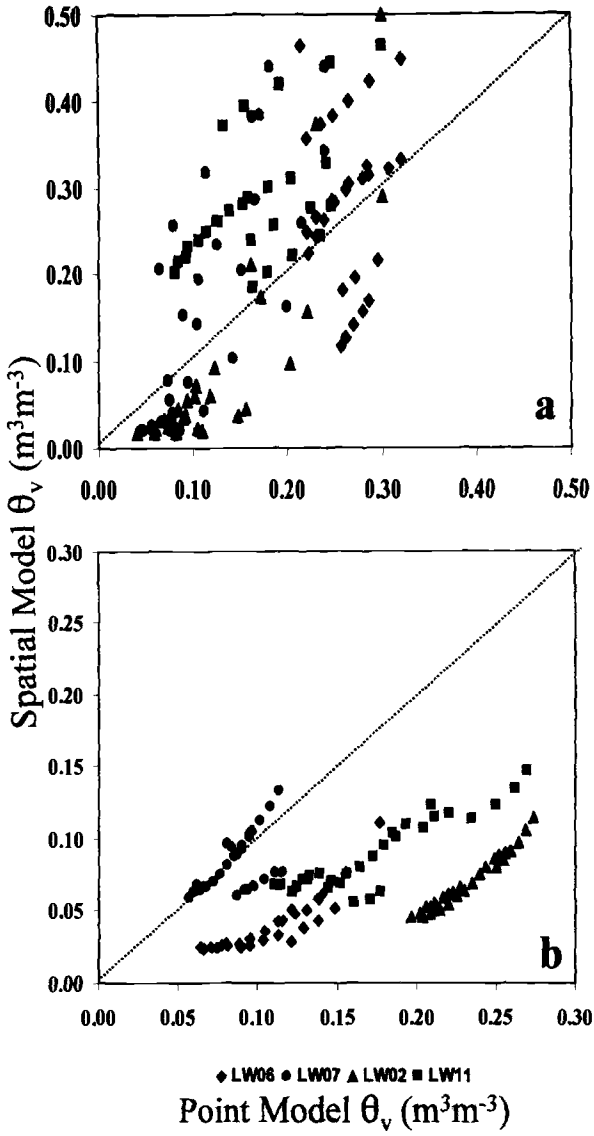


FIG. 4--*Point and spatial model estimates of volumetric water content at the surface (a) and in the total root zone (b) at four Micronet sites.*

Figure 5 is a scatterplot of estimates of surface θ_v averaged over the LWREW. This comparison is only possible for the surface layer and for the 16 days during which the ESTAR instrument was used to collect data. Linear regression of the data presented in Figure 5 yielded a r^2 of 0.91, indicating a strong linear relationship between the ESTAR and spatial model estimates of surface θ_v. However, the spatial version of the model overestimated surface θ_v relative to the ESTAR.

Time Series: Watershed Averages vs. The Four Study Sites

Figures 6 and 7 are time series plots of modeled surface and root zone θ_v, respectively, for the four Micronet sites and the average value for the LWREW. Surface and root zone θ_v for all Micronet sites and the watershed average were derived from the spatial version of the model. Also indicated on each plot is the soil texture histogram for the appropriate soil layer. For the surface layer (Fig. 6), average watershed surface θ_v generally plots as an intermediate value between sites LW02/LW11 and LW06/LW7. It would appear that the watershed value is near what might be expected for a sandy loam soil, which covers about 45% of the watershed area according to the STATSGO data.

Until day of year (DOY) 191 the average watershed root zone θ_v (Fig. 7) closely follows that of LW02, which is classified as a silt textured soil according to the STATSGO data. In fact, the watershed averaged θ_v and that for sites LW02, LW07 and LW11 all show similar trends and values until DOY 191. From DOY 191 until the end of the study period watershed θ_v corresponds more closely to site LW07, which is classified as a loam soil. The root zone soil texture histogram indicates that the largest percentage of the watershed is composed of sandy loam soils (45%), followed by silt (30%), loam (12%), sandy clay loam (7%) and sand (6%). Thus, the watershed average θ_v appears reasonable in light of the dominant soil textures of the watershed.

Summary and Conclusions

In this study, remotely sensed estimates of surface volumetric soil water content were used to initialize a simple, two-layer soil water budget model to determine the feasibility of predicting root zone soil water content. Spatial and temporal estimates of surface and root zone volumetric water content were modeled at watershed scales by integrating spatially distributed remotely sensed, soil, vegetative and meteorological data.

Verification of results from the spatial version of the two-layer model is problematic since it is impractical to obtain a truly representative data set to compare model output against. One approach to verify the spatial model output was to first evaluate the performance of the model at a point. Thus, one could assume that if the model worked well at a point that it would also work well at larger scales. Measured soil, vegetative and meteorological data were used in the point version of the model, and the model output compared well to soil water content measurements. However, point estimations are not usually representative of larger areas due to inherent spatial variability of soil properties, land cover and meteorological conditions. Therefore, a comparison between model outputs at a point and larger areas centering on that point was warranted.

Once satisfied that the model worked well at a point, distributed data sets of soils, vegetation, precipitation and initial soil moisture were used in the spatial version of the

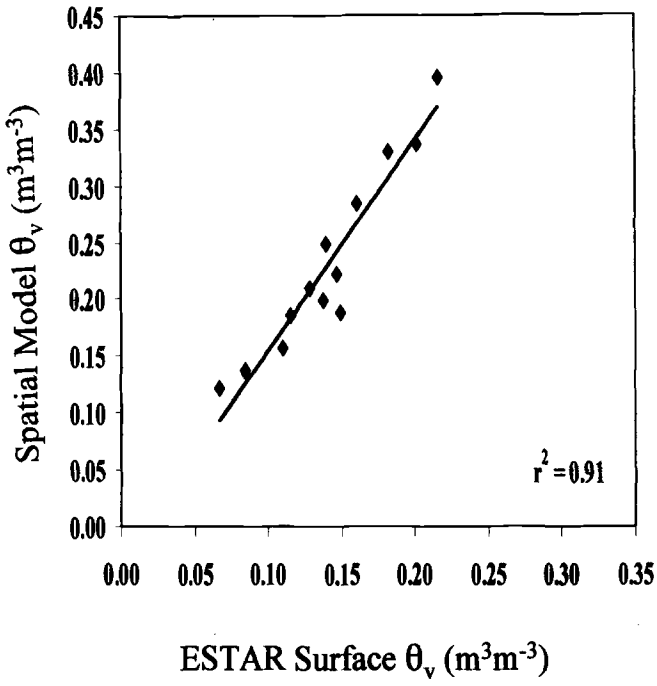


FIG. 5--*Plot of watershed averaged surface θ_v derived from the ESTAR passive microwave radiometer vs. that derived from the spatial model simulations.*

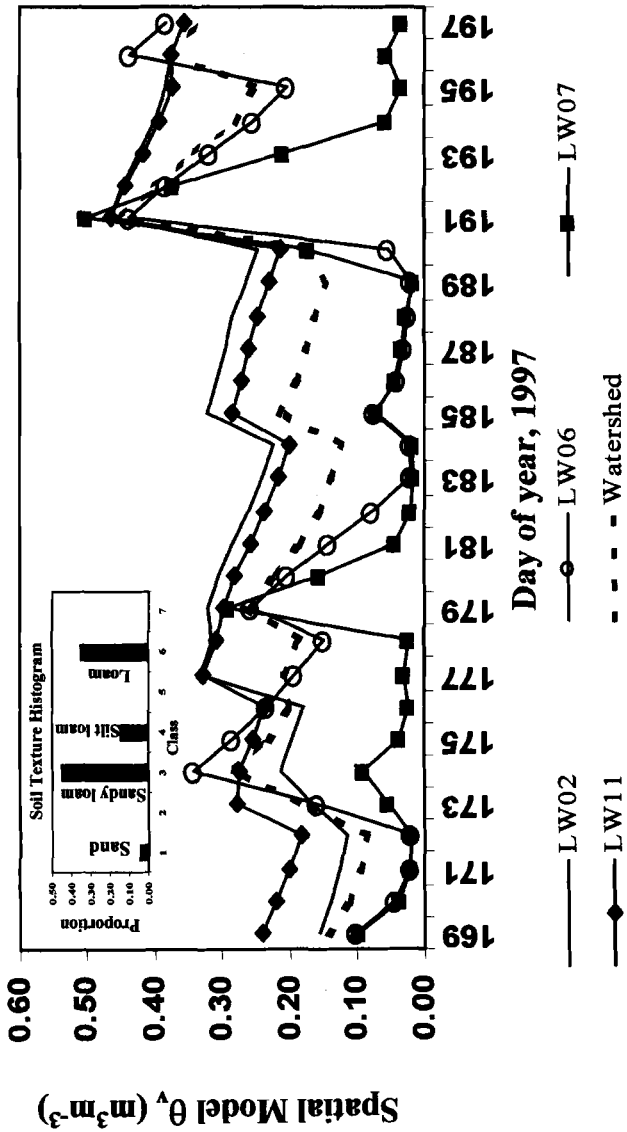


FIG. 6--*Time series plot of surface θ_v for the four Micronet sites and the average watershed value.*

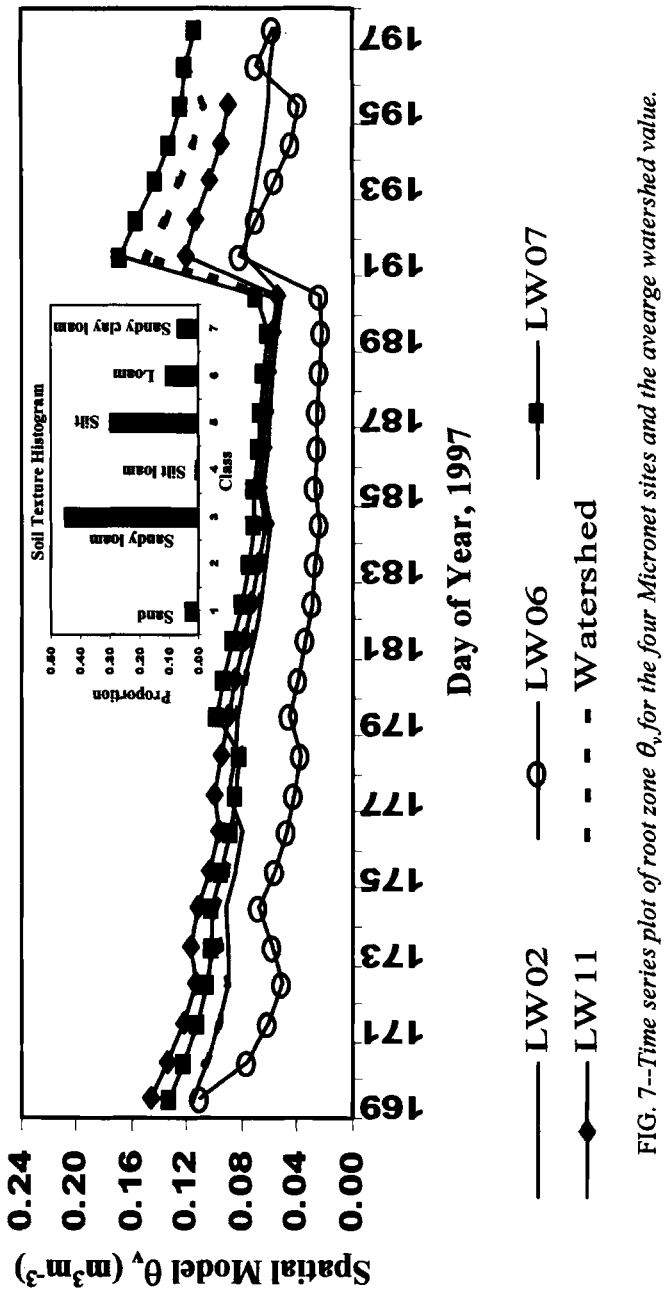


FIG. 7--*Time series plot of root zone θ_v for the four Micronet sites and the avearge watershed value.*

model to simulate distributions of surface and root zone soil moisture over an experimental watershed. Four study sites within the watershed were chosen to compare output from the spatial and point versions of the model. Although strong linear relationships were found, there were considerable differences in absolute values between the two output data sets. These differences were expected since the point and spatial versions of the model used different sources of data to supply the values for the model parameters. The point model incorporated measured data, whereas the spatial version of the model relied upon data obtained from the literature, which represented average values, and from remote sensing and GIS data obtained from various sources.

The ESTAR data used to initialize the spatial model may be in error at some locations. It is known that the ESTAR sensor underestimates surface soil water content when either the soil surface is covered by a thick layer of litter, or if the standing biomass is rather dense, such as in the case of trees (Jackson et al. 1999). Portions of the LWREW have considerable litter covering the soil surface and there are large numbers of trees along watercourses and in other areas of the watershed. The errors induced by such a contaminated ESTAR pixel may not be readily detected or even suspected if one does not know something about the area.

Discrepancies between point and spatial model output may also be due to the soil texture information used in the two models. Differences in STATSGO soil texture data and that measured at some Micronet sites were observed during the study. For example, at site LW07 laboratory-determined soil texture (at a point) indicated that the surface was a loamy sand while the STATSGO data suggested that the 1 km by 1 km area was a silt loam. Differences in soil textures lead to differences in values used to specify the soil's field capacity, wilting point, saturated water content, etc., all of which bear on how the model simulates the dynamics of soil water content.

Other discrepancies between the two model output data sets may be related to the rainfall and/or the land use classification data. The rainfall values used in the point and spatial versions of the model may at times be very different at a particular location. This difference is a result of gridding point values to produce a spatially distributed rainfall field. A rainfall value measured at a site may be different at that site after the gridding process is complete, since some algorithms do not retain the original input value but replace it with an interpolated value based upon data from the surrounding area. The land use classification data used in this study provides guidance for partitioning ET between the surface and root zone layers. It was discovered during the study that at least one location was classified as rangeland in the original 30 m TM image, but was categorized as bare soil during re-sampling to the 1 km spatial resolution.

Two other approaches were used to evaluate the reasonableness of the spatial model output. In one instance, surface θ_v estimates from the ESTAR instrument were compared against estimates from the spatial model. A strong linear correlation between the two was noted, but the modeled values were larger than that from ESTAR. The reason for the relative overestimation by the model is unclear, but may be partially explained by contaminated ESTAR pixels, as explained above. This appears particularly relevant since the model overestimation increases with increasing soil moisture. Lastly, time series plots of averaged watershed surface and root zone θ_v from the spatial model were compared with modeled (spatial model) values at four sites within the watershed. The watershed averages at the surface plotted as intermediate between those for sandy

locations and that for finer textured soils. This result appears reasonable since the largest portion of the surface soils in the watershed is sandy loam. Root zone watershed averages of 2_v also appeared reasonable, with values closely matching the finer textured loam soils.

In conclusion, it is feasible to estimate surface and root zone volumetric soil water content, at the watershed scale, by integrating data from remote sensing, GIS soils data themes, and meteorological instrumentation. Preliminary results (not reported here) indicate that assimilating remotely sensed estimates of surface soil water content on a more frequent basis could enhance simulations of soil water content. Moreover, incorporation of weather and/or climate forecasts may afford the opportunity to make short- and long-term soil water predictions, which may be useful for management of local, and regional water resources.

References

Bhumralkar, C. M., 1975, "Numerical Experiments on the Computation of Ground Surface Temperature in an Atmospheric General Circulation Model," *Journal of Applied Meteorology,* Vol. 14, pp. 1246-1258.

Deardorff, J. W., 1977, "A Parameterization of Ground Surface Moisture Content for Use in Atmospheric Prediction Models," *Journal of Applied Meteorology,* Vol. 16, pp. 1182-1185.

Engman, E. T. and Chauhan, N., 1995, "Status of Microwave Soil Moisture Measurement with Remote Sensing," *Remote Sensing of Environment,* Vol. 51, No. 1, pp. 189-198.

Jackson, T. J., 1986, "Soil Water Modeling and Remote Sensing," *IEEE Transactions on Geoscience and Remote Sensing,* Vol. GE-21, No. 1, pp. 37-46.

Jackson, T. J., Hawley, M. E., and O'Neil, P. E., 1987, "Preplanting Soil Moisture Using Passive Microwave Sensors," *Water Resources Bulletin,* Vol. 23, No. 1, pp. 11-19.

Jackson, T. J., Le Vine, D. M., Hsu, A. Y., Oldak, A., Starks, P. J., Swift, C. T., Isham, J. D., and Haken, M., 1999, "Soil Moisture Mapping at Regional Scales Using Microwave Radiometry: The Southern Great Plains Hydrology Experiment," *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 37, No. 5, pp. 2136-2151.

Le Vine, D. M., Griffis, A. J., Swift, C. T., and Jackson, T. J., 1994, "ESTAR: A Synthetic Aperture Microwave Radiometer for Remote Sensing Applications," *Proceeding of the IEEE,* Vol. 82, pp. 1787-1801.

Mehrez, B. M., Taconet, O., Vidal-Madjar, and Sucksdorff, Y., 1992, "Calibration of an Energy Flux Model Over Soils During the HAPEX-MOBILHY Experiment," *Agricultural and Forest Meteorology,* Vol. 58, pp. 257-283.

Ragab, R. 1995, "Towards a Continuous Operational System to Estimate the Root-Zone Soil Moisture from Intermittent Remotely Sensed Surface Moisture," *Journal of Hydrology*, Vol.173, pp. 1-25.

Rawls, W. J., Brakensiek, D. L., and Saxton, K. E., 1982, "Estimation of soil water properties," *Transactions of the ASAE,* Vol. 25, No. 5, pp. 1316-1320, 1328.

Southern Great Plains 1997 (SGP97) Hydrology Experiment Plan: June 18 – July18, 1997, 178 pp, URL: http://hydrolab.arsusda.gov/~tjackson/, Agricultural Research Service, Hydrology Laboratory, Beltsville, Maryland.

Zara, P. and Doraiswamy, P., 1998, "Retrieval and Mapping of Land Surface Parameters, LAI, and fPAR from Satellite Data," *American Society of Photogrammetric Engineering and Remote Sensing Annual Conference Proceedings,* pp. 539-548.

Zotova, E. N., and Geller, A. G., 1985, "Soil Moisture Estimation by Radar Survey Data During the Sowing Campaign," *International Journal of Remote Sensing,* Vol. 6, No. 2, pp. 353-364.

Spatial and Temporal Integration and Validation of Data

Frederick L. Paillet[1]

Spatial Scale Analysis in Geophysics - Integrating Surface and Borehole Geophysics in Groundwater Studies

Reference: Paillet, F. L., **"Spatial Scale Analysis In Geophysics - Integrating Surface And Borehole Geophysics In Ground Water Studies,"** *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy and Standardization, ASTM International STP 1420*, D. T. Hansen, V. H. Singhroy, R. R. Pierce, and A. I. Johnson, Eds., ASTM International, West Conshohocken, PA, 2002.

Abstract: Integration of geophysical data obtained at various scales can bridge the gap between localized data from boreholes and site-wide data from regional survey profiles. Specific approaches to such analysis include: 1) comparing geophysical measurements in boreholes with the same measurement made from the surface; 2) regressing geophysical data obtained in boreholes with water-sample data from screened intervals; 3) using multiple, physically independent measurements in boreholes to develop multivariate response models for surface geophysical surveys; 4) defining subsurface cell geometry for most effective survey inversion methods; and 5) making geophysical measurements in boreholes to serve as independent verification of geophysical interpretations. Integrated analysis of surface electromagnetic surveys and borehole geophysical logs at a study site in south Florida indicates that salinity of water in the surficial aquifers is controlled by a simple wedge of seawater intrusion along the coast and by a complex pattern of upward brine seepage from deeper aquifers throughout the study area. This interpretation was verified by drilling three additional test boreholes in carefully selected locations.

Keywords: geophysics, borehole geophysics, electromagnetic surveys, seawater intrusion

Introduction

In general, neither surface nor borehole geophysical methods can be used alone for subsurface characterization in groundwater studies. This rather broad statement is based on the observation that surface geophysical surveys almost never have enough resolution to unambiguously define subsurface conditions (Sharma 1997). Much more definitive characterization can usually be performed using borehole geophysics, but there are never enough boreholes to effectively characterize complex formations on the basis of borehole

[1] Project Chief, U.S. Geological Survey, PO Box 25046, MS 403, Denver, CO 80225.

data alone. Therefore, this discussion starts from the premise that effective characterization of subsurface hydrogeologic conditions in a heterogeneous aquifer needs to be based on an effective integration of surface and borehole geophysics with other geologic and hydrogeologic data. At least in concept, subsurface characterization can be completed by using a limited set of borehole measurements to calibrate and otherwise condition a set of surface geophysical measurements that provide complete, three-dimensional coverage of the study region.

Although the need to combine surface and borehole geophysics in site characterization seems obvious, there are few published guidelines as to how to carry out such data integration. Some researchers recommend the "toolbox" approach where a variety of geophysical techniques (the tools) are considered, and a suite of the most appropriate kinds of measurement are used to complete characterization. This study considers an analogous set of "conceptual tools" that might be used for the formulation of an effective and much less ambiguous joint integration of surface and borehole geophysics with other site data. We first list a number of such tools that might serve as a basis for the formulation of a geophysical data inversion and interpretation scheme. We then consider a large-scale site characterization study where each of these generalized conceptual tools was applied to the data integration, and where the specific contribution of each can be identified. The results show that non-invasive characterization of heterogeneous aquifers can be substantially improved by careful attention to the integration of surface and borehole geophysics during the course of the investigation.

The Conceptual Toolbox

In analogy with the many different kinds of geophysical survey equipment available, there are a number of basic concepts that can be applied to the inversion of geophysical data regardless of whether that data is electric, acoustic, or some other class of physical measurement. Because these "conceptual tools" offer a general way of interpreting almost any kind of geophysical data, they can be considered in formulating almost any subsurface investigation. We find a set of five such tools that could, in theory, be applied to any geophysical study. These conceptual tools are summarized in Table 1, where specific examples are provided by way of illustration. More detailed explanations of the conceptual tools are given as follows.

1. Scale of Investigation

Any geophysical survey made at the surface of the earth can, in principle, be made over a much smaller scale of investigation in a borehole. This concept allows the direct investigation of scale of measurement on geophysical response (Figure 1). The surface surveys average measurements over progressively larger sample volumes (defined by R_1, R_2, etc. in Figure 1) as the depth of investigation is increased. The borehole log makes

Table 1 -- *Conceptual tools used in the integration of surface and borehole geophysics.*

Conceptual Tool	Geophysical Measurement	Example of Application
Scale of investigation	Acoustic velocity	Compare bulk velocity of rock in surveys or tomographs to propagation across fractures and in unfractured rock adjacent to borehole
Regression of data	Electrical conductivity	Compare induction log to water sample data for screened intervals to calibrate electric surveys in water-quality units
Multivariate interpretation	Acoustic attenuation	Use logs to identify lithology so as to subtract background variations in intrinsic attenuation from attenuation anomalies on surveys
Inversion methods	Tomography	Parameter specification; size and shape of cells for radar tomograph inversion
Verification boreholes	Seismic reflection	Log rotary-drilled borehole to determine local depth and thickness of reflecting bed

Note:
For additional details on surface and borehole geophysics in environmental and ground water applications consult Paillet and Crowder (1996) and Sharma (1997). Readers are also referred to ASTM International, Standard Guide for Selecting Surface Geophysical Methods, (D 6429-99).

the same measurement (electrical induction, acoustic velocity, bulk density, etc.) over a small sample volume (defined by R_0 in Figure 1) as the probe is moved along the borehole. Thus, we have a means to investigate how small sub-samples within the surface survey volume contribute to the larger-scale geophysical response of the formation.

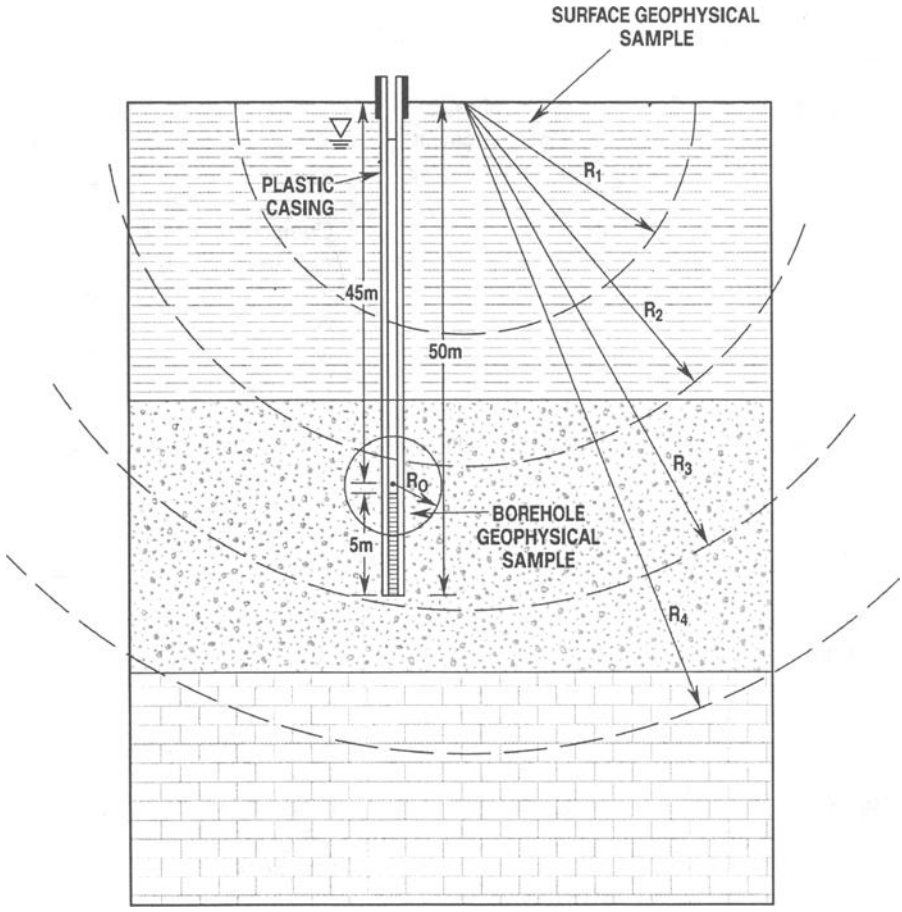


Figure 1 -- *Schematic illustration of scales associated with surface and borehole geophysical measurements compared to typical screened interval for hydraulic testing and water sample analysis.*

2. Regression of Borehole Data to Calibrate Surface Measurements

The exact relation between surface geophysical surveys and hydraulic or geologic properties of interest in the subsurface is often not well known. Because the same kind of measurement can also be made in the borehole over a smaller sample volume, the well logs provide for direct regression of a geophysical measurement with aquifer parameters given by hydraulic tests or water sample analyses. In this approach, the geophysical log response can be averaged over the screened interval in a test well (Figure 1) and this

value can be used to calibrate the surface survey in terms of the hydraulic or water quality property of interest in a particular study. A typical example is given in Figure 2, where the induction log measurement of formation conductivity is averaged over the screened interval in a sampling well to construct a relation between the electrical conductivity of the formation and electrical conductivity of the water sample.

3. Multivariate Interpretation From Standard Logs

Almost all geophysical properties that can be measured at the surface are a function of more than one subsurface variable. Given that fact, a single surface survey cannot be effectively related to one variable of interest where there is significant variation related to other properties of the subsurface. Several different geophysical logs can be run in boreholes and can be interpreted to define a physical model for the multivariate properties of the subsurface. In Figure 2, one geophysical log (natural gamma log) is used to define the aquifer. The combination of gamma and induction logs shows that formation electrical conductivity depends on both pore-water conductivity (in the aquifer), and on formation lithology (clay minerals in the overlying clay-rich alluvium, and in the underlying shale). In this example, the logs demonstrate that the regression between water conductivity and formation electrical conductivity can only be used where the surface geophysical survey interpretations apply to the sand and gravel aquifer. Effective interpretation of surface electromagnetic surveys in terms of water conductivity will only result when either surveys affected by the electrical conductivity of clay minerals in the surficial alluvium are removed from the data set or an interpretation model is used to account for the presence of this alluvium.

4. Inversion Model Characteristics in Data Inversion

The mathematical challenge of geophysical data inversion usually comes down to relating a finite number of surveys to a continuous distribution of subsurface properties. No matter whether the inversion involves one, two, or three dimensions, the continuous distribution in each dimension can be approximated as a series expansion (Parker 1994). There are an infinite number of coefficients in each such expansion. Thus, we never have enough data to form a series of equations relating the finite measurements to the infinite unknown coefficients. One solution is to truncate the series expansions to fewer coefficients than there are data points. This means that there are more equations than unknowns, and the residuals from the additional equations can be used to reduce the mean square difference between model and data. Geophysical logs provide information about the actual distribution of properties in the subsurface that can be used to determine how many coefficients to retain in the expansion, or which series of basis functions to use.

In the practical application of inversion algorithms developed for each class of surface geophysical survey, the user can determine the number of subsurface layers or cells to be

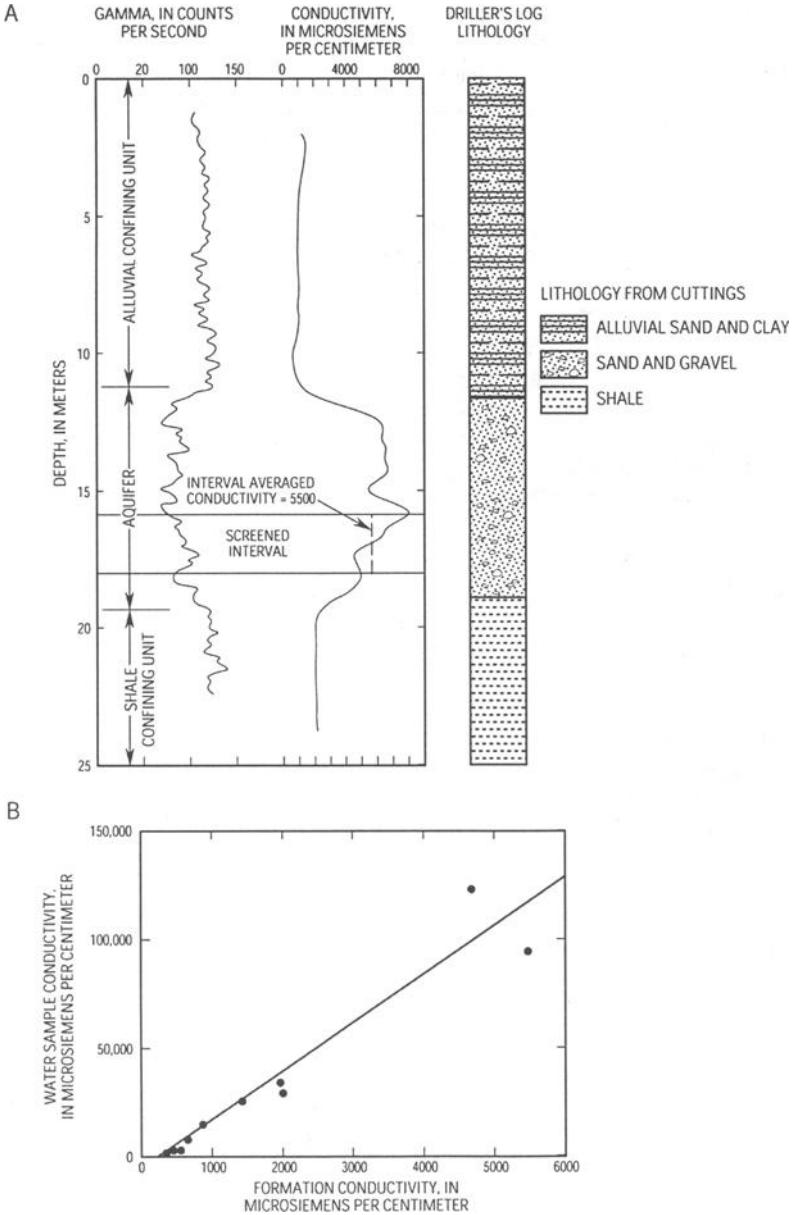


Figure 2 -- *Example of borehole induction methods used to develop a regression between water quality and formation resistivity: A) formation conductivity averaged over screened interval in a sampling well; and B) regression of electrical conductivity of water sample to formation conductivity for a series of monitoring wells.*

used in the analysis. There will always be a reduction in the residual error of the best fit solution as the number of layers or cells is increased. The geophysicist has to decide whether the improvement in the fit of the model to the data set is offset by the reduction of degrees of freedom in the analysis. There are quantitative statistical tests that can be applied to determine whether the improvement is statistically significant, but such tests generally require knowledge of the statistical properties of the subsurface. The specific information about the subsurface structure provided by geophysical logs can significantly improve the ability to formulate and interpret geophysical inversion problems. It is also known that certain geophysical measurements cannot distinguish between alternate subsurface models (for example, electrical equivalence; Sharma 1997). Geophysical logs can provide the information needed to resolve the ambivalences inherent in the selection of a specified inversion model from among equivalent models.

5. Verification Boreholes

When geophysical surveys are interpreted, the final analysis of the data set gives a prediction of subsurface properties over regions between boreholes. A statistically significant verification of the model can be obtained by identifying regions where the model predicts specific features, such as the center of a buried valley or a sharp contrast in the salinity of pore water. Geophysical logs in verification boreholes, commonly drilled at minimal expense by standard rotary drilling, then left uncased and kept open with drilling mud, can be used to verify that these features are present as predicted. When logs show that features predicted by the model actually exist, the results provide almost irrefutable evidence in support of the interpretation. Considerable care can be taken to insure that the verification boreholes are drilled in locations that effectively test the inversion model predictions, so as to maximize the impact of model verification.

The South Florida Study

Surface and borehole geophysics were combined with core descriptions, water sample analyses, and hydraulic tests to generate a predictive model for the surficial aquifers in the region surrounding the Big Cypress National Preserve in south Florida (Weedman et al. 1997, Bennet 1992). In this study, surface time-domain electromagnetic surveys (TDEM; (Fitterman and Stewart 1986, Kaufman and Keller, 1983) were used to project aquifer structure and water quality conditions identified at individual boreholes over the more than 10 km distances between individual drilling sites. Although the results of that study have been described elsewhere (Paillet et al. 1999, Paillet and Reese 2000), the south Florida geophysical data analysis provides a useful example of the contribution of borehole geophysics to the interpretation of surface geophysical surveys. In the following sections, each of the conceptual interpretation tools described in Table 1 is evaluated with respect to its contribution to the electromagnetic survey example from south Florida.

1. Scale of Investigation

Because the focus of the south Florida study was water quality and possible seawater intrusion, the electrical conductivity of the surficial aquifer was of primary interest. The relationship between electrical conductivity and formation properties could be compared at both geophysical log and surface survey scales of investigation (Figure 3). Although other information would be required to generate a useful model of subsurface properties on the basis of this combination of data, the comparison of electrical conductivity measured at two such very different scales of investigation confirms that the local variations of induction conductivity can be related to the depth-averaged measurements of subsurface conductivity given by the surface surveys. For this reason, the comparison of surface and borehole measurements of formation electrical conductivity served as an ideal starting point in the construction of a valid inversion model for the surface TDEM surveys.

2. Water Quality Regression

In the south Florida study, it was possible to relate water quality in a number of zones to local formation conductivity because there was natural flow in most boreholes after completion by installing fully screened casing and flushing of drilling mud (Figure 4). Under those flow conditions, the fluid column resistivity (0.8 ohm meters in Figure 4) could be unambiguously related to the electrical conductivity (12,500 μS/cm) of the pore water entering the borehole in the inflow interval (45-52 m in Figure 4). This analysis was repeated in all boreholes where natural flow was present, and where inflowing water ranged in conductivity from less than 1000 μS/cm to more than 14,000 μS/cm. The regression between formation electrical conductivity and pore water conductivity generated a slope of about 2.3 (the formation factor) that could be used to relate formation electrical conductivity to pore water electrical conductivity in all geophysical measurements where pore water conductivity could not otherwise be determined. This formation factor appeared anomalously low as compared to typical values of greater than 20 in consolidated sandstone (Hearst et al. 2000) but was attributed to the association of inflow with the most permeable intervals within aquifer units characterized by unusually high transmissivity values as reported by Paillet and Reese (2000). Empirical studies demonstrate that the formation factor decreases as permeability increases (Biella et al. 1983, Jorgensen 1991).

3. Multivariate Dependence of Formation Properties

In general, formation electrical conductivity depends on the salinity of pore water, the influence of pore network geometry (permeability) on ion mobility, and the fraction of

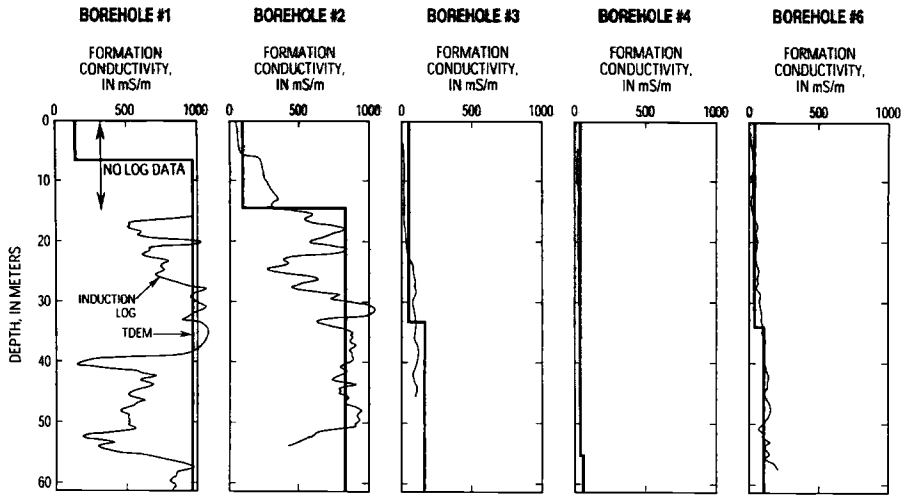


Figure 3 -- *Surface time domain electromagnetic (TDEM) surveys at borehole sites demonstrate that surveys define the electrical conductivity of the uppermost layer, and the composite electrical conductivity of underlying aquifers and confining units.*

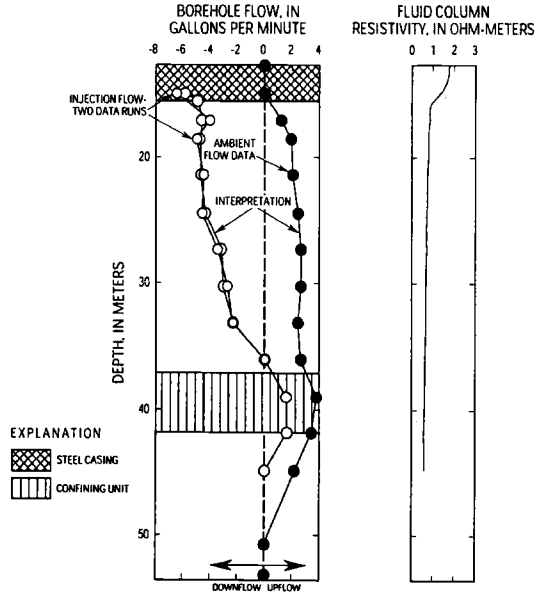


Figure 4 -- *Flow logs obtained in fully screened boreholes showed that natural upward flow existed in most boreholes; the fluid column resistivity profiles of borehole under ambient condition could then be unambiguously related to the electrical conductivity of pore water in the inflow zone or zones.*

electrically conductive minerals (clays; Biella et al. 1983, Jorgensen 1991, Kwader 1985). Geophysical logs from the south Florida boreholes indicated that the contribution of lithology and pore structure to variations in electrical conductivity were negligible (Figure 5). Although the erratic distribution of phosphatic sands caused gamma logs to be of no use in characterizing these sediments, comparison of neutron and induction logs with core lithology confirmed that clays were absent and that formation electrical conductivity and porosity trends ran parallel over discrete intervals (Weedman et al. 1997). This result indicates that the subsurface at each borehole site consists of a series of aquifer layers, each characterized by a single value of pore water salinity. Thin confining units separate aquifers of different pore water salinity, accounting for the step-wise increase in subsurface electrical conductivity. These results indicate that an effective large-scale model for the surficial aquifer is a series of aquifers of different thicknesses and containing water of differing salinity separated by thin, mineralized confining units.

4. Aquifer Structure and Inversion Layers

The layered aquifer framework interpreted from Figure 5 defines the surface electrical survey interpretation as the mapping of the aquifers and confining units identified at each borehole site over the distance between boreholes at this study site. The subsurface structure clearly indicates that model inversion formulated as a series of layers is appropriate for this situation. The comparison of logs and surveys in Figure 3 shows that the surveys effectively indicate the electrical conductivity of the uppermost aquifer layer and the depth-averaged conductivity of the series of aquifers and confining units under that uppermost layer.

An example of the aquifer inversion model constructed from the TDEM surveys along a profile between two of the boreholes at the study site is given in Figure 6. The profiles show that the inversion can be completed using two- or three-model layers for each survey, but that the profile constructed from either set of inversions shows the same structure. The profile indicates an unconfined surficial aquifer with the same water quality extends across the study site (pore-water electrical conductivity of about 400 μS/cm, and identical with the quality of the overlying surface water bodies). The underlying layer is interpreted as a composite of one or more confining units and the underlying aquifers. A wedge of seawater intrusion is interpreted on the southern side of the profile in Figure 6, corresponding with the landward limit of tidal fluctuation in local estuaries. Although not well resolved in this representation, the slope of this interface was verified by a series of more closely spaced TDEM surveys at the southern end of the TDEM profile as shown in Figure 6 (Paillet et al. 1999). Otherwise, the combination of randomly varying conductivity in the lower interpretation layer (two-layer model) and the presence of a strong upward hydraulic gradient throughout the study area indicates that subsurface variations in salinity are related to variations in the rate of upward seepage of brine and the local intrusion of seawater in the immediate vicinity of the coast.

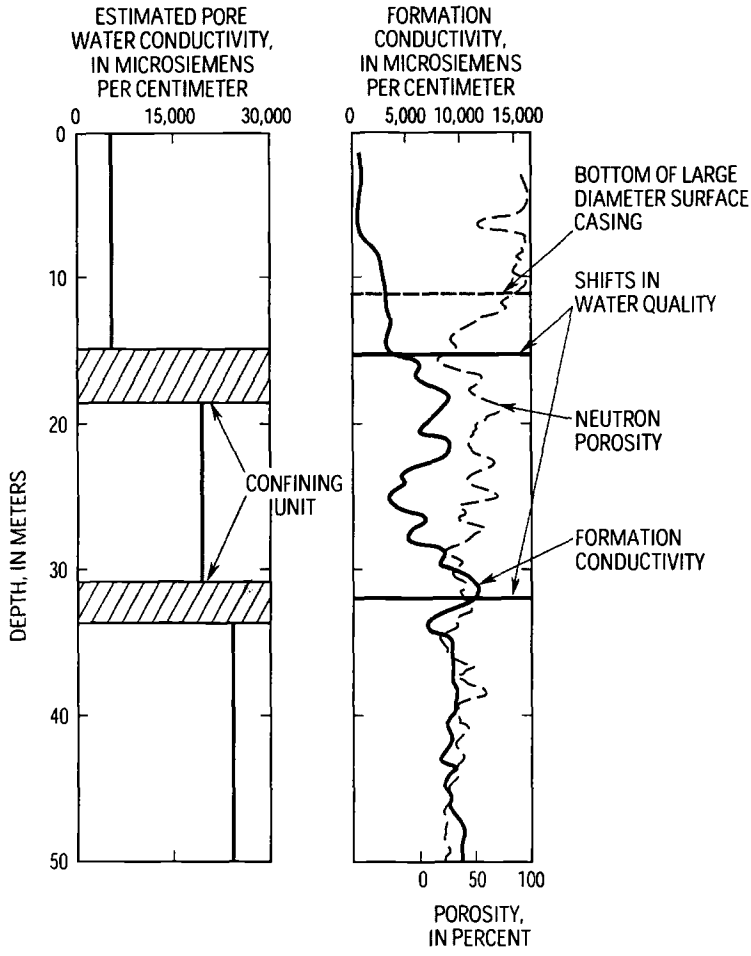


Figure 5 -- *Overlay of induction and neutron porosity logs demonstrate that the surficial aquifer separated by thin confining units into aquifers containing pore water of different salinity, and suitable for electrical modeling as a layered system.*

5. Verification Boreholes

Although the interpretation based on data such as those shown in Figure 6 appears convincing, a few additional boreholes drilled at carefully selected locations would strongly support the interpretation if they showed the depth to the bottom of the upper aquifer and the electrical conductivity of the underlying sediments were in agreement with prediction. Three such boreholes were drilled as part of the south Florida study; two of these are indicated in Figure 6. Induction logs from those boreholes agreed

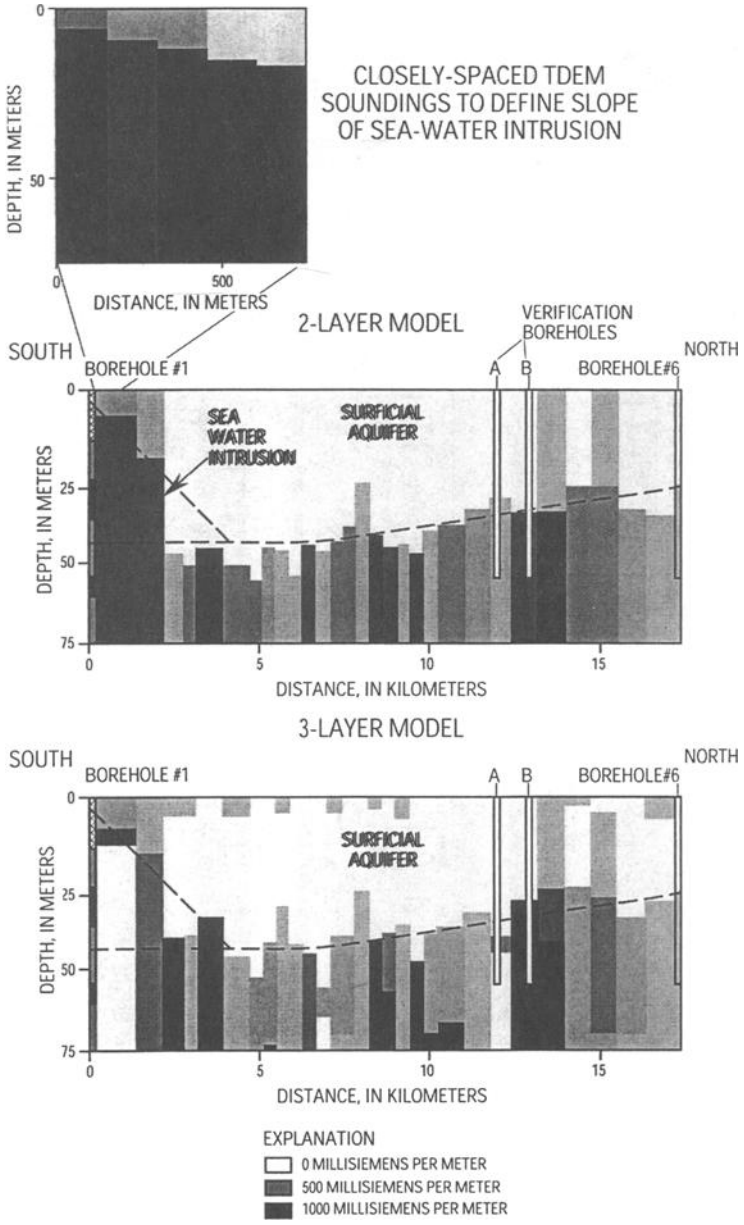


Figure 6 -- *Time domain electromagnetic survey profile across the study area for two-layer and three-layer inversion models, showing that there is no meaningful difference between the two-layer and three-layer inversions.*

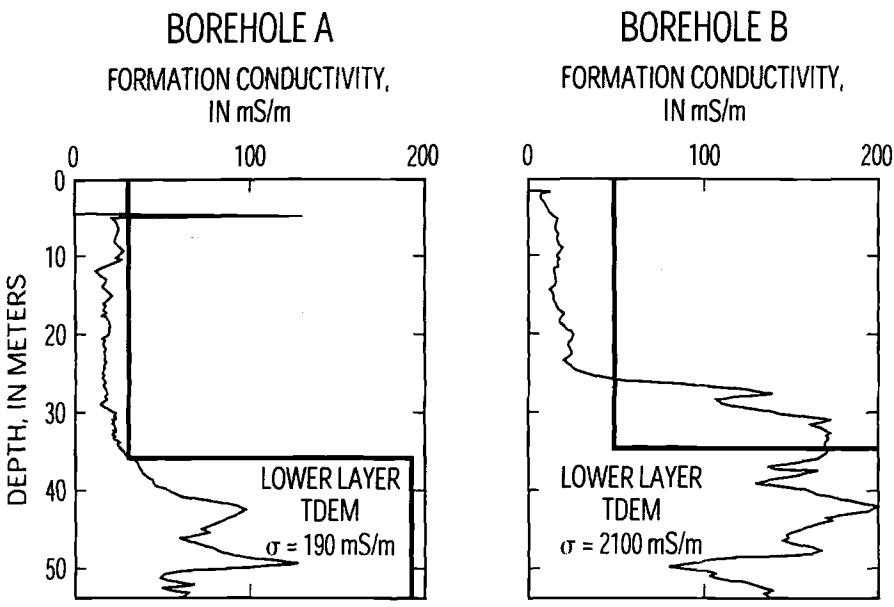


Figure 7 -- *Induction log in rotary-drilled boreholes used to verify model predictions given in Figure 6A compared to the time domain electromagnetic (TDEM) survey inversion at that location in the profile.*

quantitatively with the predictions, showing a definite step in conductivity at the predicted depth and the predicted depth-averaged electrical conductivity of the upper aquifer (Figure 7). The agreement for the underlying zone was only qualitative in that the logs showed an electrical conductivity for the lower layer significantly less than predicted, but the relative magnitude of the measured conductivity agreed with the predictions. That is, the logs in Figure 7 show that the underlying layer is significantly less conductive in borehole A than in borehole B, as indicated by the TDEM surveys. One important result is that the depth to the interface given by the logs in the verification boreholes corresponded with the average depth of the interface constructed from the average of several adjacent TDEM stations (the dashed line in Figure 6), and not the interface given by the TDEM station nearest to the drilling site. This result provides concrete justification for the otherwise reasonable but unproved assumption that the bottom of the surficial aquifer is given by the average trend of the TDEM surveys in the profile.

Conclusions

Integration of geophysical data obtained at various scales provides an effective way to

bridge the gap between localized data from boreholes and site-wide data from regional surveys. Specific conceptual approaches to such analysis are summarized in Table 1. The contribution of each of these approaches to multiple-scale geophysical site characterization was assessed at a study site in south Florida. Comparison of induction logs in boreholes with surface time domain electromagnetic surveys near borehole locations was critical in developing a model relating aquifer framework, water quality, and large-scale electrical conductivity layers for the interpretation model. Regression of pore water electrical conductivity measured in boreholes against formation conductivity given by induction logs was effectively used to interpret the salinity of pore water in the upper aquifer layer using the surface surveys. Joint analysis of the combination of surface electromagnetic surveys and borehole geophysical logs indicated that salinity of water in the surficial aquifers at the study site is controlled by a simple wedge of seawater intrusion along the coast and by a complex pattern of upward brine seepage from deeper aquifers throughout the study site. This interpretation was independently checked by drilling three test boreholes to verify the location of aquifer boundaries and the relative salinity of subsurface waters as predicted by the analysis.

References

Bennet, M. W., "A Three Dimensional Finite Difference Ground Water Flow Model of Western Collier County, Florida," South Florida Water Management District, Technical Publication 92-04, 1992.

Biella, G., Lozeij, A., and Tabacco, I., "Experimental Study of Some Hydrogeological Properties of Unconsolidated Porous Media," *Ground Water*, Vol. 21, No. 6, 1983, pp. 741-750.

Fitterman, D. V., and Stewart, M. T., "Transient Electromagnetic Sounding for Groundwater," *Geophysics*, Vol. 51, No. 3, 1986, pp. 995-1005.

Hearst, J. R., Nelson, P. H., and Paillet, F. L., *Well Logging for Physical Properties* (2nd ed.): John Wiley and Sons, Ltd., New York, 2000.

Jorgensen, D. G., "Estimating Geohydrologic Properties From Corehole Geophysical Logs," *Ground Water Monitoring and Remediation*, Vol. 10, No. 2, 1991, pp. 123-129.

Kaufman, A. A., and Keller, G. V., *Frequency and Transient Soundings*, Elsevier Science Publishers, Amsterdam, 1983.

Kwader, Thomas, "Estimating Aquifer Permeability From Formation Resistivity Factor," *Ground Water*, Vol. 23, No. 6, 1985, pp. 762-766.

Paillet, F. L., and Crowder, R. E., "A Generalized Approach for the Interpretation of Geophysical Well Logs in Ground Water Studies--Theory and Application," *Ground Water*, Vol. 34, No. 5, 1996, pp. 883-898.

Paillet, F. L., Hite, Laura, and Carlson, Matt, "Integrating Surface and Borehole Geophysics in Ground Water Studies - An Example Using Electromagnetic Soundings in South Florida," *Journal of Environmental and Engineering Geophysics*, Vol. 4, No. 1, 1999, pp. 45-55.

Paillet, F. L., and Reese, R. S., 2000, "Integrating Borehole Logs and Aquifer Tests in Aquifer Characterization," *Ground Water*, Vol. 38, No. 5, pp. 713-725.

Parker, R. L., *Geophysical Inverse Theory*, Princeton University Press, Princeton, New Jersey, 1994, pp. 55-117.

Sharma, P. V., *Environmental and Engineering Geophysics*, Cambridge University Press, Cambridge, UK, 1997.

Weedman, S. D., Paillet, F. L., Means, G. H., and Scott, T. M., "Lithology and Geophysics of the Surficial Aquifer System in Western Collier County, Florida," U.S. Geological Survey Open-File Report 97-436, 1997.

Manfred Owe[1] and Richard A. M. De Jeu[2]

The Need for Regular Remote Sensing Observations of Global Soil Moisture

REFERENCE: Owe, M. and De Jeu, R. A. M., **"The Need for Regular Remote Sensing Observations of Global Soil Moisture,"** *Spatial Methods for Solution of Environmental and Hydrolic Problems–Science, Policy and Standardization, ASTM STP 1420*, D.T. Hansen, V. H. Singhroy, R. R. Pierce, and A.I. Johnson, Eds., ASTM International, West Conshohocken, PA 2003.

Abstract: Soil moisture is an important component of the water and energy balance of the Earth's surface, and as such, is essential to many Earth science disciplines. Regular and accurate estimates of soil moisture are an important input to the study of climate change, numerical weather prediction models, drought forecasts, and agricultural predictions. Soil moisture has been identified as a parameter of significant potential for improving the accuracy of large-scale land surface-atmosphere interaction models. However, accurate estimates of surface soil moisture are often difficult to make, especially at large spatial scales. Soil moisture is a highly variable land surface parameter, and while point measurements are usually accurate, they are typically representative only of the immediate sample location, and simple averaging of point values, in order to obtain large-area spatial means, often leads to substantial errors. In addition, ground sampling is labor intensive and costly, and consequently highly impractical for long-term and/or large-scale monitoring. Since satellite remote sensing observations are already a spatially averaged value, they are ideally suited for measuring land surface parameters such as soil moisture. Passive microwave remote sensing presents the greatest potential for providing regular spatially representative estimates of surface soil moisture at global scales. But, while the optimum wavelength for soil moisture sensing is in the L-band (1.4 GHz or $\lambda = 21$ cm), such a sensor has yet to be deployed operationally. However, new and highly improved microwave retrieval techniques are being developed that maximize the information that can be obtained from less optimum sensors, such as C-band and even X-band. Progress from one such study is presented, along with preliminary results of some validation studies and plans to develop a 20-year retrospective global database of surface soil moisture. This data product will be made available to the general public through the Goddard Space Flight Center Distributed Active Archive Center (DAAC). Improved real-time estimates of surface soil moisture should greatly improve the performance of real-time models,

[1] Research Hydrologist, Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA

[2] Research Assistant, Faculty of Earth Sciences, Vrije Universiteit, Amsterdam, The Netherlands

while the development of historical data sets will provide necessary information for simulation and validation of long-term climate and global change studies.

Keywords: Soil moisture, global climate change, remote sensing, microwave

Introduction and Historical Perspectives

Soil moisture is a key land surface parameter in many Earth science disciplines. The application of soil moisture in these various disciplines usually determines the accuracy, spatial requirements, and the subsequent emphasis placed on its parameterization. Soil moisture typically exhibits a high degree of variability in both time and space. While temporal variation is reasonably well understood (i.e., seasonal variation over the course of a year), recent strong seasonal to inter-annual climate variations appear to be caused by anomalies in global circulation patterns. Spatial differences in soil moisture are far less intuitive, and are a function not only of rainfall distributions, but are due to topographic controls, heterogeneity of soil physical properties, and vegetation characteristics.

Soil moisture greatly influences the partitioning of the incoming energy into latent and sensible heat components. Soil moisture together with temperature, thus provide the link between the water and energy balances. Knowledge of the three dimensional soil moisture distribution and its relationship to evapotranspiration is key to understanding the subsequent influence of land surface processes on climate. However, despite its recognized importance, routine global observations of soil moisture, an understanding of its spatial and temporal distribution, and a thorough understanding of its affect on global circulation are still lacking.

Although interest in global change is currently high, researchers today are only rediscovering an old concept. George Perkins Marsh (1874) voiced his concern over the ability of man to alter the Earth's climate, i.e. soil moisture, precipitation, temperature, and atmosphere by modifying the surface vegetation and hydrologic regime through clearing, planting, irrigation, and draining wetlands, but was unable to provide much data in support of his theory. Thornthwaite (1956) recognized the importance of this hypothesis, and indicated the causal conditions for climatological change, such as changes in the general circulation of the atmosphere, variation in incoming radiation and changes in the Earth's features. Significant changes can be brought about by a proper combination of two or more of these factors. Geiger (1950) provided a variety of examples on the effects of human influence on the microclimate resulting from modifications to the surface. Mather (1954) also presented evidence and went into considerable detail on the probable causes of continental and regional climate variation.

Potentially damaging environmental changes that seem to have undergone remarkable acceleration in recent years include global warming, sea level rise, deforestation, and desertification. Although it is thought by some that many environmental changes are cyclic over the long term, there is strong evidence that human activity has disturbed this natural cycle. Environmental parameters are in a dynamic equilibrium, and are subject to a constant and delicate balancing act. When one parameter gets significantly out of proportion, system adjustment may result in other more adverse environmental changes. The stresses exerted on the environment as a

result of increased population growth and subsequent human activity seem to be one of the basic underlying causes of this imbalance. These stresses will continue to manifest unless a global scale effort is made to reduce their causes. Simulation studies with coupled models of the atmospheric general circulation and biosphere (Sellers et al., 1986) have demonstrated long-term adverse effects of certain environmental changes such as ozone depletion and increased atmospheric CO_2 concentrations.

Historically, hydrologists have had little information on the large-scale distribution of soil moisture in time and space. While ground-based sampling methods are typically accurate, they still are point measures, and are not always readily transformed into areal averages, especially at regional, continental, and global scales. Climatic modelling of large-scale soil moisture is frequently used as an alternative, but reliable forcing data are often lacking, especially on a global basis.

Space-based remote sensing offers potentially the greatest single contribution to large-scale monitoring of the Earth's surface. If properly utilized, satellite systems can offer the spatial, temporal, and spectral resolution necessary for consistent and continuous uninterrupted coverage of the whole Earth environment and its surrounding atmosphere. Such detailed observations are necessary in order to monitor often-subtle environmental changes. Microwave remote sensing is highly unique, in that it is the only technology which responds to the absolute volume of water contained in the environment. Radiometric temperature readings in the microwave region have been shown to yield information on soil moisture, precipitable water in clouds, precipitation, and snow. Remote sensing technology is central to the integration of the many interrelated but highly variable point scale phenomena to more useful, regionally-oriented land surface processes.

A new approach for retrieving surface soil moisture from satellite microwave radiometer data is currently being tested. The approach uses radiative transfer theory to solve for surface soil moisture. The methodology is unique, in that it does not require any field observations of soil moisture or canopy biophysical properties for calibration purposes and is totally independent of wavelength. The procedure is being tested with historical 6.6 GHz brightness temperature observations from the Scanning Multichannel Microwave Radiometer (SMMR) (Gloersen and Barath, 1977). Some background and theory of microwave remote sensing are presented, along with some preliminary results from several validation studies. Plans to develop a 20-year retrospective global database of surface soil moisture are discussed. This data product will be made available to the general public, through the Goddard Space Flight Center Distributed Active Archive Center (DAAC). It is hoped that these data will help assess seasonal, inter-annual, decadal, and longer climate variations and changes.

Theory and Background

Passive microwave remote sensing measures the natural thermal radiation from the land surface in the centimeter wave band, the magnitude of which is determined by the physical temperature and the emissivity of the radiating body. In the microwave region, the emitted radiation is extremely low as compared with longwave infrared radiation. An approximation for the Planck equation, at low frequencies ($f < 117$ GHz), is the Rayleigh-Jeans approximation, and can be shown to lead to

$$T_b \cong e_S T \quad (1)$$

where T_b is the microwave brightness temperature, T is the physical temperature of the emitting layer, and e_S is the smooth-surface emissivity. Emissivity is further defined as

$$e_S = (1 - R_S) \quad (2)$$

where R_S is the smooth-surface reflectivity. While the emissivity is lower at horizontal polarization, the sensitivity to changes in surface moisture is significantly greater than at vertical polarization (Figure 1). Conversely, at vertical polarization, the sensitivity to surface temperature is greater, subsequently forming the basis for a surface temperature estimation procedure (Owe and Van de Griend, 2001), which is discussed later. A more thorough treatment of electromagnetic theory may be found in Ulaby et al. (1986).

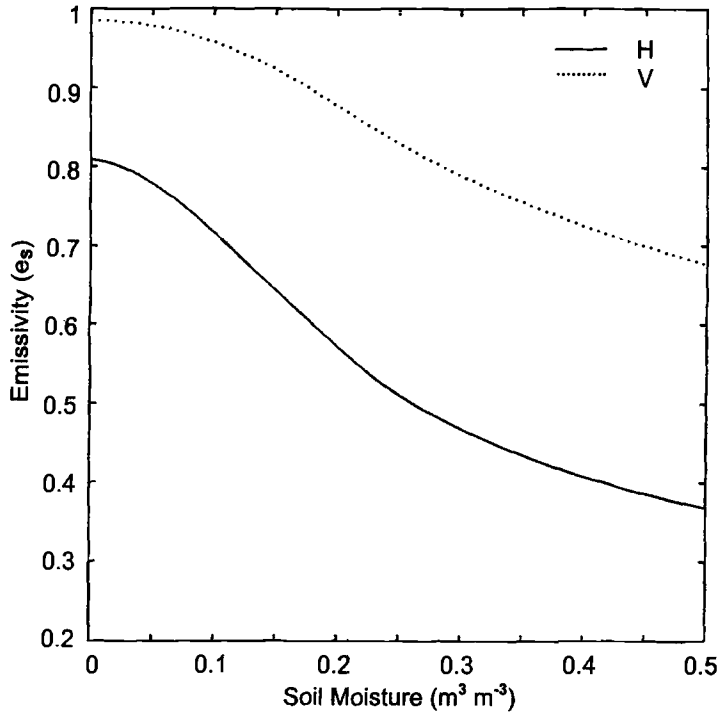


Figure 1 – *Comparison of soil emissivities at H and V polarization at a frequency of 6.6 GHz and an incidence angle of 50 degrees.*

Dielectric Constant

The microwave region is the only part of the electromagnetic spectrum that permits truly quantitative estimates of soil moisture using physically based expressions such as radiative transfer models. Microwave technology is also the only remote sensing

method that measures a direct response to the absolute volume of water in a medium. The basis for microwave remote sensing of soil moisture follows from the large contrast in dielectric constant of dry soil (~4) and water (~80) and the resulting dielectric properties of soil-water mixtures (~4 - 40) and their effect on the natural microwave emission from the soil (Schmugge, 1985). The dielectric constant is a complex number, containing a real (k') and an imaginary (k'') part (Figure 2). The real part determines the propagation characteristics of the energy as it passes upward through the soil, while the imaginary part determines the energy losses. The dielectric constant is a difficult quantity to measure in the field. Moreover, reproducing precise field conditions in laboratory soil samples makes laboratory analyses of the dielectric constant not entirely straightforward. Consequently, the validation of theoretical calculations is often difficult. Two dielectric models, which are commonly used in theoretical calculations, are the Dobson Model (Dobson et al., 1985) and the Wang-Schmugge Model (Wand and Schmugge, 1980). Schmugge (1985) presents an excellent review of basic microwave theory.

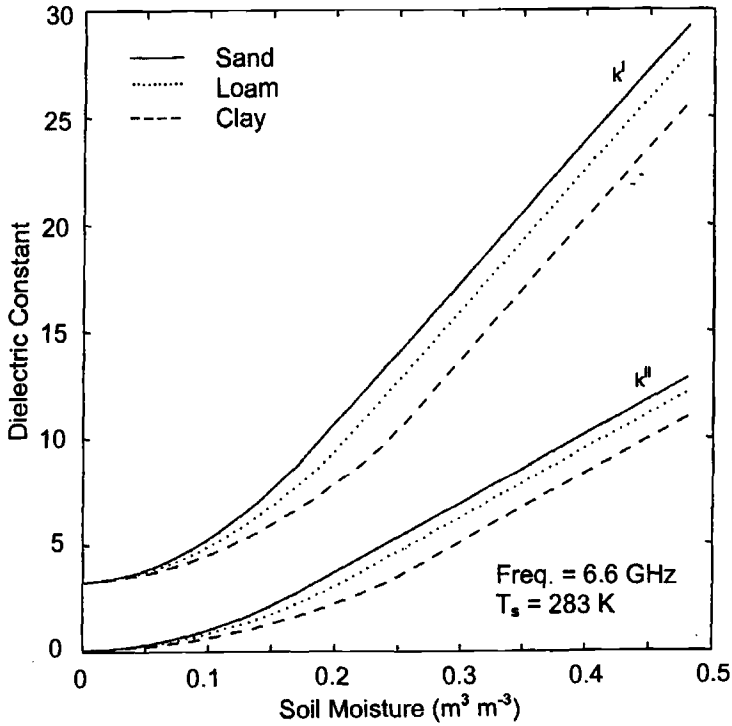


Figure 2 – *Comparison of soil dielectric constants for three generic soils.*

Soil Physical Properties

The complex dielectric constant of a soil results from a combination of the individual dielectric constants of its component parts (i.e. air, water, rock, etc.), and associated factors such as temperature, salinity, textural composition, and frequency. The relationship between the soil dielectric constant and the moisture content is almost linear, except at low

moisture contents (Figure 2). This non-linearity at low moisture contents is due to the strong bonds which develop between the surfaces of the soil particles and the thin films of water which surround them. These bonds are so strong at low moisture levels, that the free rotation of the water molecules is impeded. This water is often referred to as bound water. Therefore, in a relatively dry soil, the water is tightly bound and contributes little to the dielectric constant of the soil water mixture. As more water is added, the molecules are further from the particle surface and are able to rotate more freely. This is referred to as the free water phase. The subsequent influence of free water on the soil dielectric constant therefore also increases. Smaller particles such as irregular fine sands, silts, and clays have a higher surface area-to-volume ratio and therefore are able to hold more water molecules at higher potentials. The unique plate-like structure of clays provides an additional source of high-energy bonds and increases the soil's affinity for water. Two soils with different textural composition will exhibit different relationships between moisture content and their respective soil dielectric constants. Soils with a high clay content will generally have a lower dielectric constant than coarse sandy soils at the same moisture content, since more water is being held in the bound water phase (Figure 2).

Vegetation Effects

The effects of vegetation on the microwave emission as measured from above the canopy is two-fold. Vegetation absorbs and scatters the emitted soil radiation, but also emits its own radiation. In areas of dense canopy, the emitted soil radiation will be sufficiently attenuated, and the observed emissivity will be due largely to the vegetation alone. The magnitude of the absorption by the canopy depends upon the wavelength and the water content of the vegetation. The most frequently used wavelengths for soil moisture sensing are in the L- and C-bandwidths ($\lambda \cong 21$ cm and 5 cm), although only L-band sensors are able to penetrate vegetation of any significant density. While observations at all frequencies are subject to scattering and absorption and require some correction if the data are to be used for soil moisture retrieval, shorter wave bands are especially susceptible to vegetation influences.

Numerous canopy models have been developed to account for the effects of vegetation (Ulaby et al., 1986; Kirdiashev et al., 1979; Mo et al., 1982; Theis and Blanchard, 1988). These basic models have been modified and applied successfully by a variety of investigators, using data from primarily ground-based radiometer systems over agricultural fields (Jackson et al., 1982; Pampaloni and Paloscia, 1986; Jackson and O'Neill, 1990). Radiative transfer characteristics of vegetation can be expressed in terms of the transmissivity, Γ, and the single scattering albedo, ω. The transmissivity is defined in terms of the optical depth τ, such that

$$\Gamma = \exp(-\tau/\cos u) \tag{3}$$

The optical depth is related to the canopy density, and for frequencies less than 10 GHz, has been shown to be linear function of vegetation water content. Typical values of τ for agricultural crops have generally been given as less than one (Mo et al., 1982; Jackson and O'Neill, 1990). Theoretical calculations show that the sensitivity of above-canopy brightness temperature measurements to variations in soil emissivity decreases

with increasing optical depth or canopy thickness (Ulaby et al., 1986). This is because the soil emission is attenuated by the canopy and emission from the vegetation canopy tends to saturate the signal with increasing optical depth. This subsequently results in decreased sensor sensitivity to soil moisture variations. A transmissivity of 1 corresponds to an optical depth of 0, indicating bare soil, or at least no attenuation of the soil-emitted radiation due to an overlying canopy. Conversely, a transmissivity of 0 indicates an infinitely thick canopy, with no penetration of the soil emission through the canopy.

The single scattering albedo describes the scattering of the soil emissivity by the vegetation. The scattering albedo is a function of plant geometry, and may vary according to plant species and associations. Experimental data for this parameter are limited, and values for selected crops have been found to vary from 0.04 to about 0.12 (Mo et al., 1982; Jackson and O'Neill, 1990; Brunfeldt and Ulaby, 1984). Few values have been derived for natural vegetation, although Becker and Choudhury (1988) estimated a value of 0.05 for a semi-arid region in Africa. Van de Griend and Owe (1994) calculated a 3-year time series of both scattering albedo and canopy optical depth at both 6.6 GHz and 37 GHz for savannas of Botswana. The optical depth displayed a distinct seasonal course at both frequencies, although the values for 37 GHz were significantly higher. While the scattering albedo demonstrated considerable variability during the 3-year period, a relationship with vegetation biomass or other seasonal indicators was not observed. An average value for the scattering albedo of 0.076 was found for both frequencies.

Physical Surface Temperature

Satellite microwave observations are generally expressed as brightness temperatures, and must be normalized by the physical temperature of the surface soil as indicated in Eq. (1). The surface layer known as the soil moisture sampling depth is generally thought to determine the surface emissivity (Schmugge, 1983), and it is the temperature of this layer that should be used to normalize the observed satellite brightness. It has been noticed (Owe et al., 1982) that in semi-arid regions, nighttime brightness temperatures display a significantly higher response to variations in surface moisture content than daytime brightness temperatures. Several factors may account for this phenomenon. First, daytime surface heating is extremely high and variable in these regions. This usually causes severe drying of the surface layer. Such intense heating also increases the difficulty in making accurate spatially representative surface temperature estimates. At night, some moisture is restored to the surface as it regains some equilibrium with the underlying soil. Additionally, the temperatures of the air, soil surface, and canopy tend to approach some equilibrium, making it somewhat easier to estimate spatially averaged surface temperature at night.

Surface Roughness

Surface roughness increases the emissivity of natural surfaces, and is caused by increased scattering due to the increase in surface area of the emitting surfaces (Schmugge, 1985). Roughness also reduces the sensitivity of emissivity to soil moisture variations, and thus reduces the range in measurable emissivity from dry to wet soil

conditions (Wang, 1983). An empirical roughness model was developed by Choudhury et al. (1979), and is described as

$$e_r = 1 - R_0 \exp(-h\cos^2 u) \quad (4)$$

where h is an empirical roughness parameter, related to the root mean square (rms) height variation of the surface and the correlation length, and u is the incidence angle of the observation. Typical values for h have been suggested, ranging from 0 for a smooth surface, 0.3 for a disked field, to 0.5 for a rough plowed field.

The effect of roughness on the observed microwave brightness at 6.6 GHz for a range of surface moistures is illustrated in Figure 3. A change in the roughness parameter from 0 to 0.3 corresponds to a difference in the surface emissivity, of about 0.005 at dry conditions, to about 0.014 at saturation. There is some speculation that the effect of surface roughness is less pronounced at satellite scales, except in areas of mountainous terrain or extreme relief.

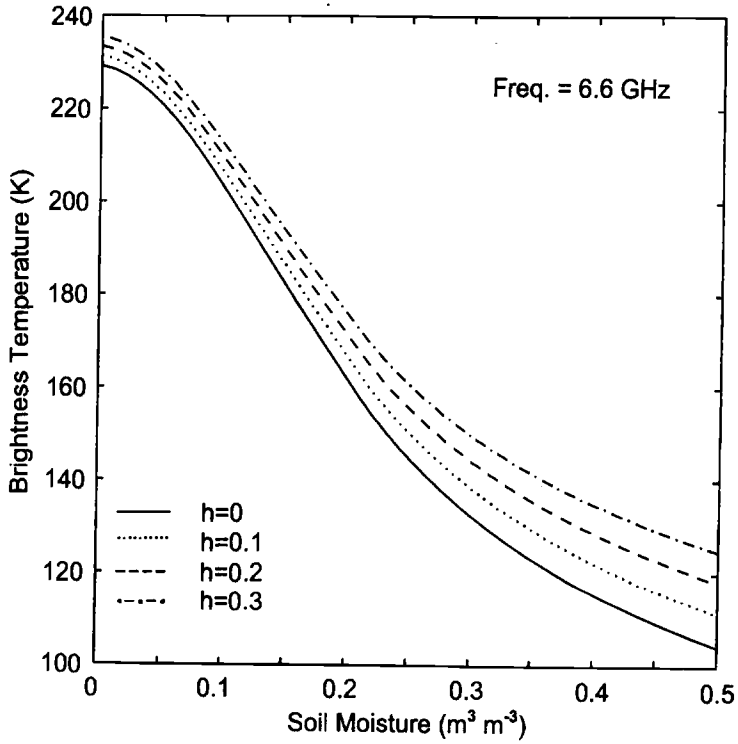


Figure 3 – *Effect of surface roughness on the brightness temperature.*

Sensor Characteristics

Emission and sensor characteristics, such as wavelength, polarization, and incidence angle also have an effect on the observed signal. A longer wavelength is able to penetrate a thicker canopy, and also has a greater soil moisture sampling depth. An

increase in the incidence angle increases the path length through the canopy, so consequently, the optical depth will be greater than at a lower incidence angle for the same canopy. While there is some evidence that differences in the transmissivity at H and V polarization are dependent on incidence angle, these differences are observed mainly over vegetation elements that exhibit some systematic orientation such as vertical stalks in tall grasses, grains, and maize (Ulaby et al., 1986; Van de Griend and Owe, 1994; Van de Griend et al., 1996). In general, the canopy and stem structure for most crops and naturally occurring vegetation are randomly oriented, especially at satellite scales. It is reasonable to assume that the leaf absorption loss factor is for the most part polarization independent. The tendency of vegetation to reduce the polarization difference with increasing biomass is the basis for the Microwave Polarization Difference Index (MPDI) (Becker and Choudhury, 1988).

Microwave Model

A new approach for microwave soil moisture retrieval has recently been developed, and is briefly described. The technique solves simultaneously for the surface soil moisture and vegetation optical depth, using a simple radiative transfer equation (Mo et al., 1982). The radiation from the land surface as observed by a satellite sensor may be expressed in terms of the radiative brightness temperature, T_b, and is given as

$$T_{b(P)} = T_S e_{r(P)} \Gamma_{(P)} + (1-\omega_{(P)}) T_C (1-\Gamma_{(P)}) + (1-e_{r(P)}) (1-\omega_{(P)}) T_C (1-\Gamma_{(P)}) \Gamma_{(P)} \quad (5)$$

where P refers to either H or V polarization, T_S and T_C are the thermometric temperatures of the soil and the canopy respectively, ω is the single scattering albedo. The first term of the above equation defines the radiation from the soil as attenuated by the overlying vegetation. The second term accounts for the upward radiation directly from the vegetation, while the third term defines the downward radiation from the vegetation, reflected upward by the soil and again attenuated by the canopy.

For purposes of simplicity, it was assumed that surface roughness would have a minimal effect over the study sites, and was set to 0, while an average value of 0.06 was used for the scattering albedo, based on previously described studies. The surface temperature of the SMMR footprint was estimated from vertically polarized 37 GHz brightness temperatures (Owe and Van de Griend, 2001).

Brightness temperature measured from space contains information on both the canopy and soil emissions and their respective physical temperatures (Eq. 1). Polarization ratios, such as the MPDI, are frequently used to remove the temperature dependence of T_b, resulting in a parameter that is quantitatively, and more highly, related to the dielectric properties of the emitting surface(s). The MPDI is defined as

$$MPDI = (T_{b(V)} - T_{b(H)})/(T_{b(V)} + T_{b(H)}) \quad (6)$$

At 37 GHz frequency, MPDI is mainly a function of the overlying vegetation, and has been shown to be a good indicator of vegetation biomass (Becker and Choudhury, 1988). At a frequency of 6.6 GHz, however, the MPDI contains information on both the canopy and the soil emissions. The relationship between MPDI, canopy optical depth

and soil moisture may be shown theoretically, and the optical depth may subsequently be defined in terms of the MPDI and the soil dielectric constant.

Eq. (5) is subsequently solved by a non-linear iterative procedure in a forward approach to derive the brightness temperature, by optimizing on the dielectric constant. Once convergence of the modelled and observed horizontal brightness temperatures is achieved, the model uses soil information on particle size distribution, porosity, and wilting point from a data base of soil physical characteristics from the Land Data Assimilation System (LDAS) (Houser et al., 2000), together with the Wang-Schmugge dielectric model, to solve for the surface soil moisture.

Results of SMMR Retrievals

The model was applied to nighttime Nimbus-7 SMMR brightness temperature data for several test sites, that were selected because of the availability of long term observational soil moisture data that can be used for comparison (Robock et al., 2000). The observational data are reported typically as point measurements of average volumetric soil moisture in the surface profile, usually the top 10 cm. While these types of soil moisture data are not the most optimum for validating microwave soil moisture retrievals, they are among the few existing global data sets that cover large areas for extended periods. Models results of soil moisture retrievals over test sites in the U.S. and Central Asia are illustrated as annual time series in Figure 4. The top two time series also illustrate precipitation for the year, for additional comparison.

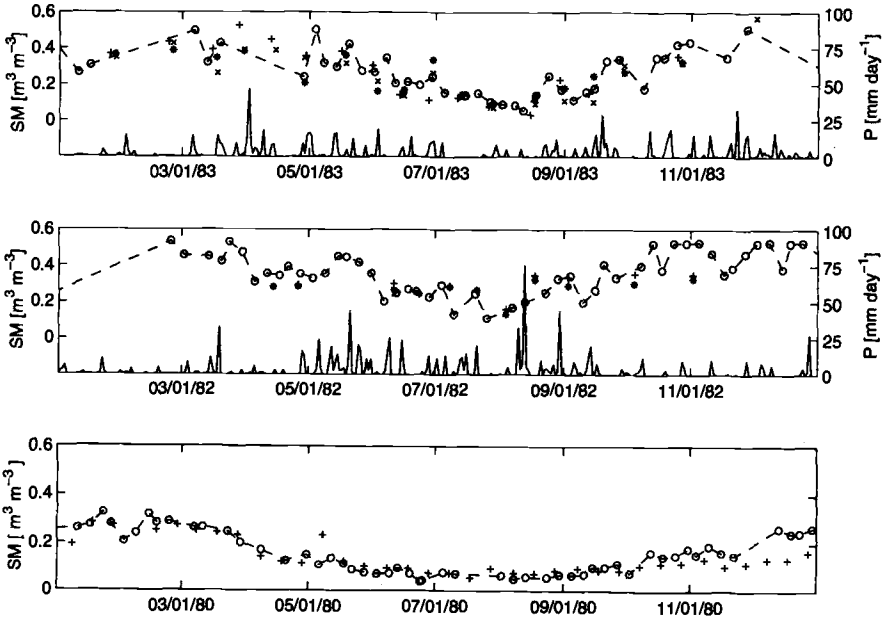


Figure 4 – *Time series of satellite derived surface soil moisture (°) and ground observations (+, *) for study sites in Illinois (1983), Iowa (1982), and Turkmenistan (1980).*

One must always keep in mind, however, several important differences when comparing the satellite-derived surface moistures with the ground observations.

- Differences in spatial resolution – The SMMR-derived surface moisture is an average value integrated over the entire footprint, whereas the observational data are point measurements.
- Differences in vertical resolution – the observational data are an average soil moisture within the top 10 cm profile, while the SMMR retrievals reflect only the moisture content of the microwave soil moisture sampling depth, which is at most only about 1 cm.
- Differences in acquisition times – Ground and satellite observations rarely occur on the same day.
- Inter-observation periods –While the SMMR observations are displayed with connecting lines, it is done so only to help in observing general trends in the time series. It is important to realize that significant changes in surface moisture frequently occur during the periods between observations, but may go totally undetected by both the satellite and the ground observations. Daily precipitation is also included in the time series to assist in understanding the observed changes in soil moisture.

Discussion and Conclusions

The importance of soil moisture in hydrological modelling and global change studies is acknowledged by most researchers. Studies have consistently shown that regular spatially representative estimates of soil moisture will greatly enhance the accuracy of crop productivity forecasts, flood and drought forecasting, numerical weather prediction models, and other large-scale land surface interaction models. While precipitation and other climate data have been monitored globally for decades, historical soil moisture observations typically exist only on a very localized scale, and usually only for very short periods of time. However, a number of data sets have recently been identified, primarily those catalogued in the Global Soil Moisture Data Bank (Robock et al., 2000), that cover greater spatial and temporal scales. While these data sets are by no means optimal, they continue to be invaluable to researchers as a validation tool, for both land surface models and satellite retrieval algorithms.

Over the last 20 years, efforts have been underway to use remote sensing as a more consistent and cost efficient way to monitor the Earth and conduct regular observations of various surface and atmospheric parameters. Certain parameters, such as vegetation biomass, snow cover, snow water equivalent, and land use classification, are measured and reported regularly, and often in near real time, through operational satellite monitoring programs. In many instances, historical global data-bases have been compiled, and have proven invaluable to researchers, especially for global climate and land-process studies.

The retrieval of soil moisture by satellite monitoring is especially complex, because of the many parameters that influence the signal observed by the sensor. Quantifying these parameters is not always straight-forward because of the often complex physical relationships which define them. Past retrieval efforts typically have solved for the vegetation component as a residual over calibration sites where spatially representative

surface soil moisture has been estimated. Empirical relations and surrogate parameters are then used to quantify the vegetation. However, such calibration procedures are usually quite error-prone, because of the lack of optimum soil moisture observational data. Several recent research efforts have developed procedures, which attempt to partition the observed radiation into its primary sources either by numerical methods such as simultaneous equations or other optimization techniques (Njoku and Li, 1999).

A methodology has been presented, that retrieves pixel average surface soil moisture and vegetation optical depth from dual polarized microwave brightness temperature observations, and has been applied to the 6.6 GHz SMMR data. The radiative transfer-based approach does not use ground observations of soil moisture, canopy measurement data, or other regional geophysical data as calibration parameters, and is totally independent of frequency. Only nighttime data was used in the study because of the greater stability of nighttime surface temperatures. Days with snow cover or when surface temperatures were below zero were eliminated from the analysis. A non-linear iterative approach is used to solve for the surface moisture.

Preliminary results from validation studies over several test sites were presented. These sites were used because of the availability of long-term soil moisture observations for comparison. Time series of the satellite-derived surface moisture compared well with the available ground observations and precipitation data. Validation studies in other regions are currently being conducted. Unfortunately, reliable, spatially averaged surface moisture data, which can be used for validation purposes are rare to non-existent, especially at satellite scales. Comparison with precipitation fields is often the only validation option available. Field experiments with the express purpose of gathering such data should be designed and implemented as a research priority. Refinements to a 37 GHz daytime surface temperature retrieval algorithm will be completed shortly, allowing for daytime soil moisture retrievals to be conducted, with the eventual goal of generating a complete retrospective daytime and nighttime global surface moisture dataset for the entire SMMR period.

Plans are currently underway, to develop a more than 20-year database of historical global surface soil moisture. Because the Nimbus-7 SMMR data are available only through mid-1987, the data will be extended with Defense Satellite Mapping Program (DSMP) Special Sensor Microwave/Imager (SSM/I) observations from the beginning in August 1987. However, the lowest frequency on the SSM/I sensor is 19 GHz, which corresponds to a wavelength of only about 1.6 cm. The effect of this is twofold. First, the soil penetration depth at this frequency is only about 0.5 cm, and second, the ability to penetrate overlying vegetation is also significantly reduced. While the SSM/I data product will be somewhat degraded, it should still be extremely useful for hydrological modelling purposes. The data can again revert back to C-band, after the launch of the new EOS Aqua satellite, currently scheduled for early 2002. The database will be in the form of daily global maps of soil moisture, and will be accompanied by average monthly error maps. These global maps will provide an estimated pixel-by-pixel error, based on the calculated vegetation optical depth. The data product will be made available to the general public through the Goddard Space Flight Center Distributed Active Archive Center (DAAC).

References

Becker, F. and Choudhury, B. J., 1988, "Relative Sensitivity of Normalized Difference Vegetation Index (NDVI) and Microwave Polarization Difference Index (MPDI) for Vegetation and Desertification Monitoring," *Remote Sensing Environment*, Vol. 24, pp. 297-311.

Brunfeldt, D. R. and Ulaby, F. T., 1984, "Measured Microwave Emission and Scattering in Vegetation Canopies," *IEEE Transactions Geoscience and Remote Sensing*, Vol. 22, pp. 520-524.

Chahine, M. T., 1983, "Interaction Mechanisms Within the Atmosphere," *Manual of Remote Sensing*, edited by R.N. Colwell, pp. 165-230.

Choudhury, B. J., Schmugge, T. J., Chang, A. T. C., and Newton, R. W., 1979, "Effect of Surface Roughness on the Microwave Emission from Soils,"*Journal of Geophysical Research*, Vol. 84, pp. 5699-5705.

Dobson, M. C., Ulaby, F. T., Hallikainen, M. T., and El-Rayes, M. A., 1985, "Microwave Dielectric Behavior of Wet Soil - Part II: Dielectric Mixing Models," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 23, pp. 35-46.

Geiger, R., 1950, *"Climate Near the Ground"*, Harvard University Press, Cambridge, MA.

Gloersen, P. and Barath, F. T., 1977, "A Scanning Multichannel Microwave Radiometer for Nimbus-G and SeaSat-A," *IEEE Journal of Oceanic Engineering*, Vol. 2, pp. 172-178.

Houser, P. R., Cosgrove, B. A., Entin, J. K., and Desetty, M., 2000, "Land Data Assimilation System," http://ldas.gsfc.nasa.gov/index.shtml, NASA/Goddard Space Flight Center, 2000.

Jackson, T. J. and O'Neill, P. E., 1990, "Attenuation of Soil Microwave Emission by Corn and Soybeans at 1.4 and 5 GHz," *IEEE Transactions Geoscience and Remote Sensing*, Vol. 28, pp. 978-980, 1990.

Jackson, T. J., Schmugge, T. J., and Wang, J. R., 1982, "Passive Microwave Sensing of Soil Moisture Under Vegetation Canopies," *Water Resources Research*, Vol. 18, pp. 1137-1142.

Kirdiashev, K. P., Chukhlantsev, A. A., and Shutko, A. M., 1979, "Microwave Radiation of the Earth's Surface in the Presence of Vegetation Cover," *Radio Engineering Electronics*, Vol. 24, pp. 256-264.

Marsh, G. P., 1874, *"The Earth as Modified by Human Action,"* Scribner and Armstrong Publishers, NY.

Mather, J. R., 1954, "The Present Climatic Fluctuation and its Bearingon a Reconstruction of Pleistocene Climatic Conditions,, *Tellus*, Vol. 6, No. 3, pp. 287-301.

Mo, T., Choudhury, B. J., Schmugge, T. J., Wang, J. R., and Jackson, T. J., 1982, "A Model for Microwave Emission From Vegetation-Covered Fields," *Journal of Geophysical Research*, Vol. 87, pp. 11,229-11,237.

Njoku, E. G. and Li, L, 1999, "Retrieval of land surface parameters using passive microwave measurements at 6-18 GHz," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, pp. 79-93, Jan. 1999.

Owe, M. and Van de Griend, A. A., 2001, "On the Relationship Between Thermodynamic Surface Temperature and High Frequency (37 GHz) Vertical Polarization Brightness Temperature Under Semi-Arid Conditions," *International Journal of Remote Sensing* (In Press).

Owe, M., Van de Griend, A. A., and Chang, A. T. C., 1982, "Surface Moisture and Satellite Microwave Observations in Semiarid Southern Africa," *Water Resources Research.*, Vol. 28, No. 3, pp. 829-839.

Pampaloni, P. and Paloscia, S., 1986, "Microwave Emission and Plant Water content: a comparison between field measurements and theory," *IEEE Geoscience and Remote Sensing*, Vol. 24, pp. 900-905.

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., Liu, S., and Namkhai, A., 2000, "The Global Soil Moisture Data Bank," Bulletin American Meteorological Society, Vol. 81, pp. 1281-1299

Schmugge, T. J., 1983, "Remote Sensing of Soil Moisture: Recent Advances," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 21, pp. 336-344.

Schmugge, T. J., 1985, "Remote Sensing of Soil Moisture," *Hydrological Forecasting*, edited by M.G. Anderson and T.P. Burt, John Wiley, New York.

Sellers, P. J., Mintz, Y., Sud, Y. C., and Dalcher, A., 1986, "A Simple Biosphere Model (SiB) for Use Within General Circulation Models," *Journal of Atmospheric Science*, Vol. 43, pp. 505-531.

Theis, S. W. and Blanchard, B. J., 1988, "The Effect of Measurement Error and Confusion From Vegetation on Passive Microwave Estimates of Soil Moisture," *International Journal of Remote Sensing*, Vol. 9, pp. 333-340.

Thornthwaite, C. W., 1956, "Modification of Rural Microclimates," *Man's Role in Changing the Face of the Earth*, W.L. Thomas, Jr., Ed., University of Chicago Press, Chicago, IL.

Ulaby, F. T., Moore, R. K., and Fung, A. K., 1982, *Microwave Remote Sensing Active and Passive*, Vol. I., Artech House, Boston, MA.

Ulaby, F. T., Moore, R. K., and Fung, A. K., 1986, *Microwave Remote Sensing Active and Passive*, Vol. III., Artech House, Boston, MA.

Van de Griend, A. A. and Owe, M., 1994, "Microwave Vegetation Optical Depth and Inverse Modelling of Soil Emissivity Using Nimbus/SMMR Satellite Observations," *Meteorology and Atmospheric Physics*, Vol. 54, pp. 225-239.

Van de Griend, A. A., Owe, M., De Ruiter, J., and Gouweleeuw, B. T., 1996, "Measurement and Behavior of Dual-Polarization Vegetation Optical Depth and Single Scattering Albedo at 1.4 and 5 GHz Microwave Frequencies," *IEEE Transactions Geoscience and Remote Sensing*, Vol. 34, pp. 957-965.

Wang, J. R. and Schmugge, T. J., 1980, "An Empirical Model for the Complex Dielectric Permittivity of Soil as a Function of Water Content," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 18, pp. 288-295.

Wang, J. R., 1983, "Passive Microwave Sensing of Soil Moisture Content: The Effects of Soil Bulk Density and Surface Roughness," *Remote Sensing of Environment*, Vol.13, pp. 329-344.

Addressing Issues of Uncertainty and Risk in Geospatial Applications

Robert G. Knowlton,[1] David M. Peterson,[1] and Hubao Zhang[1]

The Use of Decision Support Systems to Address Spatial Variability, Uncertainty, and Risk

Reference: Knowlton, R. G., Peterson, D. M., and Zhang, H., **"The Use of Decision Support Systems to Address Spatial Variability, Uncertainty, and Risk,"** *Spatial Methods for Solution of Environmental and Hydrologic Problems – Science, Policy and Standardization, ASTM STP 1420,* D. T. Hansen, V. H. Singhroy, R. R. Pierce, and A. I. Johnson, Eds., ASTM International, West Conshohocken, PA, 2003.

Abstract:

Traditional methods of characterizing contaminated waste sites and evaluating cleanup alternatives generally utilize conservative methods that may not produce optimal results. With the advent of powerful desktop computers, advanced database management tools, sophisticated graphical display capabilities, new statistical methods, as well as decision analysis methods, there is a greater opportunity to employ decision support systems to address spatial variability, uncertainty, sampling efficiency, risk and cost-benefit needs. Several decision support systems have been developed in the past few years that can address these needs directly, and help decision-makers evaluate their environmental liabilities and alternatives for action in a more efficient manner. EPA recognized the value of these new tools in the decision making process and instituted a review of Decision Support Systems. The EPA's Environmental Technology Verification (ETV) program is designed to peer-review innovative technologies for acceptability of use. ETV then publishes these reviews, along with verification certificates, to facilitate more rapid acceptance of the technologies. Two decision support tools that were evaluated by the ETV program, Sampling*FX*[2] and Groundwater*FX*[2], will be discussed in this paper, along with examples of the use of the tools for decision making purposes. The Sampling*FX* toolkit utilizes geostatistical analysis techniques and operations research methods to quantify uncertainty in the nature and extent of soil contamination, as well as optimizing the number and location of samples required for characterization. The Groundwater*FX* toolkit utilizes Monte Carlo simulation techniques and operations research methods to quantify uncertainty in the nature and extent of groundwater contamination, as well as

[1] Duke Engineering and Services, Inc., 6501 Americas Parkway, NE, Suite 810, Albuquerque, NM 87110
[2] Software by Decision*FX*, Inc., Bosque Farms, NM

optimizing the number and location of monitoring wells required for characterization, and evaluating groundwater remediation strategies.

Keywords: geostatistics, groundwater, modeling, decision analysis, decision support, uncertainty, environmental, hydrologic, operations research, sampling design

Introduction

The United States has made significant progress in characterizing and cleaning up contaminated properties over the past several decades. The cost of cleanup of our nation's contaminated lands is quite significant. The Office of Technology Assessment has estimated there may be 439000 hazardous waste sites in the country potentially requiring remediation (U.S. Congress, 1989). As of December 31, 1996, there were 1,296 uncontrolled hazardous waste sites listed on the EPA's National Priorities List (NPL) for compliance under CERCLA (Johnson and DeRosa, 1997). There are between 1500 and 3500 waste sites requiring corrective action under RCRA at treatment, storage, and disposal facilities in the U.S. (Johnson and DeRosa, 1997). The estimated total cost of cleanup between 1990 and 2020 for all waste sites in the U.S. is between $500 billion and $1 trillion (Johnson and DeRosa, 1997). Of this amount, between $106 billion and $302 billion is attributed to CERCLA sites (Johnson and DeRosa, 1997). Of course, the cleanup process will not be complete by the year 2020, and the total cost of compliance may continue to rise. These estimates assume that the standards for cleanup limits and allowable risk levels remain the same. Flexibility in cleanup goals could greatly reduce these cost estimates. There is still considerable work to be done, and new methods of analysis may help trim costs compared to conventional approaches.

Defining Nature and Extent of Contamination

One of the areas of great interest to regulatory authorities is the definition of the nature and extent of contamination, whether in soil or in groundwater. The main concern in defining the nature and extent of contamination is two fold:

- Is the sampling network design adequate in terms of numbers and locations?
- Is the interpolation between data points accurate?

The first issue has traditionally been addressed through the use of conservative statistical sampling protocols. The second issue was traditionally addressed with confirmatory sampling. In both cases there was a tendency to prescribe significant numbers of samples at potentially considerable expense.

The classic statistical approaches, which the EPA generally promotes, are based on statistical measures of the soil concentration data set and a user-specified confidence in the desired result (EPA, 1994a). The number of samples required to meet specified statistical confidence limits is dependent on the mean and variance of the population set. Spatial variability has no bearing on the prescription for sample size. In addition, the typical approach to defining sample locations is based on the size of the targets (i.e., contaminant plume area, shape, and/or size) or simply the size of the area to be addressed

for a given number of samples derived from the statistical method. These techniques generally prescribe a set sampling protocol, such as a grid or a pattern deployment for sample locations. Again, the actual spatial variability of the contaminant plume is not factored into the design of the sampling network. With larger parcels of contaminated real estate the conventional sampling approaches tend to require large numbers of samples with no bearing on the value of spatial aspects of the data.

More recently there has been a push in the environmental characterization arena to employ 'smarter' strategies to define the nature and extent of contamination. An example of this new approach is embodied in the Expedited Site Characterization (ESC) approach. The ESC technique is ideally suited to real-time characterization efforts (using in-situ field instrumentation or a mobile laboratory) in which optimal locations and numbers of samples are provided through an iterative sampling protocol. Each step in the process helps better define the knowledge base, thereby reducing uncertainties where the variability is greatest. Typically, less samples are required in this type of approach than in a conventional statistical sampling design. The resulting statistics on plume volume can be used directly for cost-benefit analyses of remedial alternatives. This methodology yields information on the uncertainty associated with the nature and extent of contamination, thereby providing decision makers flexibility in defining confidence intervals for cleanup. The uncertainty analysis also provides a sound basis for contingency estimates in the budget process. The technique is also well suited to risk-based decision making, if risk assessment methods or risk-based screening criteria are applied during the data collection process. ASTM has a standard based on ESC concepts for environmental characterization, ASTM Standard Practice for Expedited Site Characterization of Vadose Zone and Ground Water Contamination at Hazardous Waste Contaminated Sites (D6235).

A relatively new approach to optimizing the sampling network design within an ESC process involves the use of geostatistics. Geostatistics is ideally suited to the problem of defining the nature and extent of contamination because it explicitly accounts for spatial variability, whereas the conventional statistical techniques do not. When combined with an efficient operations research methodology to design the placement of sampling locations, the methodology becomes quite robust.

A new decision support tool, called Sampling*FX*[3], has been employed to analyze the nature and extent of contamination at a number of sites. The code employs two main concepts to achieve its intended goals: geostatistical analysis methods for defining spatial variability and uncertainty associated with a contaminant plume; and operations research methods for defining sampling network design (e.g., optimal number and location of samples). The geostatistical routines in Sampling*FX*, as in almost all geostatistical tools (e.g., UNCERT[4], SADA[5], etc.), are derived from the Stanford University suite of tools referred to as GSLIB (Deutsch and Journel, 1998). The difference in each tool is the selection of the individual geostatistical simulators (e.g., kriging, sequential Gaussian simulator, indicator methods, etc.) and the implementation of the output structures (e.g., what kind of maps and statistics to produce). Within Sampling*FX* the preference is to use

[3] Software by Decision*FX*, Inc., Bosque Farms, NM.
[4] Software by the Colorado School of Mines, Golden, CO.
[5] Software by the University of Tennessee, Knoxville, TN.

sequential Gaussian simulation because it simulates the range of uncertainty and variability possible in the distribution of the contamination, whereas techniques such as kriging tend to smooth or average the results, thereby dampening potential high values in the data field.

As an example of the type of analysis that can be performed with a geostatistical approach, the case of a lead-contaminated site at Sandia National Laboratories is presented. The site of interest is a 5-acre firing and testing facility. A conventional EPA-style sampling approach was applied at the site using a star and grid pattern for sampling network design. Two soil concentration thresholds were considered, representing residential and industrial land use exposure scenarios. The geostatistical analysis routines in Sampling*FX* were used to analyze the nature and extent of contamination from this conventional sampling design. Figure 1 shows the probability distribution of exceeding the two threshold limits for the two land use scenarios (i.e., 400 mg/kg for a residential land use and 2000 mg/kg for an industrial land use).

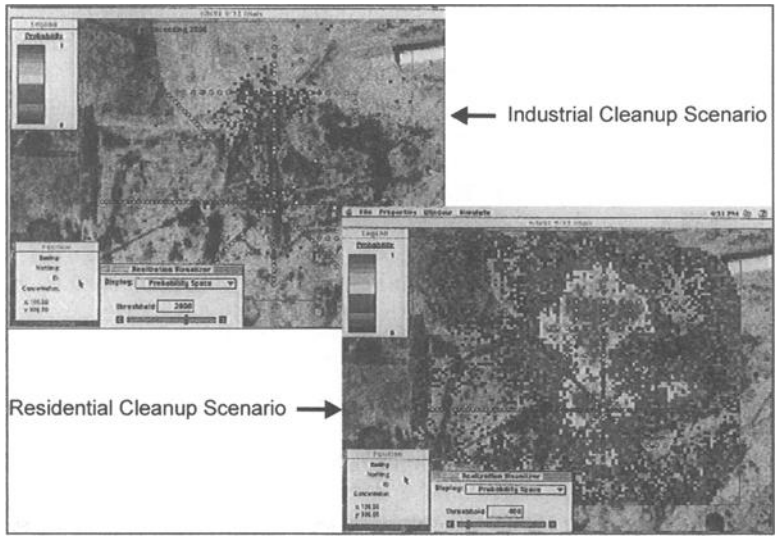


FIG. 1 —*Probability of Exceeding Risk-Based Cleanup Thresholds for Two Different Land Use Scenarios*

The cleanup volumes are markedly different for each land use scenario and the uncertainties are as well. Performing a cost analysis for cleanup of the site resulted in the estimates shown in Figure 2. The uncertainty in the residential cleanup is fairly significant (± $500K) and can be managed by selecting a confidence level for cleanup.

In addition to quantifying the uncertainty in the cleanup volume, the operations research methods in Sampling*FX* were utilized to optimize the sampling network design. The operations research methodology in Sampling*FX* utilizes probability information generated with the stochastic output of the geostatistical code combined with the spatial representation of the variance (e.g., determining where the variability is greatest) to optimally select the numbers and locations of samples needed. With an iterative, adaptive sampling approach just 65 samples would have been sufficient to get the same

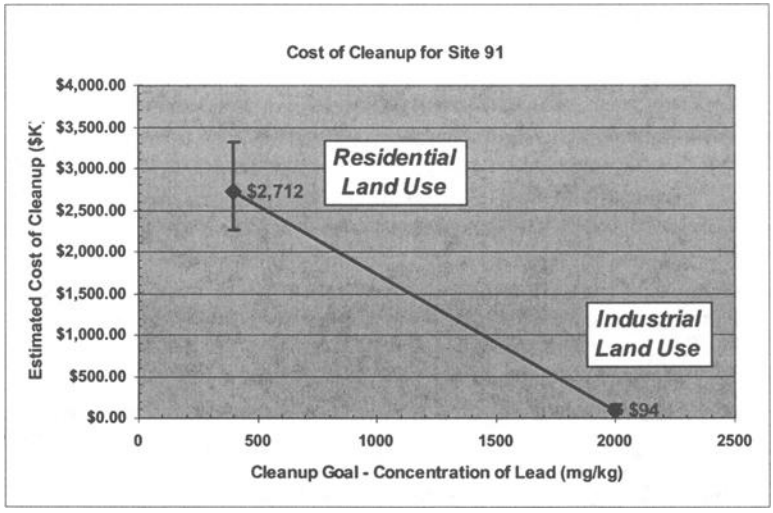


FIG. 2 — *Cleanup Cost Estimates for Two Different Land Use Scenarios.*

level of detail in the plume discretization versus the 350 samples collected with the conventional EPA approach. The logic here is that for each sampling event an analysis is performed, uncertainties estimated, and the new sampling locations chosen to efficiently reduce uncertainties. There reaches a point where additional samples do little to refine the definition of the nature and extent of contamination. In this case the plume statistics are stable after 5 rounds of sampling, with a total of 65 samples collected (as opposed to the 350 samples collected with the baseline approach).

The baseline approach at this site initially used a star pattern for sample network design, followed by a grid-sampling pattern in an area of elevated concentrations. Another traditional EPA approach that can be contrasted with these methods is a straight grid sampling method. Using design criteria from EPA's Data Quality Objectives (DQO) guidance (EPA, 1994b) to estimate the number of samples in a uniform grid yields an estimate of about 650 samples needed to adequately cover the site. Performing a geostatistical analysis of the data from a grid sampling approach, employing 650 samples

throughout the site, yields area estimates that are significantly less than either the baseline approach or the optimal sampling approach using the Sampling*FX* operations research methodology. The cost estimate for a residential cleanup scenario using the EPA uniform grid analysis is on the order of $1.7M, ± $180K. This is in contrast to the baseline and Sampling*FX* estimates of $2.7M, ± $500K. If the uniform grid sampling method were used on this site it is likely that the cleanup volume would be underestimated and confirmatory sampling would have shown the deficiency. With the uniform grid sampling approach, the final cost of cleanup would likely be greater than the baseline due to the need for a second round of mobilization for the cleanup work.

Sampling*FX* has recently been peer reviewed by the EPA's Environmental Technology Verification (ETV) program (EPA, 2000a). EPA has evaluated several decision support systems for appropriateness of use in environmental evaluations.

Groundwater Modeling Advances in Decision Analysis

In the area of groundwater flow and transport analyses there are some significant advances in available modeling codes that allow the analyst to provide uncertainty information to the decision makers. With the advent of the powerful desktop computers on the market today, the ability to perform complex computations without the need for a mainframe has become a reality. These tools can be used to help design monitor well networks and to evaluate remediation options, such as Monitored Natural Attenuation (MNA) (EPA, 1999), Alternate Concentration Limits (ACLs)[6], pump and treat operations, and permeable reactive barrier methods. Two codes are currently available to quantitatively address uncertainties in groundwater flow and transport analyses: Stochastic Groundwater Vistas (GWVistas)[7] and Groundwater*FX*[8] (the commercial version of the DOE's Groundwater Analysis and Network Design Tool, or GANDT [Knowlton et al., 2001]). In this paper, case study results are presented from analyses with the Groundwater*FX* code.

Groundwater*FX* offers the following capabilities:

- Simulation of flow and transport from a source term, through the unsaturated zone, and into the underlying aquifer;
- 3-dimensional numerical flow and transport solution in the saturated zone;
- Capability to simulate spatial variability in hydraulic conductivity distribution, and to honor any aquifer test data that is available;
- Capability to honor observed head data and water quality data in the simulations (essentially a built-in calibration method for ground-truthing the simulations to observed conditions);
- Uncertainty analysis methods (i.e., Latin Hypercube Sampling/Monte Carlo techniques); and
- Operations research method for optimizing the configuration of the monitor well network.

---

[6] Resource Conservation and Recovery Act, 42 U.S.C. § 6901 (1976)
[7] Software by Environmental Simulations International, Herndon, VA.
[8] Software by Decision*FX*, Inc., Bosque Farms, NM.

The code has a user-friendly Graphical User Interface (GUI). It is important to note that in a conventional uncertainty analysis using the Monte Carlo method thousands of model simulations are required to capture the uncertainty, due to the random nature of the sampling process. The Groundwater*FX* code uses an optimization routine called Latin Hypercube Sampling, or LHS, to reduce the number of simulations. The general rule of thumb for the number of model runs required to capture the uncertainty when using LHS is 4/3 times the number of uncertain variables. This reduces the number of modeling runs required in a typical groundwater flow and transport analysis to tens or a couple of hundred runs. Analysis times are quite reasonable, from a couple hours to a day, depending on the complexity of the analysis.

The Groundwater*FX* code has also been evaluated by the EPA's ETV program for Decision Support Systems (EPA, 2000b).

Monitored Natural Attenuation Approach

A Monitored Natural Attenuation (MNA) strategy requires that, within a reasonable time period, concentrations of the contaminants of concern be reduced below regulatory limits, or Maximum Contaminant Levels (MCLs), by natural processes. Several potential MNA processes can be considered:

- hydrodynamic dispersion of the contaminants (e.g., mass spreading and concentration reduction);
- degradation and/or decay (e.g., mass reduction);
- dilution from recharge or infiltration (e.g., areal recharge, stream/irrigation leakage); and/or
- flushing (e.g., discharge to a gaining stream).

The Groundwater*FX* code was employed coupled stream-aquifer modeling approaches to address the groundwater discharge issue. Predicted concentrations in the surface water need to be protective of human health and the environment. MNA is applicable for organic contaminants (e.g., petroleum compounds) and inorganic constituents (e.g., metals). The main difference in processes between organic and inorganic constituents is the potential for degradation. For inorganics, the degradation of contaminants of concern probably has a minimal attenuation effect because biological processes are not very effective in reducing concentrations. Dilution, dispersion, and especially flushing, are the main processes of interest for inorganics. For organic constituents, natural biodegradation processes may be present. It is also possible to augment the biodegradation process with active remedial measures (e.g., injecting hydrogen peroxide). Also, it may be advantageous to utilize a remedial measure (e.g., pump and treat) for some period of time to reduce major concentrations to manageable levels with subsequent reliance on natural attenuation. Such might be the case where major pump and treat actions have been in operation for some time, having reduced contaminant concentrations significantly, but not yielding appreciable decreases over the near-term. Time frames for achieving drinking water quality standards may be lengthy

and therefore costly in this scenario. Natural attenuation or ACLs may be an acceptable alternative for these sites.

Groundwater*FX* utilizes uncertainty analysis techniques to evaluate the likelihood of success of the strategy and to estimate statistical behavior of the future monitoring activities. It is this uncertainty/statistical analysis approach that has won favor within regulatory circles because it provides the range of possible monitoring results as opposed to point estimates, as is typically done with a deterministic approach to modeling a site.

As an example of this type of approach, consider the results of a natural attenuation analysis for a uranium mill tailings facility. The Figure 3 shows the average uranium contaminant plume distribution from an uncertainty analysis performed in 1997. The plume is discharging to the nearby stream and dilution/flushing is the dominant natural

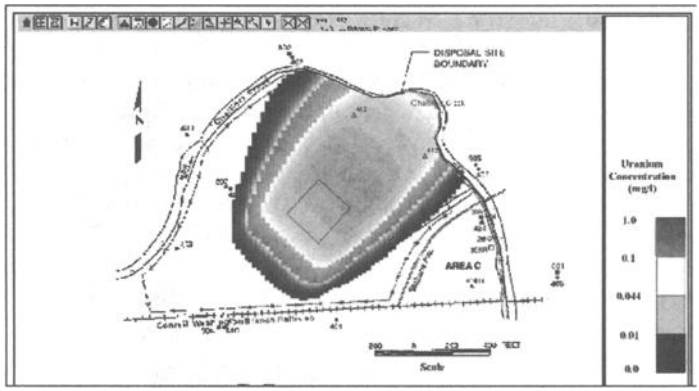


FIG. 3 – *Average Uranium Concentrations in 1997.*

attenuation mechanism. The concentrations in the stream are well within acceptable limits for both human health and ecological concerns. The color contours on the plume are such that the green-to-yellow transition represents the concentration of the MCL. Therefore, the area of yellow-to-orange color is above acceptable limits.

Figure 4 shows the average contaminant plume concentrations 30 years after the previous plot. Through time the contaminants have attenuated to the point where, on average, the concentrations are less than the MCL. However, the likelihood that the site is considered clean is not 100%.

Figure 5 shows the probability distribution at the same time frame as the previous plot; 30 years after the baseline. The interpretation of this information is that there is still a 5 to 10% probability that the concentrations may be above the MCLs at this time. In other words, on average we would expect the site to be cleaned up in 30 years, but there is still a 5 to 10% chance that it will not be within acceptable limits. To achieve essentially 100% likelihood of attaining compliance it would take approximately 5 more years beyond this time. This uncertainty analysis allows the decision maker to plan for contingencies in monitoring duration and costs.

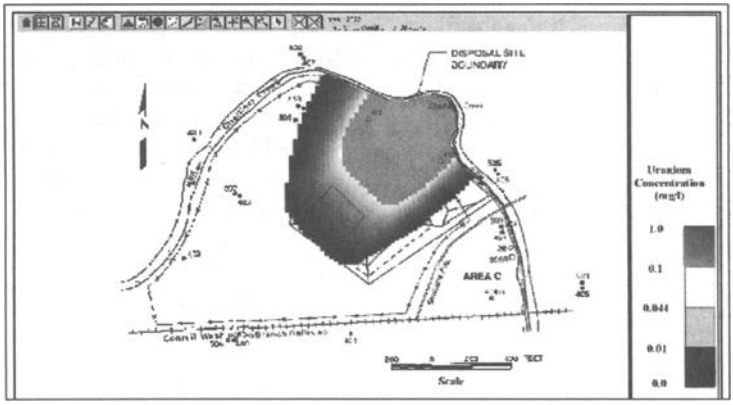


FIG. 4 — *Average Uranium Concentrations 30 Years Into the Future*

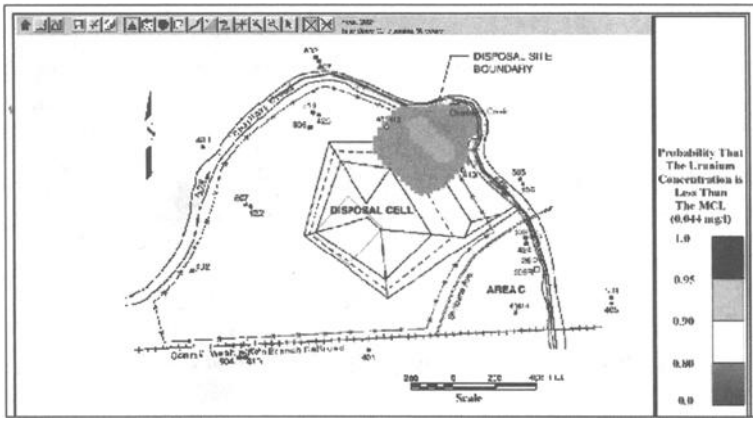


FIG. 5 — *Probability Distribution of Exceeding the MCL 30 Years After the Baseline.*

In addition to the visual depiction of the contaminant plumes just presented, the uncertainty analysis yields a statistical representation of likely concentrations in the monitoring wells through time. Figure 6 shows an example of this type of output. The

power of this type of analysis is that the future monitoring of the site can be compared to the statistical distributions in this plot and as long as observed concentrations are less than the maximums shown in the upper error bars, then the site is on track for natural attenuation. If, however, the concentrations monitored go above the uncertainty estimates then a re-evaluation is in order. If the uncertainties were addressed appropriately in the analysis, then this situation should not occur, and the future monitoring should be within the predicted limits. From a regulatory perspective this is advantageous. In a typical deterministic modeling scenario a calibrated model is used to predict concentrations at the compliance wells, yielding a single value for any given time frame of interest. If the monitored concentrations at a well are slightly above the predicted value some time in the future, it is not clear whether the site is still on track for natural attenuation. With the uncertainty analysis, we are provided likelihood estimates and a comfort range (i.e., the

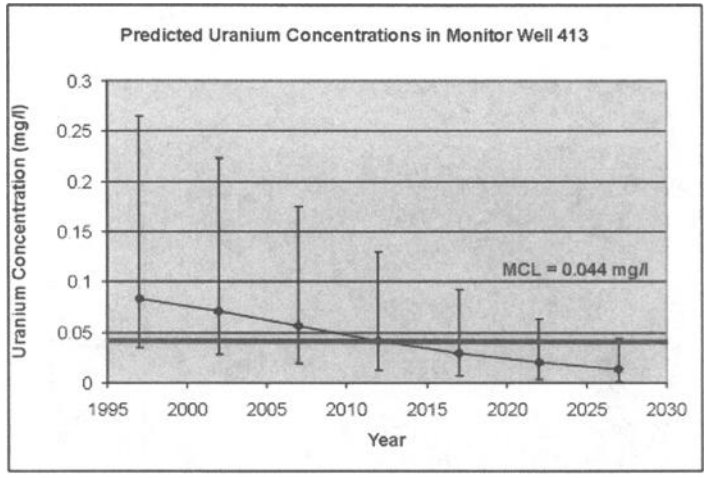


FIG. 6—*Statistical Uncertainty of Water Quality Data in a Monitor Well Through Time*

statistical spread on the predicted concentrations) for evaluating the performance of the remedy.

Uncertainty analyses have been performed to evaluate MNA options for both government and private sector clients. An analysis of the uranium mill tailings site in Riverton, WY resulted in the first natural attenuation remedy approved for a DOE Uranium Mill Tailings Remedial Action (UMTRA) site, with concurrence by the Nuclear Regulatory Commission (NRC) following EPA guidelines and rules for compliance (DOE, 1998). This type of approach is being used on other UMTRA sites, in addition to other uranium-contaminated sites.

Alternate Concentration Limits (ACLs) Approach

Another alternative remedy to active remedial operations is Alternate Concentration Limits (ACLs). The ACL concept is relatively simple: the groundwater beneath the site that is controlled by the site owner need not be cleaned up to drinking water standards, but the groundwater leaving the site at a potential point of exposure has to be safe. Figure 7 illustrates this concept. One or more Point of Compliance (POC) monitoring wells are established just down gradient of the release area, and one or more Point of Exposure (POE) monitoring wells are established at the property boundary, at a surface water discharge location, or the nearest location where a water-supply well might be located. The concentrations in the POC wells may be elevated above drinking water standards, but between the POC and the POE wells natural attenuation processes must reduce the concentrations to acceptable levels. An analysis is performed to evaluate the maximum allowable concentrations in the POC wells that would result in acceptable concentration levels in the POE wells after attenuation. The calculated concentrations in the POC wells are referred to as ACLs. As long as the problem holder maintains control of the property, provides monitoring data to demonstrate that the POE wells are protective, and the POC wells are within predicted ranges, then the ACL remedy may be maintained indefinitely. Groundwater*FX* was applied to evaluate the ACL method at a DOE uranium mill tailings site and approved by the NRC. As mentioned above, an ACL strategy is highly dependent on natural attenuation processes for successful implementation.

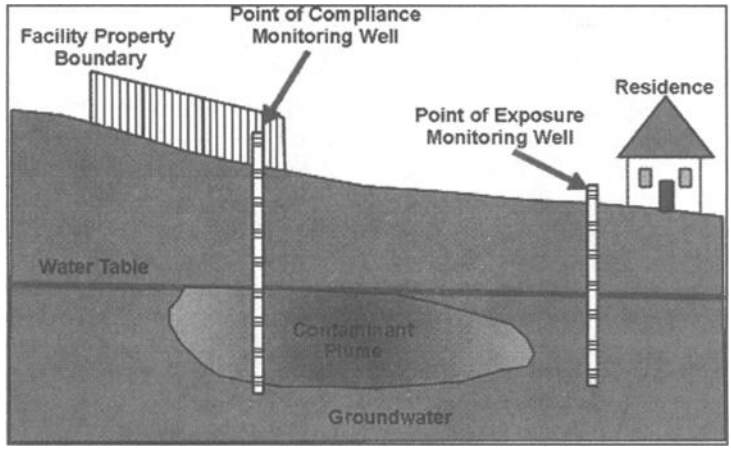


FIG. 7 — *Conceptual Diagram for the ACL Approach.*

Summary and Conclusions

The use of quantitative techniques to address uncertainty and variability for environmental analyses is becoming more advantageous and cost effective. Regulatory agencies are beginning to realize the benefits of these analyses for decision-making purposes. Demonstrated cost savings can be realized through the application of these techniques. Optimal sampling designs can be realized with decision support tools, as well as reduced time to characterize and cleanup a site. The decision maker can be provided more useful information to make defensible decisions. Budget estimates, including contingencies, are better formulated with these techniques as well. Computer codes are now available to facilitate uncertainty analyses and are fully functional on today's desktop computers.

References

Deutsch, C. V. and Journel, A. G., 1998, *GSLIB, Geostatistical Software Library and User's Guide*, Oxford University Press.

Johnson, B. L. and DeRosa, C. T., 1997, "The Toxicological Hazard of Superfund Hazardous Waste Sites", in *Reviews on Environmental Health*, Freund Publishing House, Ltd., Volume 12, No. 4, pp. 235–251.

Knowlton, R., Peterson, D., Zhang, H., Ehrhorn, T., Metzler, D. and Kautsky, M., 2001, "Uncertainty and Variability Using MODFLOW and MT3DMS in the GANDT Decision Support Tool," *Proceedings of MODFLOW 2001 and Other Modeling Odysseys Conference*, Colorado School of Mines, Golden, Colorado, pp. 293–297.

U.S. Congress, Office of Technology Assessment, 1989, *Coming Clean: Superfund's Problems can be Solved...*, OTA-ITE-433, Washington, D.C., U.S. Government Printing Office, October 1989.

U.S. Department of Energy (DOE), 1996, *Final Programmatic Environmental Impact Statement For The Uranium Mill Tailings Remedial Action Ground Water Project*, DOE/EIS–0198, October 1996, prepared by the U.S. Department of Energy, UMTRA Project Office, Albuquerque Operations Office, Albuquerque, New Mexico.

U.S. Department of Energy (DOE), 1998, *Final Site Observational Work Plan for the UMTRA Project Site at Riverton, Wyoming*, Document Number U0013801, February 1998, prepared by the U.S. Department of Energy, Albuquerque Operations Office, Grand Junction Office, Grand Junction, Colorado.

U.S. Environmental Protection Agency (EPA), 1994a, *Statistical Methods for Evaluating the Attainment of Cleanup Standards, Volume 3: Reference-Based Standards for Soils and Solid Media*, EPA 230-R-94-004, June 1994.

U.S. Environmental Protection Agency (EPA), 1994b, *Guidance for the Data Quality Objectives Process*. EPA QA/G4, Office of Research and Development, EPA/600/R-96/055.

U.S. Environmental Protection Agency (EPA), 1999, *Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action, and Underground Storage Tank*

Sites, EPA Office of Solid Waste and Emergency Response (OSWER), OSWER Directive 9200.4-17P (April 21, 1999).

U.S. Environmental Protection Agency (EPA), 2000a, *Environmental Technology Verification Report, Environmental Decision Support Software, DecisionFX, Inc., SamplingFX,* EPA/600/R-00/038, March 2000, U.S. Environmental Protection Agency, Environmental Sciences Division, National Exposure Research Laboratory, Las Vegas, Nevada.

U.S. Environmental Protection Agency (EPA), 2000b, *Environmental Technology Verification Report, Environmental Decision Support Software, DecisionFX, Inc., GroundwaterFX,* EPA/600/R-00/037, February 2000, U.S. Environmental Protection Agency, Environmental Sciences Division, National Exposure Research Laboratory, Las Vegas, Nevada.

David T. Hansen [1]

Status of Standards and Guides Related to the Application of Spatial Methods to Environmental and Hydrologic Problems

Reference: Hansen, D. T., **"Status of Standards and Guides Related to the Application of Spatial Methods to Environmental and Hydrologic Problems,"** *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy, and Standardization, ASTM STP 1420*, D.T. Hansen, V. H. Singhroy, R. R. Pierce, and A. I. Johnson, Eds., ASTM International, PA, 2003.

Abstract: Geographic information systems (GIS) and related technologies are active tools in addressing environmental and hydrologic problems. The application of spatial methods to environmental and hydrologic problems has been supported by the rapid development of computer systems, software, and the internet. The present computing environment has become increasingly dynamic and the use of spatial methods to address local issues has been rapidly growing.

The fields of environmental studies and hydrology have long been active areas for the development of standard guides and practices. ASTM has been active in the development of standards for site characterization, modeling of surface and ground-water systems, and the application of statistical and geostatistical techniques. Standards related to spatial data are more recent. These include the content of digital spatial data and procedures for capturing and recording surface location and elevation. Work has also been done on defining characteristics for particular sets of data in the United States. This includes developing base data layers and classification schemes for some thematic layers. Some work has been done on analysis methods.

The use of spatial methods to address site-specific issues will be assisted by further development of the standards process in this dynamic environment. Data management, data integration, and data quality are common issues in this environment. Common methods, practices, or guides that address these issues will assist in the application of spatial methods to site-specific issues.

Keywords: geospatial standards, geostatistics, spatial statistics, classification standards, federal geographic data committee, national spatial data infrastructure, open GIS, site characterization, modeling, global positioning system (GPS), remote sensing, graphical user interface.

[1] Geospatial Scientist, Soil Scientist, MPGIS, U.S. Bureau of Reclamation, 2800 Cottage Way, Sacramento CA. 95825-1898.

open GIS, site characterization, modeling, global positioning system (GPS), remote sensing, graphical user interface.

Introduction

For this discussion, spatial methods applied to environmental and hydrologic problems involve the use of a computing environment such as geographic information systems (GIS), remote sensing, and related technologies. All involve people, computer hardware and software. They are used to collect and organize data where coordinate values of x, y, and z are key attributes of location for digitally representing real-world features. Capture may be direct with global positioning systems (GPS), or indirect from remotely sensed data or other map or coordinate sources. Data with defined coordinate systems can be manipulated in the same coordinate space to model environmental and hydrologic characteristics for analysis.

Computer and telecommunication protocols before 1990 have provided the framework for the rapid growth in the use of computers and the internet. Over the past ten years, a variety of standards have assisted in the rapid development and deployment of computer applications with graphical user interfaces. GIS and remote sensing applications have made active use of these developments. Standards directly related to geospatial data provide a basis for common terminology, common data structures and definitions. They support support access to and application of geospatial data. At present, there is relatively little guidance on the integration and application of geospatial data for solving site-specific environmental and hydrologic problems.

Existing Standards

Development of graphical user interfaces and the internet has made data and analysis tools readily available to almost anyone. Commonly accepted computer standards have assisted in integrating data across computer platforms and between computer applications. Data and applications can now be easily be transferred and accessed between systems. They can also be accessed across the internet. This set of commonly accepted protocols and standards have assisted in the development and dissemination of networking systems and applications across all computing environments. GIS and remote sensing data have taken advantage of these information technology standards for accessing spatial data and for developing geospatial applications. A small but growing number of standards are related to digital geospatial data development. A few are related to geospatial analysis, geospatial applications, or geospatial methods. Major standard topics related to geospatial data are identified in Table 1. Some standards overlap into more than one topic area.

Table 1 – *Standards Related to Digital Geospatial Applications*

Major Topic	Standard or Protocol
Computing Software / Hardware	TCP / IP
Data Access & Transfer	File Transfer Protocol (FTP) Z39.50 Component Object Model (COM) Structured Query Language (SQL) Mark up Languages (HTML, XML, GEOML) Spatial Data Transfer Standard (SDTS)
Coordinate Control	Geodetic Control Data Profiles for Cadastral, Imagery and Remote Sensing Height Modernization (Elevation)
Data Content	Content Standards for Digital Geospatial Metadata Geospatial and Geopositional Accuracy Soils Geographic Data Standard Geodetic Control Utility Data Data Profiles for different formats: Vector Cadastral Imagery and Remote Sensing CADD
Data Development	Vegetation Classifications Geologic Map Symbolization Frame Work Data Object Models
Data and Site Analysis	Geostatistics Site Characterization and Modeling

Computer Software / Hardware, Data Access and Transfer

Two organizations focused on developing standards and guidelines exclusively related to geospatial data are the Federal Geographic Data Committee (FGDC) of the United States government and the Open GIS Consortium. In 1993, the Federal Geographic Standard Committee (FGDC) issued the Spatial Data Transfer Standard also known as FIPS 173 with a vector profile. The intent of this standard was the encapsulation of GIS vector data and data quality information in a format for transfer between systems. Subsequently FGDC has issued profiles for spatial data in other formats such as imagery and CADD. These and related FGDC standards are available from their Web site (Status of FGDC Standards 2001).

The Open GIS Consortium is an international industry consortium whose goal is to achieve interoperability between commercial geoprocessing software. This group strives to normalize encoding of information between software systems, naming of features and feature relationships, and schemas for describing data. They have been active in developing open interfaces between software systems. This has included development of extensions to general digital formats and interfaces to handle geospatial data. This includes geography markup language as an extension to HTML and XML. This consortium has supported adoption of nonproprietary formats for the representation of features. It has been active in developing tools for accessing and performing spatial operations over the Web (Open GIS Consortium, 2001).

Coordinate Control

The underlying control for GIS data in the United States is either the geodetic network of the National Geodetic Survey (NGS) or the various cadastral surveys for a particular locality. NGS has the responsibility for establishing and maintaining the geodetic reference system for the United States (Yeager et al.). With cooperators from other Federal, State, local agencies and surveyors a reference set of monuments have been established. Horizontal and where available vertical coordinate values are reported to the national database available via the Web (NGS Data Sheets 1999). These points represent the underlying control for all of the national map series. Figure 1 shows the control point set near Courtland, California and the corresponding location depicted on a digital raster graphic (DRG) of the USGS 7.5 minute topographic sheet. For elevation data in California and elsewhere, NGS has been active in coordinating development of control networks for height modernization (Zilkoski et al. 1999). NGS has developed guidelines for using GPS derived ellipsoid heights to establish vertical control networks (Zilkoski et al. 1997).

Figure 1- *NGS Courtland Control Point*

Cadastral surveys are another major source of control for digital geospatial data in the United States. The survey record serves as the basis for constructing the digital

representation of ownership parcels and other locally developed data. For public lands in the United States, this record first began in 1785 with surveys by the General Land Office, now the Bureau of Land Management (BLM). These surveys cover public lands and subdivision of those lands for transfer to private ownership. Local governments have carried on the practice with further subdivision and identification of parcels. These local surveys are the basis for highly accurate digital representations of city and urban environments. The surveys are often the control for other detailed data sets such as water, storm water and sewer systems, street networks, and other utilities. This local level of highly detailed information based on survey measurements generally has been loosely linked or integrated with regional or national data.

At the federal level, BLM has responsibility for standards related to cadastral (Status of FGDC Standards 2001). These surveys were tied to a geographic coordinate at the initial starting point or the prime meridian of the survey. Monuments were placed at section corners and other corners in these surveys. Except for meridians in the public land survey and the effort of local and private surveyors to tie their survey to a geodetic control, the geographic coordinates of cadastral surveys are of unknown quality. BLM has been capturing geographic coordinates for existing monuments of the Public Land Survey. This is the basis for the geographic coordinate database (GCDB). This database provides the basis for digital representation of the public land survey with an estimate of the accuracy or reliability of the corner location (GCDB Data Preparation 2000). BLM is in the process of integrating GCDB data with control established by local units of government in their surveys. This will provide a basis for integrating local survey based data into regional and national data.

Data Content

Shortly after adoption of the Spatial Data Transfer Standard, FGDC adopted the Content Standards for Digital Geospatial Metadata. This standard was also reviewed and adopted in slightly modified form by ASTM Committee D18 as ASTM Content Specifications for Digital Geospatial Metadata (D 5714). In the near future, it is expected that this FGDC standard will be replaced by a related standard adopted by the International Standards Organization (ISO). These standards provide a common lexicon of elements and format for describing and documenting geospatial data.

In addition to the content of metadata, FGDC has been active in identifying content for specific groups of data. This includes identifying key elements for biological data with the Biological Data Profile. Wetland and riparian classification systems of the U.S. Fish and Wildlife Service have been adopted. It has adopted an initial hierarchical classification schema for vegetation. There is also a data standard for soils. Content specifications have also been prepared for digital orthoimagery and remote sensing, and utility data. These and related standards are available at the FGDC web site (Status of FGDC Standards 2001). These standards assist in integrating data from different sources.

Data Development

In 1994, the Federal government identified the outlines for a National Spatial Data Infrastructure (NSDI) (Clinton 1994). The vision for NSDI is coordination in the development of geospatial data and open access to that data. Toward this vision, a network of internet nodes have been established for locating and accessing geospatial data. Executive order 12906 recognized the need for common data layers identified as framework data. Table 2 identifies several framework data layers.

Table 2 – *Major Framework Data Layers for the United States*

Thematic Framework Layers	Description
Digital Raster Graphics (DRG)	Digital image of the USGS 7.5 minute topographic map series
Digital Orthoquads (DOQ)	Digital images of orthorectified high altitude aerial photography with a raster cell size of about one meter on the ground.
National Elevation Data (NED)	DEM for United States at a resolution of 10 to 30 meters with derivatives for hydrologic modeling known as NED-H.
Geodetic Control	Reported geodetic control points for the National Geodetic Survey with horizontal and some elevation values for the 1984 North American Datum and the 1988 North American Vertical Datum
Geographic Coordinate Data Base (GCDB)	Geographic coordinates for the corners established by the Bureau of Land Management for the public land survey system
National Hydrology Data (NHD)	Surface hydrology network with stream center lines for graphic display, water quality reporting, and hydrologic modeling. Initial release of this data is at a scale of 1:100,000 based on the river reach system of the Environmental Protection Agency.
Geographic Names (GNIS)	Commonly accepted names for geographic features with location by USGS
National Land Cover Data (NLCD)	Land cover data with hierarchical labeling on a 30 meter basis based on remote sensing interpretations.
Watershed Data Base (WBD)	Watersheds down to sub watershed units

These data layers are based on common definitions to meet common requirements between organizations. Most of these framework layers are accessible via the NSDI nodes set up under the Executive Order.

Many of these layers are designed to work together. For example, the development of NHD requires updating the geographic names data. NHD itself nests within the watershed database. In the area of water utilities and surface water, the digital representation has evolved into an object model. These objects have properties and behaviors that define how they match and combine for network analysis of the entire system (GIS in Water Resources Consortium 2000).

Frequently these thematic layers have been developed cooperatively by several organizations. Data developed by different organizations follow a common schema for data content. This data can be combined to produce a common product. The intent in their development is to produce a seamless data layer that can cover broad areas without additional processing. These layers provide a basic standard for the database structure and attributes for that theme. Accessible over the internet, this data provides the basis for further data development at a regional or local level. Increasingly, the expectation is that several of these layers will be updated with more current and accurate local data.

Data Analysis and Site-Specific Applications

In this discussion, geospatial analysis includes the broad spectrum of data query and display, spatial operations such as overlay, buffer, and spatial modeling. Spatial analysis is one of the tools used for site-specific applications for environmental and hydrologic problems. Issues involved in site analysis are similar to issues faced in developing geospatial applications. Site characterization and evaluation incorporate data from a variety of sources, collected at different times, with different characteristics. The process of assessing, manipulating, and applying data in site characterization is similar to that used in developing geospatial applications. Geospatial analysis of environmental and hydrologic problems is essentially a modeling process. It begins with a conceptual model, the collection and evaluation of data, construction of a model, and evaluation of that model.

ASTM has been extensively involved in the development of methods, practices, and guides for environmental site characterization. These areas include geotechnical methods, soil, surface water, and ground-water characaterization, monitoring, and modeling. Ivan Johnson describes the process in the development of these standards (Johnson 2001). This body of standards covers detailed site investigation planning, characterization, sampling, and reporting. Geostatistics, spatial statistics, or other spatial methods may be used to distribute or evaluate the distribution of the parameters of interest across the landscape. Several case studies have been published by a variety of professional organizations. ASTM has held a symposium on the application of geostatistics to environmental and geotechnical problems (Rouhani et al. 1996). ASTM Committee D18 has also been active in sponsoring several symposia

on applications of GIS and remote sensing to environmental problems (Singhroy et al. 1996), and Johnson et al. 1992).

Issues Related to Site Specific Analysis of Environmental and Hydrologic Problems

The computer desktop can be used to view and access data from a variety of sources any where in the world. Besides viewing and displaying data, it is a platform for spatial analysis. Computing platforms and software have advanced to the point where geospatial applications can be run on the computer desktop. Entire geospatial analysis packages can be placed on CD-Rom or accessed over the internet. Stakeholders, regulators, and the public can do site-specific analysis with data from a variety of sources. In this symposium, we have seen several applications of geostatistics and decision support systems addressing soil and groundwater contamination (Gadgil et al. 2001, Knowlton et al. 2001). This rapid development has out paced our ability to successfully address many issues related to data management, to documentation of the process, and to evaluation of the data or product.

Commonly accepted methods and practices on the use of geospatial data and tools would assist existing and new users in the application of spatial methods to environmental and hydrologic problems. Topics for these methods include basic issues of data management, integration of data from multiple sources, and evaluation of data produced by the process. These issues overlap. Data management and documentation of spatial analysis are increasingly important in the current dynamic computing environment. Integration of geospatial data is a common process for almost any application. Integration of locally developed data with regional and national data is increasing common. Evaluation of the products produced by GIS applications is an important step for any site-specific application.

Data Management and Documentation in Applications

Data used and produced by desktop applications and decision support systems need to be identified and managed. Information on this data provides the basis for evaluating the quality and appropriateness of the data for the application and for reviewing the results. A desktop application by Ducks Unlimited for the Central Valley of California allows community groups to model their local environment. This application permits the user to run a suitability analysis for wetland habitat development (Ducks Unlimited, 1998). In the application, the user selects data layers, assigns weights or recodes these layers, and combines these weights to produce a surface of suitability for wetland habitat. Figure 2 shows a weighting scheme in this application. Combining these recoded layers produces a new data layer of suitability values. The analysis can be run rapidly for a variety of scenarios with different data layers and weighting schemes.

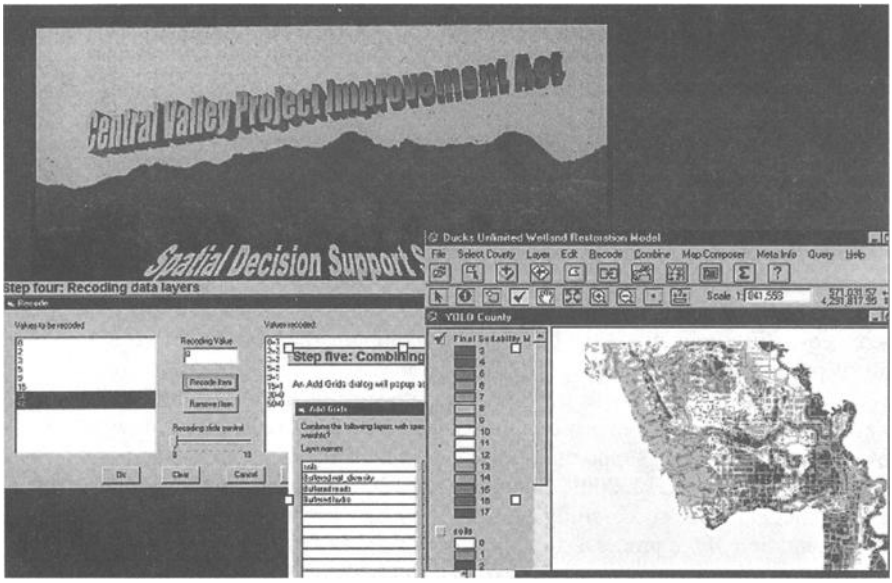


Figure 2 - *GIS Model for Wetland Habitat Suitability*

This application is available on CD-Rom. Other applications are accessible over the internet. These applications may include several processing steps. Steps can include reclassifying data or selection of subsets of data. The source data may be combined or manipulated. Any of these steps produce additional data. To effectively utilize these applications, users need to be able to keep track of the themes used, the processing steps followed, and the data produced. This information provides the basis for evaluating and comparing scenarios. Key data files and processing steps need to be identified to be able to reproduce scenarios.

Common practices or guides for identifying key data, the steps followed, and the outputs produced would assist in development of this and other spatial applications. These are recognized as key elements in the FGDC Content Standards. Implementation the documentation of data and processes used in dynamic environments has not been addressed. Information needs to be retained on the data used, the rating or weighting criteria assigned to the data, and the results obtained as part of the system.

Integration of Data

Data representing a particular theme often come from a variety of sources. They represent different time periods with different levels of accuracy. Many site-specific applications include data from regional or national sources as well as locally

developed data. Common accepted digital capture methods identifying key information to be retained would assist users in integrating data from a variety of sources. For example, identifying information that should be captured along with the coordinate values from a GPS unit would assist the user who is integrating the GPS data with other digital themes. Locally developed data may be high quality surveys. For the spatial application, these may be integrated with similar data where positional accuracies are more uncertain or unknown. It is recognized that this will be the case for several thematic layers in the national framework of the United States. These are referred to as high resolution data layers. Among these layers are NED and NHD. Initially, the seamless layer is developed from 1:24,000 scale data. These sources often represent different time periods with different spatial resolutions. These national data layers are periodically updated as more accurate data is available.

For flood analysis and hydrologic modeling in the Delta region of California, NED is being combined with more accurate data developed locally. Figure 3 shows elevations for an area along the Sacramento River of California from NED and from lidar using remote sensing technology. The resolution of the NED data is 30 meters with elevation intervals of about 2 meters. The resolution of the lidar data is sub meter with sub meter elevation values. The display of NED data is on the left. the display of lidar data is on the right.

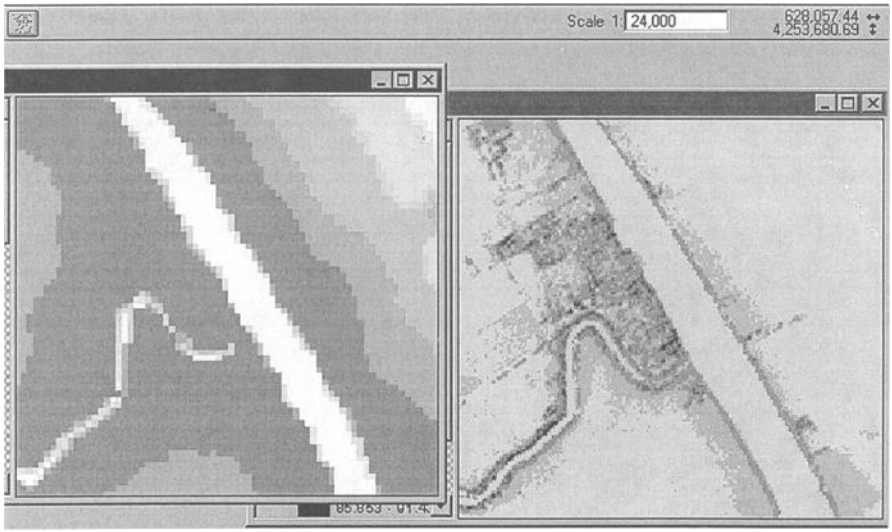


Figure 3 - *Depiction of Surface Elevation from NED on the Left and Radar Imagery on the Right, Sacramento River, California*

This is an area of active subsidence. There have been significant changes in surface elevation since the date of the sources for NED. The high resolution lidar data represents the current surface. Integration of these datasets into a seamless layer for analysis requires identifying common points between the data layers and transforming the layers into a common coordinate system. Values from one layer must then be

substituted for values in the other layer where appropriate. This adjustment will be based on a recent network of high quality elevation control has recently been developed following procedures of NGS (Zilkoski and D'Onofrio 1999).

NED is a relatively simple dataset representing point elevation values. More complex data such as NHD represent greater challenges for integrating locally developed data into a national or regional database. Consensus guides for performing this integration and discriminating the sources of values within the seamless layer would assist in the development and application of the data.

Evaluation of Data Quality

Data quality is an assessment reported as components of metadata. This assessment is to cover both the spatial representation of the feature including its geopositional accuracy and the attributes or data carried for those features. Evaluating the quality of the attributes carried for the spatial features usually follows common procedures. FGDC has identified standards for evaluating and reporting the accuracy of point data representing true points and survey level data (Status of FGDC Standards 2001). The Open GIS Consortium has identified specifications for encoding the quality or accuracy of data types (Open GIS Consortium 1999). There is no clear consensus on evaluating and reporting the quality of other spatial data types.

Figure 4 shows the digital boundaries from two different sources displayed on the top of a digital orthophotoquad.

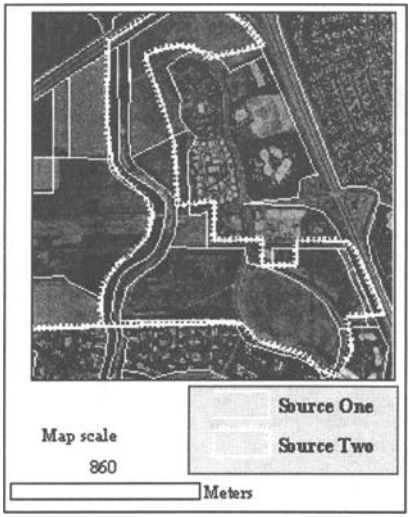


Figure 4 - *Land Use from Two Sources Plotted on a Digital Orthophotoquad.*

The narrow white line represents one digital source. The broader dashed line represents another source. Each data source was originally compiled at a scale of 1:24,000. They represent similar types of land use. The display shows clear differences in the boundaries between the data sources. A simple overlay of these datasets would produce a myriad of sliver polygons. Another method to combine these sources is to generate a new base map and then capture the landuse types from each source. In this method, polygons are generated from the digital image. The codes from each source are then captured for these new polygons. This eliminates sliver polygons. The new base is from the digital orthoquad. It controls the spatial resolution of the resulting product.

Evaluating digital data for an application requires an understanding of data quality for both the spatial representation and the attributes. In many spatial applications, several data layers are combined in an analysis. For site-specific applications, these sources will include regional data developed at various scales, survey grade data as well as data captured with GPS, remote sensing, or aerial photography. The accuracies of some of this data will not be known. Identifying common methods for evaluating data for various applications would assist in their use. This evaluation should include some methods for estimating the accuracy or uncertainty in the spatial representation. This will improve the confidence in applying these tools to site-specific applications.

Summary

In June of 1999, the Federal Geographic Data Committee hosted a National Geodata Forum for the United States. This forum involved several hundred people from across the United State representing the Federal Government, Native American indian tribes, states, regions, and local governments. The vision of this forum was to be able to integrate the digital information from all these organizations to address issues at the community level. While this goal is approachable now, it is elusive when attempting to integrate information from all levels to address site-specific issues. Currently features do not always match at the database level or in the spatial representation. The forum recognized that major steps had been taken with the development of standards, NSDI nodes for data sharing, and framework data requirements and data. It also recognized that further work was needed to understand requirements for data, software, and users at the local application level.

The past decade has seen the development of standards that have assisted in the documentation, development, and dissemination of both digital geospatial data and tools to apply that data to national, regional, and community issues. There is a large set of standards for site-characterization, evaluation, and modeling for environmental and hydrologic issues. Applying and integrating these standards in our dynamic computing environments require further work. Further coordination is needed to effectively integrate and utilize these data and tools. This includes involvement of all governmental levels, business, professional organizations, academia and community groups. This information must be able to match or blend with information collected

from different sources, over different time periods, and which represent different spatial resolutions or accuracies. Groups must be able to come to consensus on appropriate methods for evaluating and using this information to address environmental or hydrologic problems at the site level. There is a role for a variety of organizations to be involved in this process.

References

ASTM, 1997, *ASTM Standards Related to Environmental Site Characterization*, ASTM International, West Conshohocken, PA.

Clinton, William J., April 13 1994, "Coordinating Geographic Data Acquisition and Access, The National Spatial Data Infrastructure," *Executive Order 12906*, Washington, DC.

Ducks Unlimited, 1998, *User's Guide Central Valley Project Improvement Act GIS Model Version 1.0*, Ducks Unlimited, Inc. 3074 Gold Canal Drive, Rancho Cordova, CA.

Gadgil, J. M., P. A. Sullivan, J.B. Kool, and S. Mayer, 2002, "Application of Gesostatistical Techniques for Risk Based Soil Cleanup," *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy, and Standardization, ASTM STP 1420*, D. T. Hansen, V. H. Singhroy, R. R. Pierce, A. I. Johnson, Eds., ASTM International, West Conshohocken, PA.

GCDB Data Preparation, June 2000, Geographic Coordinate Data Base, http://www.blm.gov/nhp/what/lands/cadastral.html, Bureau of Land Management, Washington, DC.

GIS in Water Resources Consortium, June 2000, *ArcGIS Hydro Data Model*, David R. Maidment, Editor, http://www.crwr.utexas.edu/giswr, Center for Research in Water Resources, University of Texas at Austin, TX.

Johnson, A. I., Petersson, C. B., and Fulton, J.L., editors, 1992, *Geographic Information Systems (GIS) and Mapping - Practices and Standards, ASTM STP 1126*, ASTM International, West Conshohocken, PA.

Johnson, A. Ivan, 2002, "ASTM Standards Related to Uncertainty and Risk," *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy, and Standardization, ASTM STP 1420*, D. T. Hansen, V. H. Singhroy, R. R. Pierce, A. I. Johnson, Eds., ASTM International, West Conshohocken, PA, 2001.

Knowlton, R. G. M. Peterson and H. Zhang, 2002, "The Use of Decision Support Systems to Address Spatial Variability, Uncertainty, and Risk," *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy, and Standardization, ASTM STP 1420*, D. T. Hansen, V. H. Singhroy, R. R. Pierce, A. I. Johnson, Eds., ASTM International, West Conshohocken, PA.

National Geodetic Survey Data Sheets, March, 1999, NGS Data Sheets: http:\www.ngs.noaa.gov/datasheets.html, National Geodetic Survey, Washington, D.C.

OGC's Role in the Spatial Standards World, January 2001,

URL://www.opengis.org/datasheets/Dat01Stds000127.html Open GIS Consortium, USA.

Open GIS Consortium, 1999, "The OpenGIS Abstract Specification, Topic 9, Quality, Version 4", http://www.opengis.org/tehno/specs.html, The Open GIS Consortium, USA.

Rouhani. S., Srivastava, R. M., Desbarats, A. J., Cromer, M. V., Johnson, A. I., Eds., 1996, *Geostatistics for Environmental and Geotechnical Applications. ASTM STP1283*, ASTM International, West Conshohocken, PA.

Singhroy, V. H., Nebert, D. D., Johnson, A .I., 1996, *Remote Sensing and GIS for Site Characterization; Applications and Standards, STP 1279*, ASTM International, West Conshohocken, PA, 2001.

Status of FGDC Standards, January 2001, URL: http://www.fgdc.gov/standards/status/textstatus.html, Federal Geographic Data Committee, Washington, DC.

Zilkoski, D., D'Onofrio, J. D., and Frakes, S. J., November 1997, *NOAA Technical Memorandum NOS NGS-58, Guidelines for Establishing GPS-Derived Ellipsoid Heights, NGS Standards: 2 cm and 5 cm, Version 4.3*, http:\\www.ngs.noaa.gov/PUBS_LIB/NGS-58.html, National Geodetic Survey, Silver Spring, MD.

Zilkoski, D., and D'Onofrio, J. D., 1999, *Geodetic Phase of NOS San Francisco Bay Demonstration Project*, http:\\www.ngs.noaa.gov/initiatives/HeightMod/Geodetic/sfbay_geodetic.htm, National Geodetic Survey, Silver Spring MD.

Yeager, J. A.,Laoine, L. A. , and Spencer, J. F., Jr., *The Contribution of Geodetic Data to the National Spatial Data Infrastructure*, http://www.ngs.noaa.gov/PUBS_LIB/NSDI_Lapine.html, National Oceanographic and Atmospheric Administration, National Geodetic Survey, Silver Spring, MD.

Development of Standard Data Sets

Marti E. Ikehara[1]

Application of GPS for Expansion of the Vertical Datum in California

Reference: Ikehara, M. E., **"Application of GPS for Expansion of the Vertical Datum in California"** *Spatial Methods for Solution of Environmental and Hydrologic Problems - Science, Policy, and Standardization, ASTM STP 1420*, D.T. Hansen, V. H. Singhroy, R. R. Pierce, and A. I. Johnson, Eds., ASTM International, West Conshohocken, PA, 2003.

Abstract: Global Positioning System (GPS) surveying is helping to expand the sparse coverage of geodetic stations referenced to the North American Vertical Datum of 1988 (NAVD 88) in California. Accurate and precise vertical control is an important aspect of solutions to various hydrologic problems such as floodplain mapping and aqueduct gradient determinations. The National Geodetic Survey (NGS) has developed guidelines for GPS procedures that will result in local accuracies for ellipsoid heights of 2 cm at 95% confidence. The NAVD 88 datum in California is being expanded in segments of GPS networks that are observed by multiple-agency partnerships. Networks between San Francisco and Sacramento have been completed, and expansion in the Central Valley is being pursued. Determination and maintenance of accurate vertical heights in the Central Valley is particularly difficult because of large historical and unknown current effects of land subsidence resulting from massive quantities of groundwater withdrawal for several decades prior to the 1980's. Other NGS goals are to establish more geodetic and tidal datum connections, and to provide real-time water level information to harbor pilots for safer and more efficient vessel loading and navigation.

Keywords: geodetic surveying, vertical control, Global Positioning System

Introduction

In California, Global Positioning System (GPS) surveying is helping to expand the sparse coverage of geodetic stations referenced to the North American Vertical Datum of 1988 (NAVD 88). This datum has superceded National Geodetic Vertical Datum of 1929 (NGVD 29). Accurate and precise vertical control is an important aspect of solutions to various hydrologic problems such as floodplain mapping and aqueduct gradient determinations. The challenge of determining and maintaining high-quality accurate vertical control has increased with the growing emphasis on the regional, rather than local, scope of problems. In past, the scope of a problem area might have been a few kilometers in diameter. We now recognize that, because systems in the natural world are

[1] Geodetic Advisor for California, National Geodetic Survey, 1727 30th St., MS-35, Sacramento, CA 95816.

interdependent, the scope of a problem and/or solution area might be many tens of kilometers in diameter or length. Measurements of geodetic control in a regional network in today's environment must balance efficiency with accuracy needs.

After many pilot projects to test and prove the concept, the National Geodetic Survey (NGS) has developed guidelines for GPS surveying procedures that will result in ellipsoid height local accuracies with a relative error of less than 2 cm at 95% confidence (Zilkoski et al.1997). This means that the error for the difference in heights of adjacent stations determined in this network will seldom exceed 2 cm. If the project is properly surveyed with the highest quality geodetic control of the National Spatial Reference System (NSRS) (Doyle, D. R. 1994), the network accuracy is 5 cm at 95% confidence. This means that the error in the relative difference in heights of two bench marks surveyed to similar high-accuracy standards and tied to the NSRS will not exceed 5 cm.

What are ellipsoid heights? Are they relative to "mean sea level"? No, ellipsoid heights are not relative to mean sea level (NGVD 29) nor to NAVD 88, the official vertical datum in North America. Because GPS satellites orbit around the Earth, the measurements we get from GPS data are referenced to the ellipsoid that approximates the earth's shape. In the U.S., most GPS surveyors choose the North American Datum of 1983 (NAD 83) ellipse as the reference datum; the vertical component of GPS data is called the ellipsoid height. To get GPS-derived heights that are comparable to orthometric (leveled) heights (H), a geoid height (N) is subtracted from the ellipsoid height (h); these components are illustrated in Figure 1. The geoid height is the difference between the ellipsoid and the global geoid (an equipotential surface). The estimations of N for a geoid model are dependent on topographic relief of the earth's surface and gravity values that change due to subsurface rock density variations.

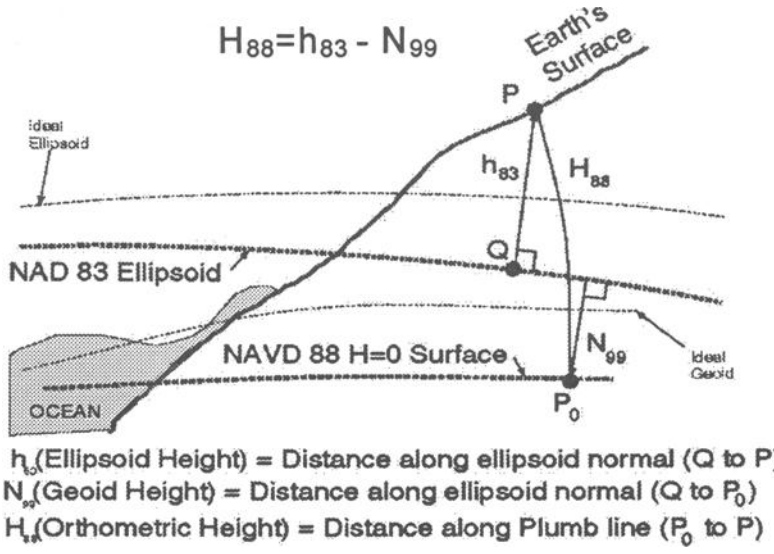


Figure 1. *Orthometric height (NAVD 88) is computed by subtracting geoid height(GEOID99) from ellipsoid height (NAD 83).*

A geoid model, which is based on gravity, leveled heights, and GPS heights at essentially the same location, is used in GPS processing to convert ellipsoid heights to orthometric heights. The current NGS model, GEOID99, utilizes leveled elevations referenced to NAVD 88. Decades ago we learned that the ocean level varies nearly 3 ft from coast to coast in North America, but the discrepancy could be averaged across the country without causing misclosures in a surveyor's project. As the more precise GPS surveying methods (100 times better than previous methods) become relatively commonplace, particularly at continental and regional scales, NGVD 29, often equated with 'mean sea level', will become obsolete.

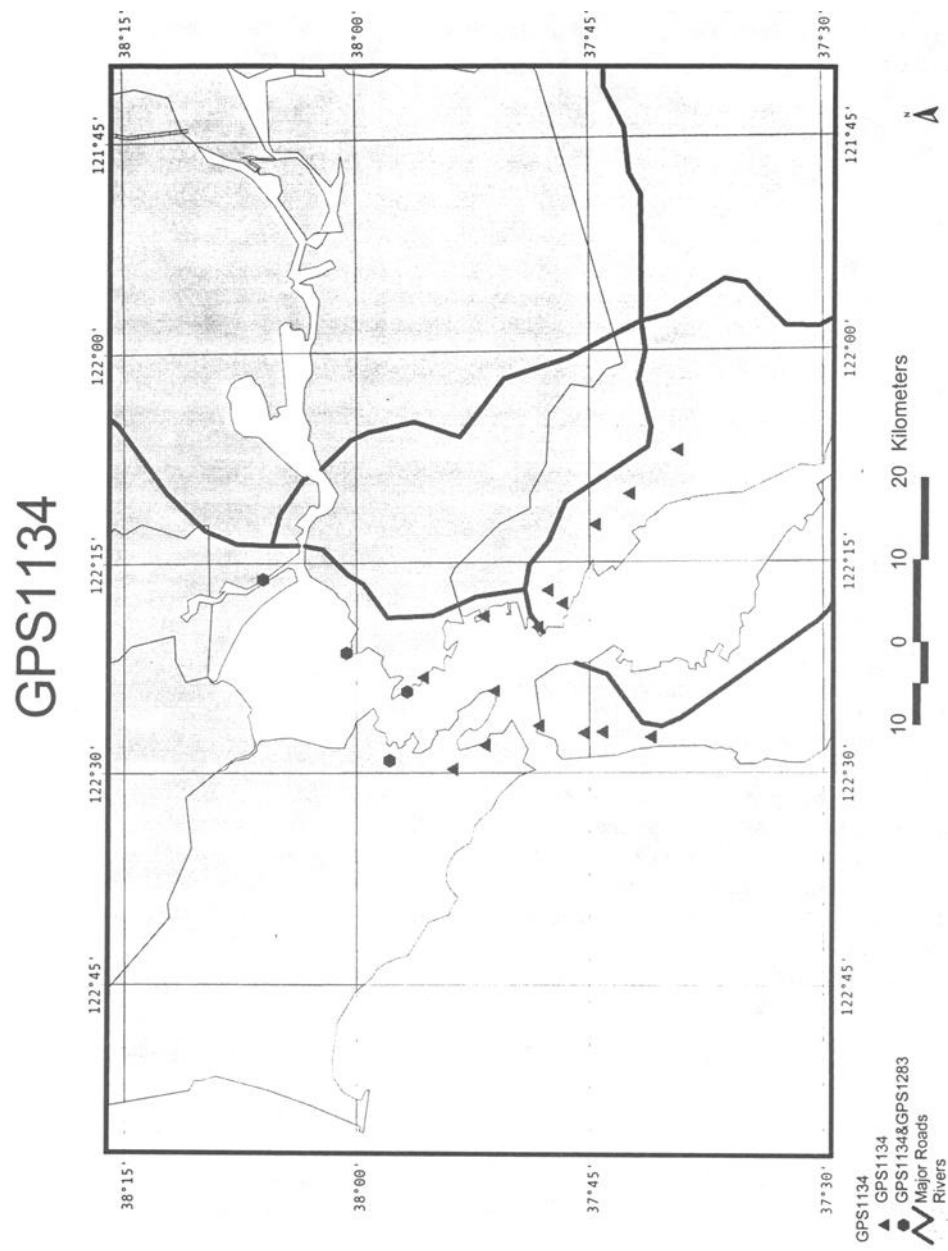
In addition to defining the new vertical datum, NAVD 88, NGS also helps people understand the relationship between the two geodetic vertical datums and between NAVD 88 and various tidal datums. NGS cooperates with the Centers for Operational Oceanographic Products and Services (CO-OPS) and Coast Survey (CS) --all of which are offices within the National Ocean Service-- to ascertain the relationship between NAVD 88 and the local tidal datums. Mean sea level and other water levels, or datums, such as MLLW (mean lower low water) or MHW (mean high water) are defined at each tide gauging station after a specific period of data are recorded and analyzed. The relationship between a local tidal datum and NAVD 88 can be defined if the bench mark(s) established at the tide station are surveyed with other bench marks that have been adjusted to the NAVD 88 datum.

NAVD 88 Expansion using GPS Surveys

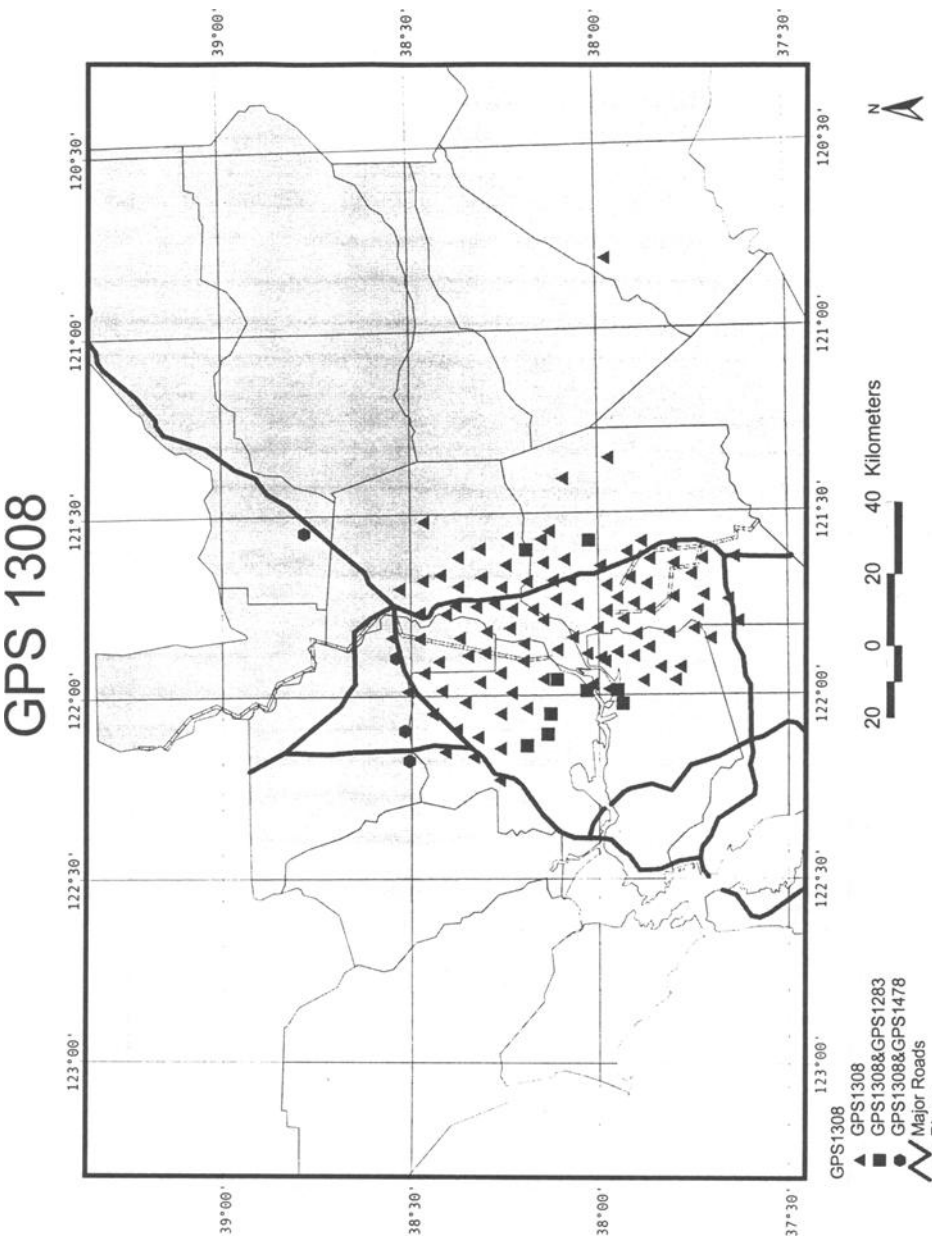
In California, the NAVD 88 datum is being expanded in segments of GPS networks that are observed by multiple-agency partnerships. This paper will focus on the efforts that have resulted in the expansion of the datum in northern California between San Francisco (S.F.) and Sacramento. It is instructive to note that the need for the expansions was driven by agencies whose missions relate to hydrology--both freshwater and tidal. For example, in the Bay Area, the S.F. Bay Conservation and Development Commission (BCDC) and the California Coastal Commission were among the cooperating agencies. For the Sacramento Delta network, the U.S. Bureau of Reclamation and the California Department of Water Resources played key roles in the organization and execution of the surveying.

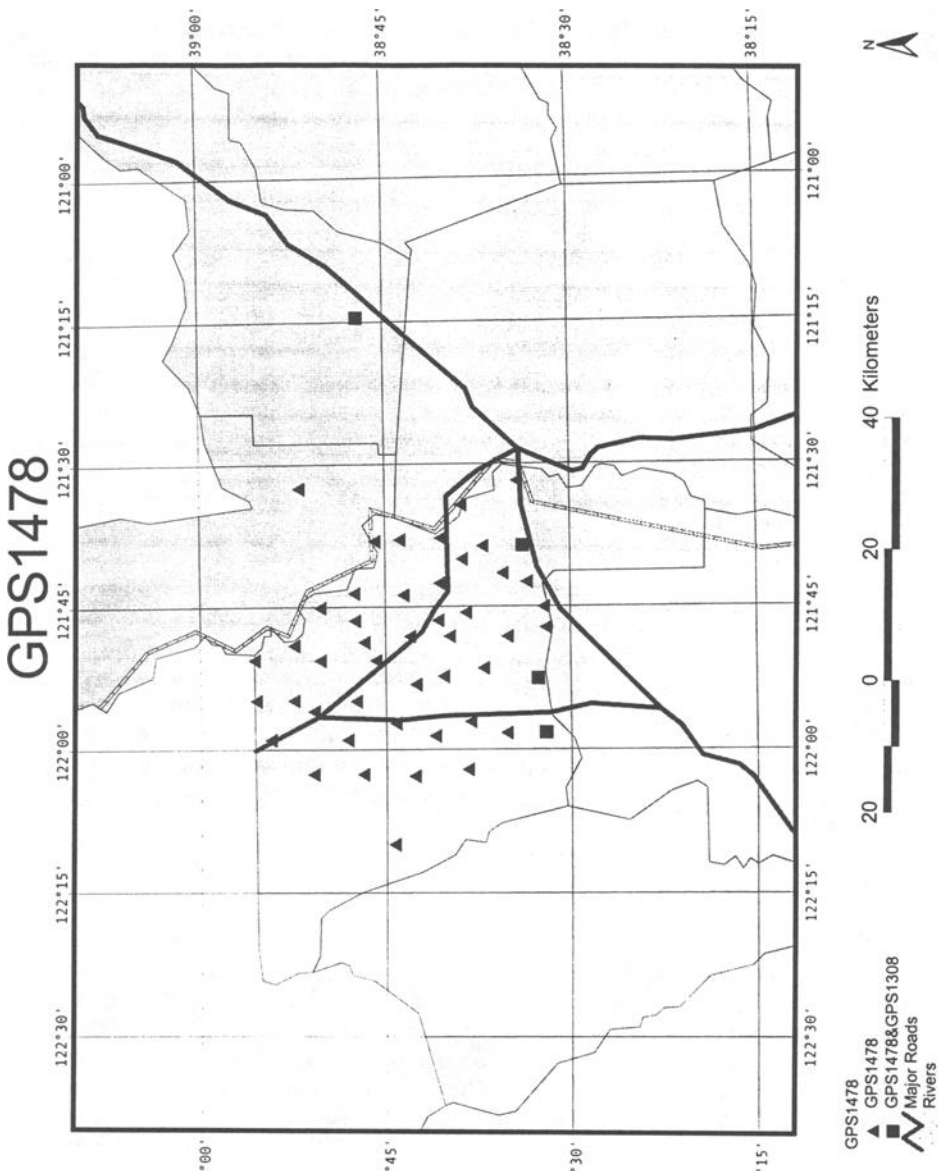
There are currently 4 networks that contributed to the expansion of NAVD 88 in the larger Sacramento-San Francisco region. They are tabulated as follows:

Year	Network Name	Project No.	No. of Pts.	Main Counties
1995	North SF Bay	GPS1134	20	S.F., Marin, Alameda
1996	San Pablo/Suisun Bays	GPS1283	43	Marin, Sonoma, Solano, Contra Costa
1997	Sacramento Delta	GPS1308	112	Sacramento, Yolo, Solano, Contra Costa, San Joaquin,
1999	Yolo County	GPS1478	53	Yolo


GPS1134
38°15'
38°00'
37°45'
37°30'
122°45'
122°30'
122°15'
122°00'
121°45'
GPS1134
GPS1134
GPS1134&GPS1283
Major Roads
Rivers
10 0 10 20 Kilometers
N

GPS1283
123°3'
122°44'
122°25'
122°6'
121°47'
121°28'
121°9'
38°54'
38°35'
38°16'
37°57'
37°38'
GPS1283
GPS1283
GPS1283&GPS1134
GPS1283&GPS1308
Major Roads
Rivers
10 0 10 20 Kilometers
N


GPS 1308
39°00'
38°30'
38°00'
37°30'
123°00'
122°30'
122°00'
121°30'
121°00'
120°30'
GPS1308
GPS1308
GPS1308&GPS1283
GPS1308&GPS1478
Major Roads
Rivers
20 0 20 40 Kilometers
N


GPS1478
122°30'
122°15'
122°00'
121°45'
121°30'
121°15'
121°00'
39°00'
38°45'
38°30'
38°15'
GPS1478
GPS1478
GPS1478&GPS1308
Major Roads
Rivers
20
0
20
40 Kilometers
N

The extent of each network is shown in Figures 2 and 3. Those points that are included in more than one network are indicated by an asterisk; the symbol for those surveyed in just one network is a triangle. National Continuously Operating Reference Stations (CORS) are not distinguished separately. At CORS sites, a permanent antenna mount has been established and the GPS unit is recording measurements every 30 seconds continuously. Data for the CORS in the NSRS are available at the NGS website: www.ngs.noaa.gov/CORS/ and data for many CORS worldwide can be found at the Scripps Orbit and Permanent Array Center website: www.sopac.ucsd.edu. Except for the Yolo County network, each of the networks included several tidal station bench marks so that the relationship between the tidal datums and NAD 83 could be determined.

Each subsequent network utilized some of the same points in networks previously surveyed. After the Sacramento Delta survey was completed, all the data from the first three networks was combined in a single least-squares surveying adjustment, thus assuring that the ellipsoid heights of points in any one of the networks meet the 5-cm network accuracy at 95% confidence. Surveying practices and office procedures for each of these networks complied with the "Guidelines for achieving 2-cm ellipsoid heights." Leveling for a few pairs of bench marks in the Yolo County network separated by <5 km was done about the same time as the GPS observations for a comparison. Differences between leveling and the GPS-derived orthometric heights for these four baselines were <1.5 cm, confirming that the 2-cm standard was achieved.

There are many specifications involved with the field and office components of designing, observing, and processing a network with the intent of getting precise ellipsoid heights. The document (Zilkoski et al, 1997) describing these specs can be found at: http://www.ngs.noaa.gov/PUBS_LIB/NGS-58.html. Some of the more important ones are mentioned here. The distance between adjacent geodetic control stations in the local network was <10km. Another critical part of the guidelines requires that each of these baselines be re-observed on a different day and different time of day to allow the satellite configuration to change significantly. The difference in the vertical component of the two baseline vectors should not exceed 2 cm or the observation must be done a third time and achieve the 2 cm criteria.

Summary

California has a lot of crustal motion, both horizontally and vertically, so establishing a horizontal or vertical datum with one set of measurements is not sufficient. In order to maintain accurate control that reflects the shift in position--whether horizontally or vertically or both--of the land surface, periodic surveys have to be done to update the three-dimensional coordinates. Determination and maintenance of accurate heights in the Central Valley are particularly difficult because of large historical and unknown current effects of land subsidence due to groundwater withdrawal.

What hydrological and environmental applications or studies could utilize accurate vertical and/or tidal datums? One application that requires both is wetland restoration. Wetland restoration of tidally-influenced habitats, which have become more prevalent in the past few years in the Bay Area and Delta, could require accuracies of just a few centimeters. Precision and accuracy are critical when the slope of the proposed restoration area is gradual, whether the restored land is in-situ or involves earth-moving

plans. Accurate and precise vertical control is an important aspect of solutions to various hydrologic problems, such as floodplain mapping and aqueduct gradient determinations--another situation which requires centimeter-level accuracies for many kilometers. In the Bay Area, consideration is being given to establishing real-time CORS that would provide near-centimeter coordinates for positioning and navigation rather than sub-meter accuracies such as available through the Coast Guard DGPS (differential GPS) sites. Establishing real-time CORS on buoys or channel markers at critical commercial passages would allow accurate real-time water levels to be transmitted to ship pilots. Providing precise real-time water levels to harbor pilots at commercial ports would enable safer navigation and more efficient vessel loading and unloading. Vessels are generally "underloaded" because the water level of the destination harbor at the arrival time has to be estimated and a sufficient safety factor must be incorporated so that the vessel doesn't go aground.

With the help of several dozen agencies and private businesses in this region, the networks between San Francisco and Sacramento have been completed, and expansion is being planned. For the existing networks in this area, re-surveys will need to be scheduled to monitor the presence, extent, and magnitude of land subsidence or other crustal motion effects. Plans are underway to complete the southern S.F. Bay loop this year under the auspices of the California Spatial Reference Center (CSRC). Additionally, expanded coverage of the NAVD 88 datum in the Central Valley (north and south of the Delta network) is a goal being pursued to help with solving the environmental and hydrologic problems associated with the legal and equitable distribution of fresh water for human consumption and natural needs.

References:

Doyle, D. R., 1994, "Development of the National Spatial Reference System," URL: http://www.ngs.noaa.gov/PUBS_LIB/develop_NSRS.html, National Geodetic Survey, Silver Spring, MD, 1 October, 2001.

Zilkoski, D. B., D'Onofrio, J. D., and Frakes, S. J., November 1997, "Guidelines for Establishing GPS-Derived Ellipsoid Heights (Standards: 2 cm and 5 cm), Version 4.3, NOAA Technical Memorandum NOS NGS-58, Silver Spring, MD.

Vern H. Singhroy [1] and Peter J. Barnett[2]

Satellite Based Standardized Terrain Maps: A Case Study

Reference: Singhroy, V. H., and Barnett P. J., **"Satellite Based Standardized Terrain Maps: A Case Study,"** *Spatial Methods for Solution of Environmental and Hydrologic Problems -- Science, Policy, and Standardization, ASTM STP1420,* D. T. Hansen, V. H. Singhroy, R. R. Pierce, and A. E. Johnson, Eds., ASTM International, West Conshohocken, PA, 2002.

Abstract: The Ontario Geological Survey and the Canada Centre for Remote Sensing are currently preparing a series of satellite based terrain maps for a 250 000 square kilometer area of the boreal forest region in Northwest Ontario. The purpose of this provincial and federal mapping program is to produce a series of 1:100 000 standardized, satellite based engineering-terrain maps that will be published as a provincial map series. The terrain maps are being used to plan forestry roads and other civil engineering works in support of forest harvesting programs in the region. They are also used to verify forest productivity models in the boreal forest. This paper presents the interpretation methodologies and examples of the satellite based standardized terrain maps. Our results show that image maps produced from a combination of DEM and TM can provide a base on which to interpret and overlay engineering terrain units at a scale of 1:100 000 for large areas of the boreal forest regions in northern Canada. This method will result in considerable savings in time and cost when compared to traditional air photo methods.

Keywords: Terrain Maps, Boreal Forest, Landsat, DEM

Introduction

Northern Ontario Engineering Geology Terrain Studies (NOEGTS) were completed for the part of the Canadian boreal forest region south of latitude 51° N. Each study included a 1:100 000-scale terrain map that was based almost entirely on the interpretation of air photographs with limited field checking. The legends of the maps contain information on surface material type, landform, topography (relief) and drainage conditions. These maps provide useful information concerning the landscape for forest management and civil engineering.

[1] Research Scientist, Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, ON, Canada K1A 0Y7, vern.singhroy@ccrs.nrcan.gc.ca

[2] Sedimentary Geologist, Sedimentary Geoscience Section, Ontario Geological Survey, 933 Ramsey Lake Road, Sudbury, ON, Canada P3E 6B5, peter.barnett@ndm.gov.on.ca

The vast area of boreal forest north of 51° N latitude, for the most part, has no equivalent maps of the terrain conditions. Such maps would be expensive to create using traditional air photo interpretation and field investigations because of the large areal extent and limitations of access. As a result, this study attempts to create engineering geology terrain maps using the integration and interpretation of various types of remotely sensed imagery, digital elevation models and their derivatives and appropriate geological depositional models. The geobotanical remote sensing techniques developed by [*1, 2, 3, 4*] were used as a basis for interpretation of satellite based remote sensing for terrain mapping in densely vegetated areas.

In this paper, we provide an image based methodology and interpretation to produce engineering terrain maps. The standard procedures developed in the Trout Lake area will be used to produce produce image based terrain maps for the entire region.

Geology

We have selected the Trout Lake area to test several image processing and interpretation techniques that will provide us with the guidelines to produce standardized satellite based terrain maps of the region. The Trout Lake test site, located on the Canadian Shield, is an irregular bedrock-dominated terrain underlain by Precambrian igneous and metamorphic rocks. The area was last glaciated during the Wisconsinan Stage by the Nouveau Québec Sector of the Laurentide Ice Sheet. During deglaciation of the area, a large, ice-contact glacier-fed lake fronted the ice margin (glacial Lake Agassiz). The interactions of the glacier with the glacial lake controlled depositional environments and hence the distribution and types of landform, origin (glacial, glaciofluvial or glaciolacustrine), material type and sediment distribution.

Direct glacial sedimentation was deposited primarily under the glacier (till) or along the ice margin or grounding line (flowtills). Subglacial landforms include till plains and fluted till plains, drumlins and eskers. Ice-marginal landforms include one large end moraine, the Lac Seul Moraine [*5*] which contains deltas and subaqueous fans, and grounding line fans commonly referred to as DeGeer moraines. The Lac Seul Moraine is considered by some to be a very large grounding line fan produced during a catastrophic release of stored meltwater within the Laurentide Ice Sheet [*6*].

Sedimentation into the glacial lake was dominated by density underflows as a result of density differences between the incoming sediment-charged meltwater and the lake water. In this setting, the distribution of fine-grained sediments is controlled by lake-bottom topography and elevation, with sediments being preferentially deposited in topographically low areas. Abandoned shoreline features, beaches, bars and spits of glacial Lake Agassiz are well developed along the Lac Seul Moraine and other isolated hills of ice-contact sediments. Several levels of this lake are recorded. Wave and current action in this lake has also produced large areas of bedrock outcrops, washed clean of any pre-existing glacial sediment. Many small wetlands occur in bedrock-dominated terrain and in areas of low relief, grounding line fans. Larger areas of wetlands occur within the broad plains underlain by glaciolacustrine fine-grained sediments.

Methodology

DEM Generation

Hydraulically conditioned digital elevation models (DEMs) were created for the entire area at the Provincial Geomatics Service Centre of the Ontario Ministry of Natural Resources (MNR). MNR used 1:50 000 NAD83 digital topographic maps with a contour interval of 10 m and grid cell resolution of 25 m to derive the DEMs. Vectors marking lakes and rivers, derived from the DEM, were used to geometically rectify the imagery used in this study to the Universal Transverse Mercator (UTM) projection.

The DEM generation is a two step process beginning with the creation of a continuous digital drainage network. The first step involves the creation of virtual water segments, vectors that connect stream arcs through the many lake polygons in the area, using primarily automated techniques followed by checking the created network to remove loops and breaks. Finally watershed discharge points are located and all arcs within each virtual watershed are aligned such that the flow is toward the discharge point. The second step includes the checking and editing of the contour information based on 1:50 000 scale digital National Topographic Series maps and combining it with the created drainage network. The software program ANUDEM is used for the DEM construction. ANUDEM uses elevation information, the drainage network and water polygons in the interpolation process which creates a higher quality representation of the land surface as a result of incorporating drainage enforcement into the DEM algorithm. The DEMs were processed for the entire area in four batches and the individual 1:100 000 scale tiles of the test areas clipped out subsequently.

In addition to interpreting the DEMs directly, several derived products were created to also aid in interpretation of landform, material-type, relief and drainage conditions. These included hillshaded DEMs, slope, aspect, and elevation range models. These derived products are produced using ArcInfo™ and ArcView™ software packages. Landsat 7 Extended Thematic Mapper (ETM) multispectral and panchromatic data, IRS-1C panchromatic and RADARSAT data were also used in the interpretation. Initial geological analysis was based on individual image maps of one degree of longitude and one-half degree of latitude. Image maps were created using each of the four sensor types. A method was then devised to combine the sensors with the shaded DEM data, and image maps of combined products (ETM multispectral + DEM, ETM panchromatic + DEM, IRS 1-C panchromatic + DEM, RADARSAT SAR + DEM), were also created. The color or black and white image maps were printed out at a scale of 1:100 000. The maps were used in the field and interpreted visually by integrating the field information.

Creation of Image Maps

Geometric correction was performed on each of the three optical datasets separately, as each dataset had a different pixel size. The Landsat 7 ETM multispectral data had a pixel size of 30 m; the ETM panchromatic data had a pixel size of 15 m; and the IRS-1C

panchromatic data had a pixel size of 5.8 m. The data had been ordered from suppliers based on a NAD83 datum. For the optical data, the first step involved the collection of ground control points (GCPs) based on the drainage vectors provided with the DEM. GCPs were usually based on such features as the junction of two streams, the junction of a river and a lake or on very small features such as islands or lakes. An effort was made to obtain a network of GCPs that were as regularly spaced as possible throughout the image. As a new GCP was added to the GCP set, the residual, or RMS, error (total, x, y) for the GCP was displayed. Some of the points in the dataset were subsequently deleted or moved. GCP selection using the IRS 1-C 5.8 m resolution data was more difficult than with the TM data because the fine resolution of the imagery made it difficult to establish the exact location of a particular stream junction, lake junction or centre of an island or lake. The next step in the preprocessing of the optical data entailed the registration and projection based on a first-order cubic convolution resampling, using the ground control points. The resampled image retained the same pixel size as the raw image, but was rectified to the proper UTM zone and spatially registered to the vectors of the DEM.

The 25 m spatial resolution RADARSAT S2 data were orthorectified using GCPs based not only on the drainage vectors to establish planimetric locations, but also on elevation data derived from the DEM.

After the ETM, IRS-1C and RADARSAT images were geometrically corrected, they were enhanced (a linear contrast stretch was used for all of the images), then converted into individual images in map format. Each image was printed in color or in black and white at 1:100 000, with a scalebar, title and UTM coordinates, lat/long coordinates and a UTM grid. Several 3-band combinations of the Landsat TM data were compared, and a Band 4-5-7 composite image (as an R-G-B combination) was found to be the most useful for obtaining surficial geology information.

A shaded DEM can be obtained from the original DEM using GIS or image analysis software to simulate the effect of various sun azimuths and altitudes. For instance, one can reproduce the same illumination conditions that existed during the acquisition of an optical image. The sun azimuth and altitude, listed in the header of the optical image, can be used as inputs for the hill-shading routine.

Several methods were investigated of combining the shaded DEM image with each of the geometrically corrected images. The method yielding the most visually effective results involved adding scaled versions of the DEM and the images together. The shaded DEM was first resampled to the spatial resolution of the satellite image. The 16-bit shaded DEM and the 16-bit RADARSAT SAR data were scaled to eight bits. If scaling must take place, a linear contrast stretch is applied at the same time as scaling to the DEM and/or satellite image. The linear stretch is based on the minimum and maximum values in the histogram of the DEM and the satellite image. If the satellite image has not already been enhanced at this point, a linear contrast stretch must be applied to the individual channels.

The satellite image and the shaded DEM (which are now contrast-stretched and displayed at a common number of bits and at the same spatial resolution) must be multiplied by scalar factors which sum to 1. For instance, a factor of 0.625 applied to the three TM bands and a factor of 0.375 applied to the shaded DEM were found to result in an effective combined image. After the multiplication, the grey levels of the TM bands thus ranged from 0 to 159 (0 to 0.625(255)) and the grey levels of the shaded DEM ranged from 0 to 96 (0 to 0.375 (255)). Then the scaled satellite image and the scaled, shaded DEM are added together to produce an enhanced satellite-DEM combination, which is then displayed. A further contrast stretch is applied if necessary to the combination image. In the case of the TM Band 4-5-7 color image, three new channels were produced (scaled TM 4 + scaled DEM; scaled TM5 + scaled DEM; and scaled TM7 + scaled DEM). The three new channels were displayed as red, green and blue and a new contrast stretch was applied to the combination and saved. The scalar factors of 0.625 (for satellite image) and 0.375 (for shaded DEM) were applied to all of the optical and SAR images in the enhancement described previously to create enhancements. The resulting satellite image-DEM combination is intended purely for display, not for quantitative image analysis. It effectively combines the qualities of the original satellite image but adds an exaggerated topographic component to the data, which is effective for displaying features such as hills, valleys, ridges and linear discontinuities. The satellite-DEM enhancements were then printed out as 1:100 000 black and white and color image maps and used as ancillary data, in combination with the previously-described satellite image maps, to plan fieldwork and to derive surficial geology information. The enhancement procedures are outlined in Figure 1.

Discussion

Information on the various legend components: material, landform, topography and drainage, were derived from the interpretation of DEMs, their derivatives and remotely sensed imagery through the use of established landform/sediment relationships or models. The DEM and derivatives, such as hill-shaded DEM, slope, aspect and elevation range models in conjunction with Landsat 7, RADARSAT, and IRS-1C satellite images were used to produce 1:100 000 image-based terrain maps similar to the NOEGTS maps (Figure 4).

Analysis of the DEM, the derived hill-shade, slope and aspect maps aid in mapping the various landscape elements (Figure 2A to D). The hill-shaded DEM provides a general overview of the terrain (Figure 2B). Areas of glacial and glaciolacustrine sedimentation occur as smooth areas on this image; rock-dominated terrain has a very irregular appearance. The Lac Seul Moraine, for example, stands out as a dominant, linear topographic high. On the slope map (Figure 2C), the moraine has steep slopes on the proximal side and gentler slopes on the distal side. Ice-marginal deltas and subaqueous fans have their own characteristic shape but slope relationships are similar to that of the end moraine. The granular nature of the material associated with these landforms (predominantly ice-contact stratified sediments) and their positive relief makes these sites very well drained and allows them to support the growth of jack pine (*Pinus banksiana*).

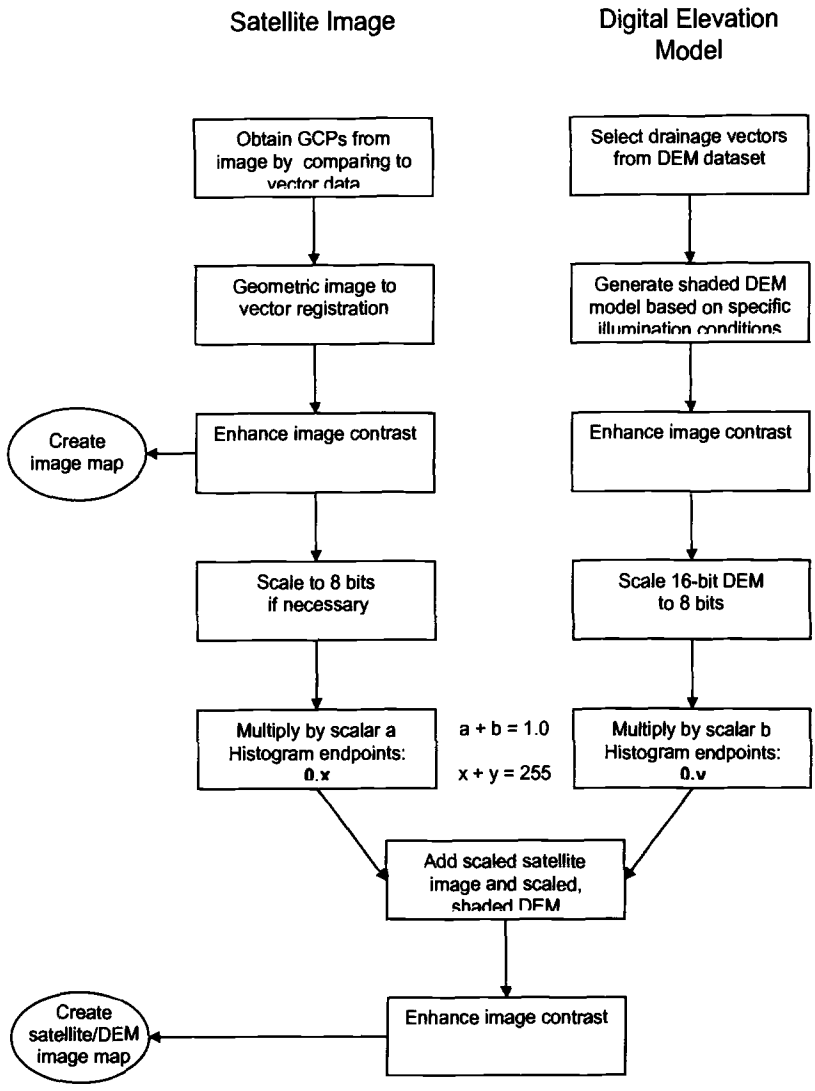


Figure 1 — *Flowchart outlining the creation of a merged satellite/DEM image map*

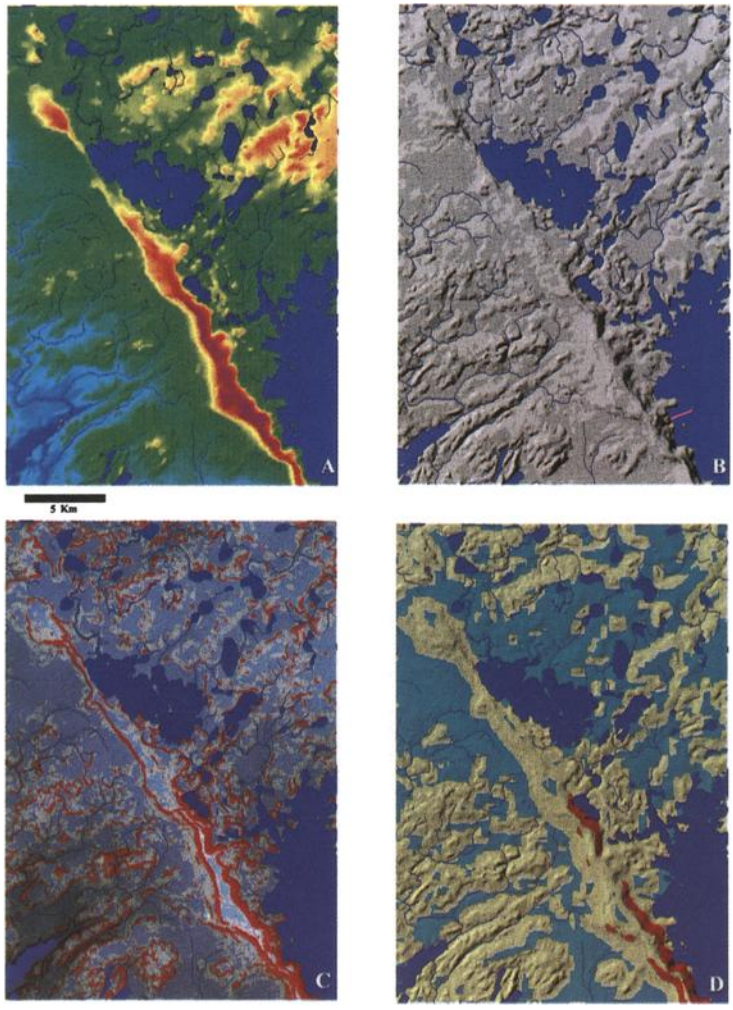


Figure 2 — *Examples of the Digital Elevation Model and derivative maps used in the interpretation of terrain units in part of the Trout Lake area: A – DEM, B – hill-shaded DEM, C – slope map with DEM as intensity on the colour and D – elevation range map with hill-shade as intensity on the colour.*

Glaciolacustrine sediments and subglacially deposited till form plains which are recognizable on the DEM and slope maps. In the Trout Lake area, plains of lower elevation tend to be underlain by glaciolacustrine materials and plains at higher elevations are commonly underlain by till. Only a few of the larger generally low-relief linear grounding line fans are recognizable on the DEM, however, these forms can be seen as promontories and/or linear islands within the area lakes and as linear tonal variations on the TM and IRS images.

Areas of irregular topography and short complex slopes are either bedrock-dominated areas or stagnant ice deposits (small fans and kettles). Separation of these two area types as well as several of the others discussed above can be done in association with the remotely sensed images based on forest variability.

The relief component of the legend can be derived using a 30x30 pixel filtering technique of the range in elevation values in the DEM. The resultant data can then be classed into the low (<15m; blue), moderate (>15m to <60m; yellow) and high (>60m; red) relief classes used in the NOEGTS maps (Figure 2D). Figures 3 A to D show the LANDSAT 7 image bands 4, 5 and 7 (Fig. 3A) integrated with the hill-shaded DEM (Fig. 3B), the IRS image (Fig 3C), and the IRS image plus DEM (Fig. 3D) of the Trout Lake area.

An engineering terrain and surficial materials interpretation maps of a smaller area are provided in Figures 4 and 5. Examples of most of the terrain units in the region are found in this Figure. The engineering terrain interpretation used materials, landforms, topography and drainage as the main classification elements. Landforms and topography are provided by the DEM and SAR images, whereas materials and drainage/moisture are provided by the remotely sensed information. There is a high geobotanical correlation between the well-drained sand plains and gravel and the dense jack pine forest cover; as such, these areas are easily identifiable terrain units (sgME/MD). Other terrain units such as organic terrains (pOT), rock knobs (RN) flat lake plain ((LP) and hummocky ground moraine (MG) are easily interpreted from remotely sensed data. The most useful image sources for terrain mapping are the Landsat/DEM composite and the IRS/DEM composite. Although the high resolution of the IRS does provide the details for terrain mapping at 1:50 000, we have decided to use the Landsat /DEM composite as an image base for cost considerations, suitable for terrain mapping at 1:100 000.

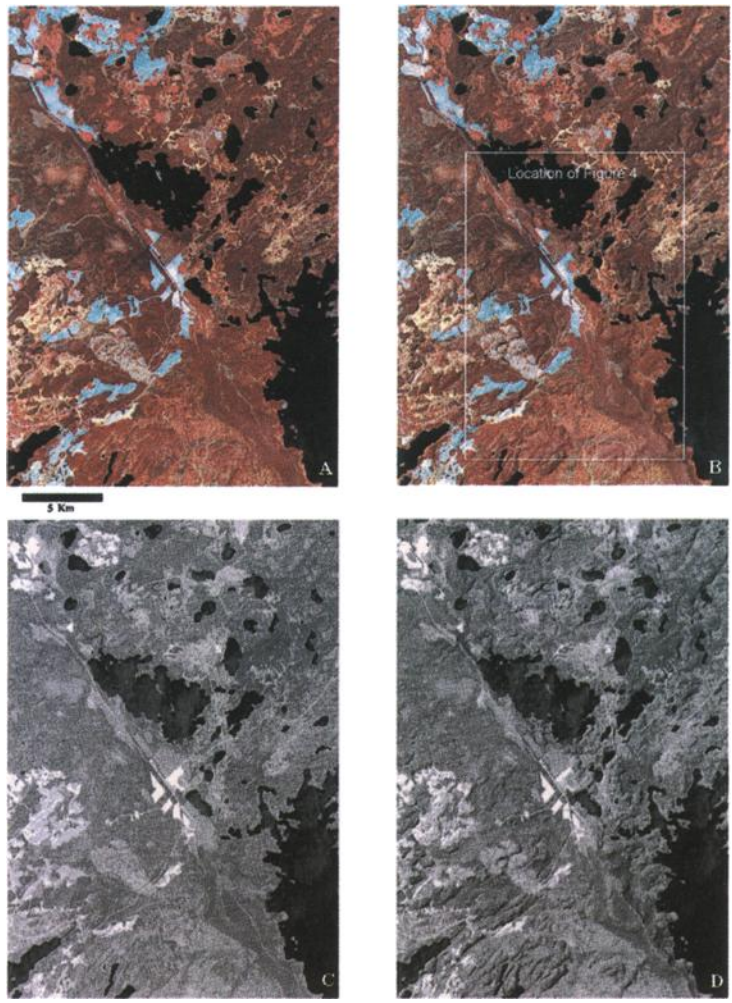


Figure 3 — *Examples of imagery and merged image/DEM products used to interpret terrain units. A – TM bands 4, 5 and 7. B – TM bands 4, 5 and 7 merged with a hill-shaded DEM. C – IRS-1c panchromatic band. D – IRS-1c panchromatic band merged with a hill-shaded DEM.*

Figure 4 — *A subarea of the Landsat 4-5-7/DEM merged image shown in Figures 2 and 3 (boundaries are outlined in Figure 3B) displaying engineering terrain units interpreted from remotely sensed imagery, the DEM and DEM derivatives. Terrain unit classification is based on four components: materials, landform, relief and moisture. As an example, the unit classified as sgME/MD (labelled above as (3)) has four components; sg indicates that it is composed of the materials sand and gravel; ME indicates that the landform it comprises is an end moraine; M (following the slash) indicates that the relief is moderate; and D indicates that moisture conditions are dry. A key is provided below.*

Materials	Landform	Relief	Drainage
sg– sand and gravel sL- lacustrine silt ts – silty till p - peat	ME – end moraine LP – lake plain GK – glaciofluvial kame MG – ground moraine	L – low M – moderate H – high	W – wet M – moist D - dry

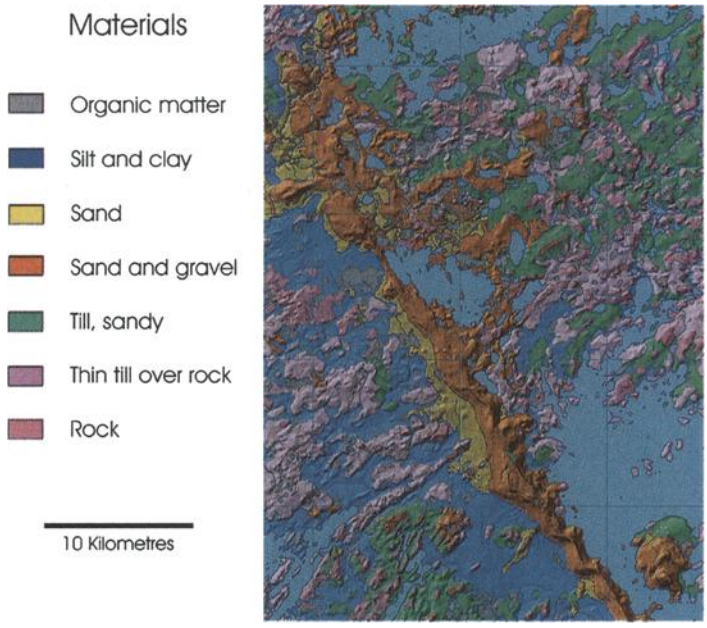


Figure 5 — *Surficial Materials Map of the Trout Lake area, Northern Canada.*

The DEM and the DEM/ TM composites (Figures 2 and 3) provide insights on the regional glacial history. For instance, the Lac Seul Moraine extends diagonally through the center of the Figures 2 and 3. A large ice-marginal delta, part of the Lac Seul Moraine (sgME/MD), is located in the southeast corner of the image. The delta provides information on the level of glacial Lake Agassiz during moraine formation (locally approximately 470m a.s.l.). Several lower shorelines of this lake can also be seen ringing the delta and moraine. At the northwest corner of the image an example of a subaqueous fan occurs along the morainic ridge.

Southwest of the moraine, the terrain consists of bedrock-controlled hills (RN) with varying thickness of till cover separated by low relief plains underlain by glaciolacustrine fine-grained sediments (LP). In places the cover of till is thick enough to subdue the bedrock topography and is then mapped as till. Northeast of the Lac Seul Moraine the terrain is dominated by smaller ice-marginal forms composed of sand and gravel as well as fields of DeGeer moraines composed of till, ice-contact gravel and sand or both (MG).

Conclusions

This study has shown that TM combined with DEM data can provide a standardized satellite image base map to interpret and overlay engineering terrain units at a scale of 1: 100 000 for large areas of the boreal forest regions in northern Canada. These methods can be applied in many other areas and will result in considerable savings in time and cost when compared to traditional air photo mapping methods

References

[1] Singhroy, V.H., Kenny, F.M. and Barnett, P.J., "Imagery for Quaternary Geological M apping in Glaciated Terrains", *Canadian Journal of Remote Sensing*, Vol. 18, No. 2, 1992, pp. 112-117.

[2] Singhroy, V.H., Slaney, R., Lowman, P., Harris J. and Moon, W., "RADARSAT and Radar Geo logy in Canada", *Canadian Journal of Remote Sensing*, Vol. 19, No. 4, 1993, pp. 338-351.

[3] Singhroy, V.H., "Environmental and Geological Site Characterization in Vegetated Areas: Image Enhancement Guidelines", *Remote Sensing and GIS: Applications and Standards,* ASTM Special Technical Publication #1279, V. Singhroy, D. Nebert, and A. Johnson, Eds., American Society for Testing and Materials, West Conhohocken, PA, 1996, pp. 5-17.

[4] Graham, D.F., and Grant, D.R., "Airborne SAR for Surficial Geological Mapping", *Canadian Journal of Remote Sensing*, Vol. 20, No. 3., 1994, pp. 319-323.

[5] Prest, V.K., "Red Lake–Lansdowne House Area, Northwestern Ontario, Surficial Geology", *Geological Survey of Canada, Paper 63-6*, 1963.

[6] Sharpe, D.R. and Cowan, W.R., "Moraine Formation in Northwestern Ontario: Product of Subglacial Fluvial and Glaciolacustrine Sedimentation," *Canadian Journal of Earth Sciences*, 27, 1990, pp. 1478-1489.

National Data

Wolfgang-Albert Flügel[1] and Michael Märker[2]

The Response Units Concept and Its Application for the Assessment of Hydrologically Related Erosion Processes in Semiarid Catchments of Southern Africa

Reference: Flügel, W.-A., Märker, M., **"The Response Units Concept and Its Application for the Assessment of Hydrologically Related Erosion Processes in Semiarid Catchments of Southern Africa,"** *Spatial Methods for Solution of Environmental and Hydrologic Problems—Science, Policy, and Standardization, ASTM STP 1420*, D.T. Hansen, V. H. Singhroy, R.R. Pierce, and A. I. Johnson, ASTM International, Conshohocken, PA, 2003.

Abstract: Proper management of valuable land resources is of paramount importance especially in regions affected by natural hazards. The sustainable development of land resources depends on the understanding of the processes and dynamics active within the landscape. In Southern African countries water shortage and water quality issues related to soil erosion are a major problem affecting the population in rural and urban areas. Consequently, during the last decade increasing attention has been focussed especially on such issues, and an increasing number of integrated hydrological and erosion studies, including the development and application of respective integrated regionalization concepts, is reflecting this development. The present study deals with the regionalization of spatially distributed hydrological related erosion processes in the catchments of the Mkomazi river (KwaZulu-Natal, South Africa) and the Mbuluzi-river (Kingdom of Swaziland). It was carried out within the framework of an interdisciplinary EU-funded project developing an Integrated Water Resources Management System (IWRMS) in semiarid catchments of Southern Africa. Within this project the concept of "Response Units (RUs)" was applied and adapted as Erosion Response Units (ERUs) to regionalize the distribution of hydrologically induced soil erosion in space and time. ERUs are landscape model entities identifying relative homogeneous hydrological related erosion processes, thus providing a spatially distributed model structure for regionalization. The examples from Southern Africa presented in this paper discuss the methods used to delineate such Response Units integrating remote sensing and GIS techniques.

1 Institut für Geographie, Friedrich-Schiller-Universität Jena, Löbdergraben 32, 07743 Jena, Germany.

2 Dipartimento di Scienza del Suolo e Nutrizione della Pianta – Università degli Studi di Firenze, Piazzale delle Cascine 14, 50144 Firenze, Italy.

Keywords: response units (RUs), hydrology, erosion, modeling, southern Africa, GIS, remote sensing

Introduction

The study presented herein deals with the delineation of Erosion Response Units (ERUs) in the catchments of the Mkomazi-river (KwaZulu-Natal, South Africa) and in the Mbuluzi river (Kingdom of Swaziland) and is part of the interdisciplinary EU-funded project *"Integrated Water Resources Management System (IWRMS)"* for semiarid catchments of Southern Africa. A central objective of *IWRMS* is to support catchment managers and decision makers to improve the regional strategic planning of water resources by optimizing water use, thus satisfying the demands of competing stakeholders while protecting water and land resources.

Response Units (RUs) are modeling entities of a regionalization concept representing landscape units characterized by unique association of geology, soil, topography, micro climate and land use, and are delineated by means of GIS overlay analyses. Controlled by these homogeneous configurations, of their natural capital components, they consequently have respective individual hydrological and erosion process dynamics. Similar approaches have been presented in the past by Leavesley et al. (1983) and Beran et al. (1990) and the concept of Hydrological Response Units (HRUs) was furthermore developed and validated by Flügel (1995) and Bongartz (1999). Adapted for modeling and regionalization of erosion caused by runoff the approach was furthermore extended by Flügel et al. (1999) and Märker et al. (1999) who introduced and tested the concept of Erosion Response Units (ERUs).

ERUs were delineated in both study catchments to regionalize the distribution of erosion processes and related landscape features to quantify the susceptibility of the river basins in terms of: (1) erosivity, that depends on rainfall properties such as intensity and duration, and (2) physiographic landscape characteristics such as land use, erodibility and geomorphology. By means of remote sensing techniques the distributed physiographic and anthropogenic catchment properties such as land use and settlements were classified. ERUs were delineated by overlaying and reclassifying the relevant data layers by means of a GIS accounting for the physiographic and management heterogeneity of the respective river basins.

Studied Catchments

Mbuluzi River

The Mbuluzi river basin originates in the Ngwenya hills in Swaziland and flows through the North/Central part of the country into Mozambique, running through all of

the physiographic regions of Swaziland. The river drains an area of about 3100 km^2, ranging from the border with Mozambique upwards (Fig. 1). It includes three landscape units: (1) the Highveld area (1066 - 1500 meters above see level) characterized by steep slopes with average gradients exceeding 18 percent; (2) the Middleveld (610 - 760 meters a.s.l.) with median slopes of 12%; and (3) the Lowveld (125 - 364 m a.s.l.) with gentle relief and moderate slopes of 3%. The mean annual rainfall ranges from 700 to 1200 mm (905 mm, Kwaluzeni), with the main rainfall season in summer (October to March). Kiggundu (1986) calculated a rainfall erosivity (EI_{30}) of 450 kJmm/m² hr (after Wischmeier and Smith 1978). The upper catchment is drained by the upper Mbuluzane River (A ~ 221 km^2) and the Mhlambanyoni River (A ~ 42 km^2). The latter was selected as a representative test area (Fig. 1) for the basin of the Mbuluzi river and is characterized by extensive and deep gully systems (Fig. 2).

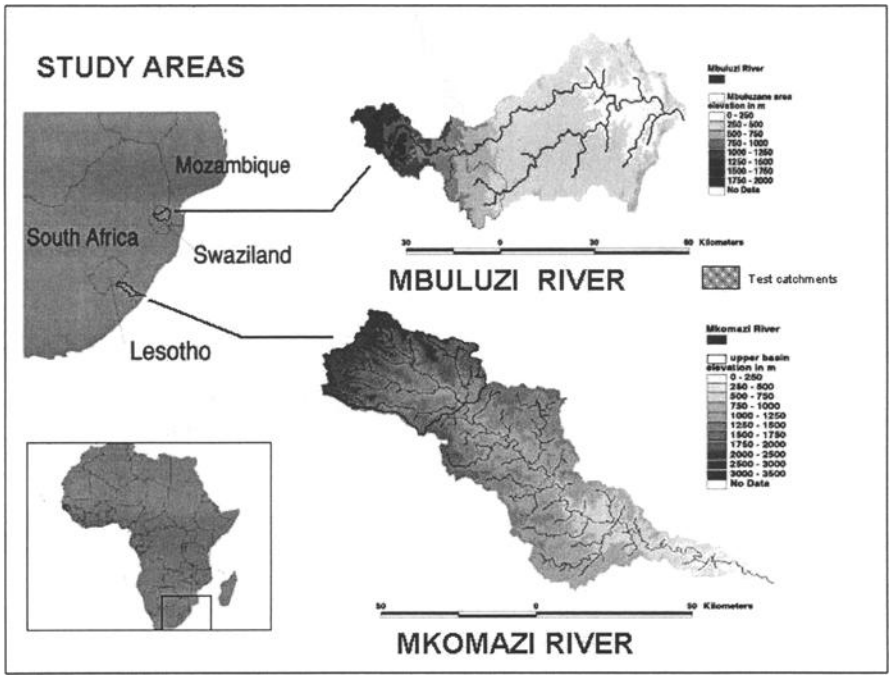


Fig. 1 - Location of the Mkomazi river and Mbuluzane river study catchments

The geology is dominated by granites and some areas are composed of precambrian sediments and volcanic outcrops. Granite and granitic gneisses with outcrops of dolerite and gabbro were found in the Middleveld. The Lowveld area is composed of sedimentary and volcanic rocks of the Karroo sequence. Due to intensive chemical weathering the lithology is decomposed into a thick granodioritic saprolite layer and a system of amphibolite and serpentite dykes (Felix-Henningsen et al. 1993; Mushala 2000).

The main soil types in the Highveld and Middleveld are deep, acid and well drained red and yellow ferrisolic and ferralitic soils, often with stone lines. In the lower Middleveld grey or red light textured soils from granite and gneiss are quite common. Meanwhile the Lowveld is characterized by weathered red, brown and black clays originating from basalt rocks (Murdoch 1970).

Land use in the upper parts of the Mbuluzi river basin is mainly rangeland and bushland, with some small-scale farms and subsistence agriculture around the rural dwellings. Intensive sugar cane plantations dominate the lower part, with irrigation and bush lands in the Lebombo region. The catchment is quite densely populated and as pasture is the dominant land use overgrazing is widespread.

Fig. 2 - Deep gully developed in saprolite material in the Mbothoma area about 15 km north of Manzini (Swaziland)

Upper Mkomazi river

The upper Mkomazi river catchment in the province of KwaZulu-Natal (South Africa) stretches from the Drakensberg escarpment down towards the Indian Ocean. The sources of the Mkomazi river are situated at altitudes of approximately 3300 m above see level in the upper Drakensberg area. The flow length is about 160 km from Northwest to Southeast, and the mouth of the Mkomazi river is located 40 km Southwest of Durban. The upper part of the catchment, which was selected as representative test area (Fig. 1) and tributary rivers are the Nzinga, Loteni, Mkomanzana and Elands river. The Mkomazi river drains an area of about 4400 km² and can be subdivided into four physiographic

zones: (1) the coastal lowlands up to 500 m a.s.l.; (2) the interior lowland area ("middle berg area") from 500 to 2000 m a.s.l.; (3) the mountain area up to 2500 m a.s.l.; and (4) the highlands, with elevations up to 3300 m a.s.l.

The climatic conditions in the semi-arid catchment are characterized by high seasonality with dry winters and a summer rainfall season. The mean annual precipitation varies between 1000 mm and 1800 mm in the upper Drakensberg down to values of less than 700 mm in the central areas, which are the most arid ones in the catchment (Seuffert et al. 1999; Tyson et al. 1976). The maximum rainfall occurs in the summer months between February and March. In the upper catchment the mean January temperature reaches 21 °C versus 24 °C on the coast (Durban). The monthly minimum temperatures vary between 10 °C in the "High Berg" area of the Drakensberg and 16,5 °C in the coastal parts. In the winter months, July to September, frost can occur especially in the mountain areas.

The geology of the upper Mkomazi river catchment is dominated by the Drakensberg escarpment. The oldest outcropping lithology are the Permian dark grey shales, siltstones and sandstones of the Escourt Formation. The successive formation, which also belongs to the Beaufort group, is the Triassic Tarkstad Formation, consisting of fine to medium grained sandstones and mudstones. Various sand and mudstones of the Triassic Molteno, Elliot and Clarens formations build up the next layers. A thick sequence of basaltic lava of the Drakensberg formation was deposited on top of this sedimentary series during the Jurassic period. All the above mentioned formations were disturbed by injections of dolerite as both dykes and sills. Some partly consolidated colluvial deposits (Masotcheni formation) and alluvial material of the Quaternary age were found in the middle and lower parts of the hill slopes (Linstrom 1979). These colluvial-alluvial materials are the result of several "cut-and-fill" cycles, probably due to climatic fluctuations of short duration (Botha 1996), and gully erosion is a quite common feature in this colluvial material.

The vegetation in the test area of the upper Mkomazi basin (Fig. 1) is a result of the altitude and the long history of grassland burning by man (Garland 1987). The vegetation can be distinguished into the middle mountain, subalpine and alpine belts, which are all dominated by *Themeda* species with pockets of shrub and woodland or *Protea* savanna (Killick 1963). The main land use in the test area is unimproved grassland with scarce patches of agriculture and forest plantations. Again overgrazing and consequent forms of sheet and gully erosion are the most eye catching features when traveling through this region.

Response Units Approach

Erosion

If one wants to characterize and quantify the erosion processes and their hydrological related dynamics at the catchment scale, both the physiographic properties and the human

management have to be considered. As not all of the integrated erosion processes are fully understood (e.g. subrosion, suffusion) it is consequently quite difficult to simulate the effects of erosion for the basin in total. However, it is commonly understood that erosion processes have to be analyzed according to their different temporal and spatial scales ranging from the micro scale rill-interrill erosion up to gully systems at the macro scale. As listed in Table 1 they are also related to respective time scales ranging between hours and days to years and decades.

Table 1- *Dominant time and spatial scales for water related erosion processes and landforms*

Erosionprocesses and forms	**Time scale**	**Spatial scale**
Gullies	Single event -continuous	Slope - catchment
Rills	Single event	Plot - slope
Interrill	Single event	Plot - slope
Tunnelling and piping	Single event - continuous	Slope - catchment
Badlands	Continuous	Slope - catchment
Mass movement	Single event - continuous	Slope - catchment

Response Units: HRUs and ERUs

To integrate these different spatial and temporal scales a regionalization concept such as the three-dimensional Response Units (RUs) is required. This is capable of associating the physiographic components and geomorphic erosion features distributed within the catchment with related process dynamics active on various time scales.

Hydrologically induced erosion processes and related geomorphic landscape features, such as gullies, can only be analyzed taking into account the hydrological dynamics of the drainage basin. As shown by Flügel (1995) and Bongartz (1999) the latter can be modeled successfully by applying the regionalization concept of Hydrological Response Units (HRU). This concept is based on the fact that specific associations of SVAT components (SVAT = Soil Vegetation Atmosphere Transfer interface) can differentiated within a catchment. They in turn can be associated with corresponding homogeneous hydrological process dynamics controlling the response of the HRUs to rainfall input, and distributing it into evapotranspiration, groundwater recharge and runoff generation. In other words: the way in which the system's unique HRU entities react on rainfall input depends on their specific configuration of SVAT components and the conditions of their respective status variables.

Such SVAT components are geology, soil, topography, microclimate and land use and are represented in a GIS by respective data layers. In terms of erosion such RU landscape entities having each unique association of SVAT components controlling their individual interlinked hydrological and erosion dynamics are defined as Erosion Response Units (ERUs):

"ERUs are distributed three dimensional terrain units, which are heterogeneously structured but each have unique configurations of SVAT components such as geology, soils, topography, micro climate and land use inducing respective individual erosion process dynamics having an internal variance which is negligible, if compared to adjacent units."

As can be concluded from this definition, ERUs also reveal information on the dominant morphologic erosion landscape features generated when transformation the precipitation input into the corresponding system's sediment output by surface and subsurface runoff. The ERU concept if applied to the entire river basin is representing a conceptual model, which conceives the catchment as a heterogeneous association of spatially distributed unique landscape entities, named ERUs. They have different erosion potential controlled by their individual configuration of SVAT components and associated human management. ERUs therefore can be applied as spatial modeling objects to regionalize erosion process dynamics while preserving the distribution of both the natural capital and the human environment.

Erosion Reference Units (ERefUs)

ERUs are delineated within a GIS by means of overlay analyses and reclassification techniques using erosion related criteria derived from a thorough geomorphic landscape analyses. This was done in the selected test areas of each study catchment (Fig. 1) by means of a combined approach integrating aerial photography interpretation and erosion classification with the findings from detailed field mapping. Based on the method proposed by van Zuidam (1985) and applying the criteria listed in Table 2 and Table 3 stereo-aerial-photographs in 1:30.000 scale (Mkomazi river: 1996; Mbuluzi river: 1990) were used to evaluate land degradation from classified vegetation cover. As a result Erosion Reference Units (ERefUs) were identified for each test area having specific erosion landscape features and associated intensities of land degradation.

Table 2 - *Frequency and density of rill and gully erosion features (after Van Zuidam, 1985)*

Depth (cm)	Spacing (m) between rills and gullies				
	< 25	25 – 50	50 - 150	150 - 500	> 500
5 – 50	Moderate	Slight			
50 - 150	Severe	Moderate	Slight		
150 - 500	Severe	Severe	Moderate	Slight	
> 500	Severe	Severe	Severe	Moderate	Slight

The classified ERefUs were transformed into digital format and georeferenced to the scale of the 1:50.000 topographical map yielding the erosion intensity map for the KwaThunzi test area in the upper Mkomazi catchment (Fig. 3), which was subsequently extended to the Mkomazi river basin in total.

Table 3 - *Classification of erosion types and status with respect to vegetation cover (modified from van Zuidam 1985)*

Erosion type	Status	Vegetation cover (%)	Erosion class
Slight sheet wash down	No	> 90	1
Rill-interrill; shallow gullies	slight	> 75	2
Rill-interrill; shallow–medium to medium deep gullies	slight moderate	> 75	3
Rill; medium deep gullies	moderate	51 – 75	4
Rill; medium deep to deep gullies; landslides	severe	26 – 50	5
Rill; deep gullies; badlands, severe mass movements	very severe	< 25	6

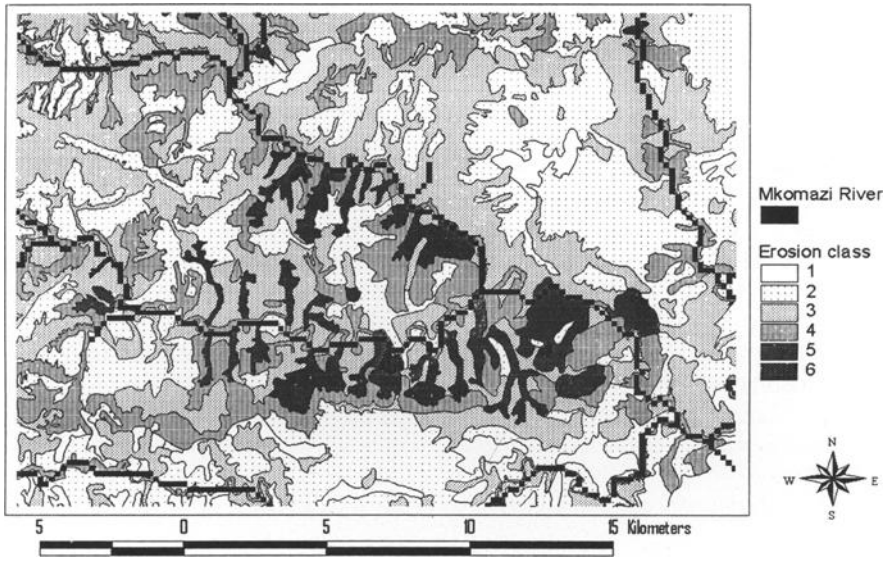


Fig. 3 - Erosion intensity map (1:50.000) for the KwaThunzi test area in the upper Mkomazi river catchment showing six distinct erosion classes defined in Table 4

GIS delineation of ERUs

The SVAT components required for the delineation of ERUs are topography, geology, land use, soils and erosion features, and are represented within the GIS by related data layers stored in raster format. Depending on the resolution of the available Digital

Elevation Model (DTM) they had pixel sizes of 200 x 200 m (4 ha) in the Mkomazi river basin and 25 x 25 m (0.0625 ha) in the Mbuluzi river catchment respectively.
These data layers were reclassified as follows applying the relevant erosion related parameters listed in Table 4 to reduce the number of classes within each data layer to an reasonable size:

(1) The national land cover classification (Thompson et al. 1996) was revised and reclassified into 6 land cover classes.
(2) Slope aspects were subdivided into four classes (North: 315-45°; East: 45-135°; South: 135-215° and West: 215-315°).
(3) Landscape morphology was parameterized by means of slope gradient (°), erosive slope length (m) and curvature (concave and convex). Erosive slope length was calculated and classified into three classes according to the cumulative frequency of the values: < 30 m; 30 – 60 m and >60 m. Because of the scarce DEM information (4 ha; 0,0625 ha pixel area) slope curvature was differentiated into the two groups concave and convex only.
(4) Soil texture is of paramount importance for erosion and for the Mkomazi river basin has been derived from published land type maps. Because of the high correlation between soil texture and geology in the Mkomazi river basin these two parameters were combined in a single layer of lithology and soil texture. For the Mbuluzi river catchment soil texture and lithologic information was obtained from the Swaziland Soil Map (Murdoch 1970). Contrary to the Mkomazi classification, here only five classes, that is, alluvium, clay, loam, rock outcrops, and sand were reasonable.

These reclassified data layers of SVAT components were combined with the ERefUs using overlay analyses and successively adding the different layers. The procedure is schematically shown in Fig. 4 for the test area of the Mbuluzane catchment in Swaziland. After each overlay the results were reclassified by referring them to the area of the respective ERefUs, and generalizing all classes with less than 2%.

Upscaling of ERU Information

The Erosion Response Units (ERUs) delineated by this procedure correspond to the present erosion dynamics and related landscape conditions. They comprise defined parameter combinations, applied as criteria to represent the heterogeneous system of the selected test areas in terms of morphology and erosion response. Consequently, the transfer of the ERU concept from the selected test areas to the entire river basins of the Mkomazi and Mbuluzane river will classify their erosion process dynamics and quantify their present distributed erosion status. Furthermore it will permit the evaluation and classification of the catchment's susceptibility to erosion according to the criteria listed in Table 4. The result of this exercise, that is, the regionalized erosion dynamics and status, is shown for the Mkomazi river basin in Fig. 5. Subsequently the catchments susceptibility to erosion was derived and classified from this spatial ERU distribution (Fig. 6) by applying the classification criteria listed in Table. 4.

Table 4 - *GIS data layers used for overlay analyses and associated Erosion Reference Units (ERefUs)*

Class	ERefUs	GIS data layer			
		Aspect	Landuse	Slope morphology	Geology and Soils
1	Slight sheet wash down	North	Unimproved grassland	Concave/convex < 1° ; > 60 m	Alluvium, Sand, Loam, Clay
2	Rill-interrill; shallow gullies	East	Shrub, bush, forests	Convex 1 - 5° ; > 60 m	Partly consolidated sediments (Masotcheni Formation)
3	Rill-interrill; shallow–medium to medium deep gullies	South	Wetland, open water	Concave 1 - 5° ; > 60 m	Basalt, Dolerite, Shales, Mud-Siltstone, Diamectites, Sand
4	Rill; medium deep gullies	West	Cultivated: commercial and subsidence	Convex 5 - 10° ; > 30 m	Basalt, Dolerite, Shales, Mud-Siltstone, Diamectites, Sand
5	Rill; medium deep to deep gullies; landslides	All	Urban	Concave 5 - 10° ; > 30 m	Gneiss, Granite, Diorite, Sandstone, Loam
6	Rill; deep gullies; badlands, severe mass movements	All	Degraded unimproved grass- and bushland	Concave/convex > 10° ; < 60 m	Gneiss, Granite, Diorite, Sandstone, Sand, Clay

Discussion

In the catchment of the Mbuluzi river in Swaziland altogether 40 ERUs and in the Mkomazi river basin in South Africa a total of 57 ERUs were classified. They represent the spatial distribution of the different erosion types and their intensities. The latter were classified according to Table 4 into six classes calibrated against specific ERefUs and can be interpreted as follows:

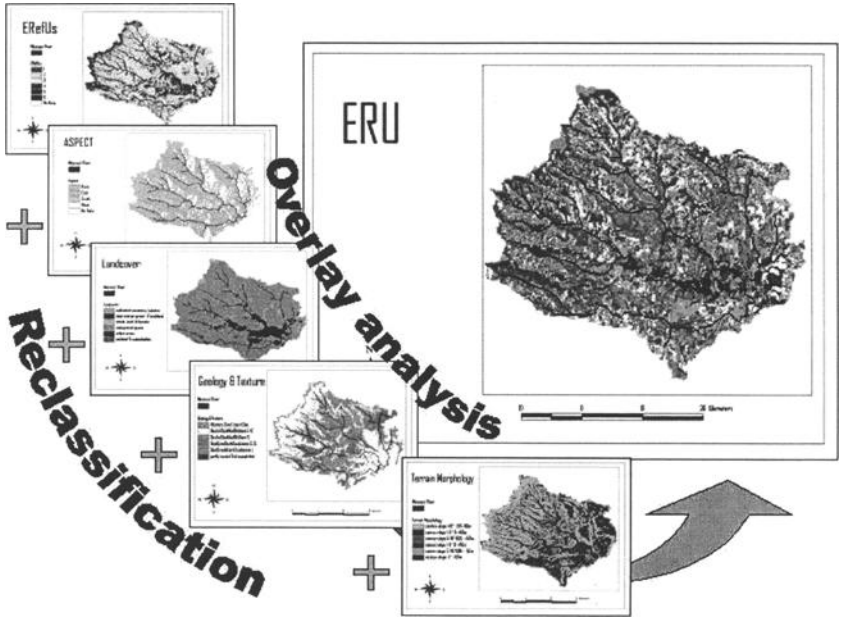


Fig. 4 - Overlay procedure of SVAT components used for the ERU delineation in the Mbuluzi river catchment

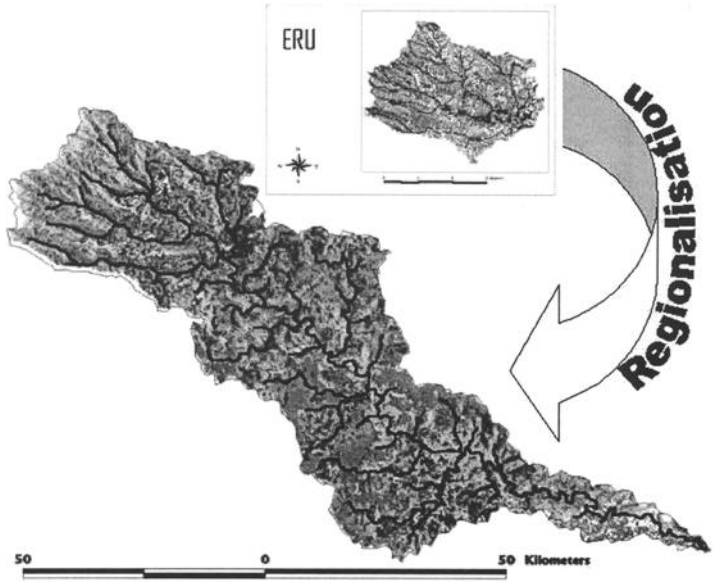


Fig. 5 - Regionalized erosion by distributed ERUs in the Mkomazi river catchment

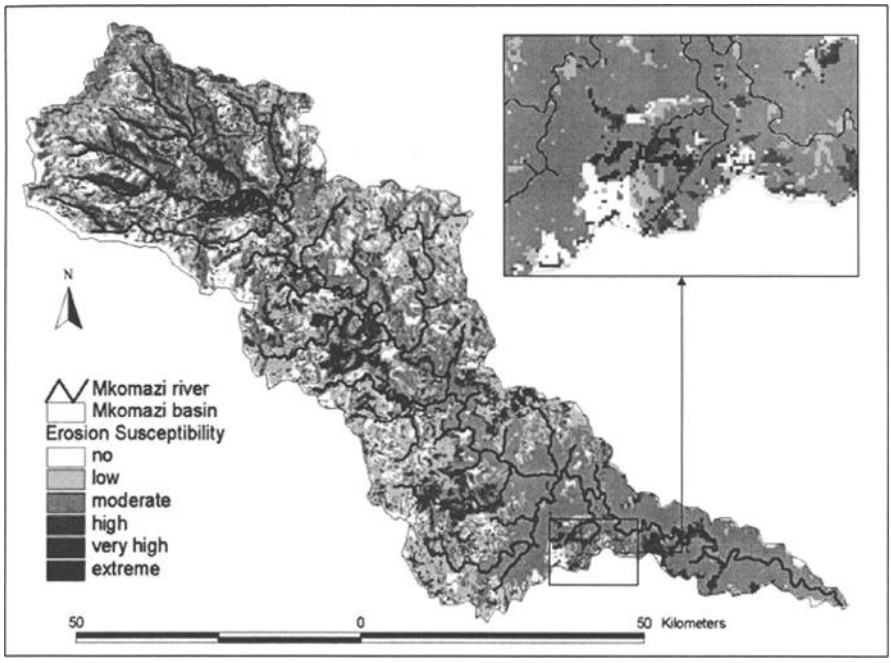


Fig. 6 - Erosion susceptibility derived from ERU distribution in the Mkomazi river catchment

(1) In the Mbuluzi river basin severe gully erosion was identified near Mbothoma in the upper part of the Mbuluzane river catchment and in the Mhlambanyoni catchment. These gullies are classified in erosion class number six and they are clearly visible in the 1:50.000 scale. About 8% of the Mhlambanyoni basin is directly affected by severe erosion (classes 4, 5 and 6), whereas 40% of the area shows significant erosion features such as deep linear and rill-interrill erosion of class 2 - 5. The zone of intensive erosion is running along a North-South stretching system of amphibolite/serpentite and dolerite/granophyre dykes. The main lithology consists of highly erodible saprolites (Mushala et al. 1994, Scholten et al. 1995). It is a densely populated area with a high livestock concentration. Consequently overgrazing is a common problem, especially on communal land like the Mbothoma area. Cattle tracks and pathways are visible in the aerial photographs and analyses of sequential photograph time series proved, that gullies often develop along such pathways and tracks.

(2) The upper Mkomazi catchment shows severe erosion mainly in the densely populated rural areas where the lithology consists of partly consolidated sediments of the Masotcheni formation or shales, siltstones and mudstones. Severe deep gully erosion (classes 4, 5, 6) affects about 13% of the entire upper basin. Whereas 90 % of the upper Mkomazi river area is affected by erosion (classes 2 – 6).

(3) Both catchments show a low erosion risk for southern exposed slopes, and very

high erosion risk is limited to unimproved grassland.

(4) Whereas in the Mkomazi river catchment the slopes with high erosion risk are steeper than 10° and have an erosive slope length shorter than 60 m in the Mbuluzi river catchment the slopes prone to erosion are 5 –10° steep and have an erosive slope length around 30 m.

(5) The differences in the ERU combinations observed in the two catchments are mainly due to differences in topography, soils and geology. The Mbuluzi catchment shows a high erosion risk for soils with loamy texture, often overlaying saprolites, whereas in the Mkomazi river catchment the partly consolidated sediments of the Masotcheni formation have a high erosion risk. Loamy clay sediments and basaltic lithologies, as well as sandy granodioritic material, are only subject to erosion when degraded by overgrazing.

As erosion type and intensity is specific to each individual ERU their spatial distribution also indicates the catchment's susceptibility to erosion, which was analyzed for the Mkomazi river basin and is shown in Fig. 6. In the upper part of the basin the erosion classes were derived from the mapped ERefUs. The erosion sites identified in the lower basin are mainly located in areas with dense informal settlements and lithologies with high k-factors, i.e. erodibilities according to Wischmeier and Smith (1978). Only in areas with rock outcrops or escarpments was the DEM not detailed enough to identify these areas correctly. As a result, the susceptibility delineated from the ERUs was greater than the one actually observed in this area during the field validation.

The window in Fig. 6 is showing a linear morphometric structure, which was classified by the ERU based approach as having a high erosion susceptibility represented by the erosion landscape feature "medium-deep to deep gully erosion." This theoretical derived classification was confirmed during a consequent field campaign by the existing gully shown in Fig. 7, and proved the validity of the developed approach.

Conclusions

The regionalization concept of Erosion Response Units (ERUs) was successfully applied within two southern African test catchments having different erosion processes dynamics and respective intensities. The ERU approach proved to yield realistic information about the dominant erosion processes, the catchment's susceptibility to erosion, and associated geomorphological landscape features. The criteria to delineate ERUs were identified in selected test areas using Erosion Reference Units (ERefUs) integrating information about the prevailing erosion processes and their respective intensities.

Detailed information about topography and lithology, as well as land cover information was derived with remote sensing techniques such as aerial photography interpretation (API). GIS was used to integrate the relevant SVAT components as respective data layers and to delineate ERUs by means of overlay analyses and reclassification.

The erosion processes considered in this study were interrill-rill erosion processes as well as deep gully erosion. The spatial distribution of dominant gully erosion in the

study areas provides evidence that the erosion dynamics must be included in the calculation of sediment yield, especially if the lithology consists of saprolites, which are highly vulnerable to erosion if degraded by overgrazing.

The methodical approach was validated in the lower part of the Mkomazi river basin where a linear erosion structure with high erosion susceptibility was classified by the ERUs. This classification was confirmed as a deep gully structure (Fig. 7) in a consequent field campaign

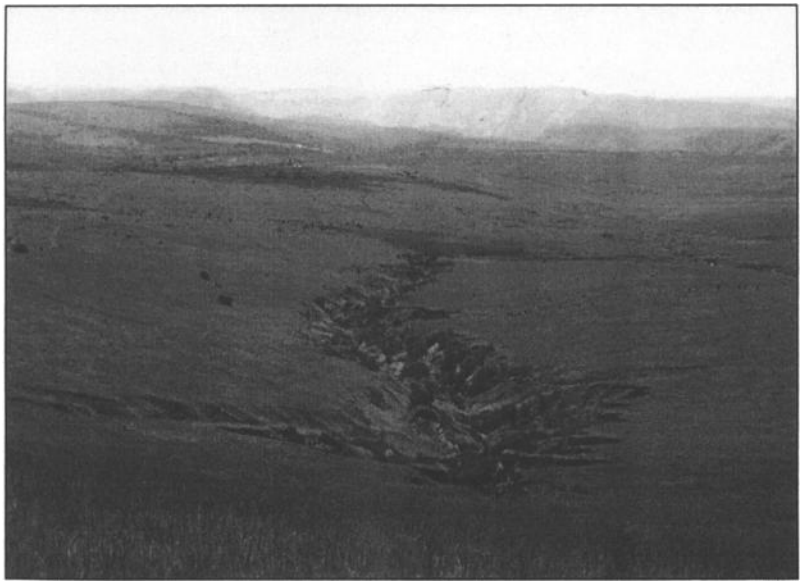


Fig. 7 – Gully system 20 km east of Ixopo (KwaZulu/ Natal, RSA) identified using the ERU concept

References

Beran, M. A., Brilly, M., Becker, A. and Bonacci O., 1990, "Regionalisation in hydrology," *Proceedings of the Ljubiljana Symposium* April 1990, IAHS Publication No. 191.

Bongartz, K., 1999, "Ableitung von Flächen homogener Systemantwort (HRUs) zur Parameterisierung hydrologisch relevanter Prozesse am Beispiel eines Thüringer Vorfluters," *Leipziger Geowissenschaften*, Vol. 11, pp. 123-128.

Botha, G. A., 1996, "The geology and palaeopedology of late quarternary colluvial sediments in northern Kwazulu/Natal," *Memoir of the Geological Survey of South Africa*, 83.

Felix-Henningsen, P., Schotte, M. and Scholten, T., 1993, "Mineralogische Eigenschaften von Boden-Saprolit-Komplexen auf Kristallgesteinen in Swaziland (Südliches Afrika)," *Mitteilungen der Deutschen Bodenkundlichen Gesellschaft,* Vol. 72, pp.1293-1296.

Flügel, W. A., 1995, "Delineating Hydrological Response Units (HRUs) by GIS analysis regional hydrological modelling using PRMS/ MMS in the drainage basin of the River Bröl, Germany," *Hydrological Processes* Vol. 9, pp. 423-436.

Flügel, W.A, Märker, M., Moretti, S., Rodolfi, G. and Staudenrausch, H., 1999, "Soil erosion hazard assessment in the Mkomazi river catchment (KwaZulu/Natal – South Africa) by using aerial photo interpretation,". *Zentralblatt für Geologie und Paläontologie;Teil I.*, Heft 5/6, pp. 641-653.

Garland, G. G., 1987, "Erosion risk from footpaths and vegetation burning in the central Drakensberg," *Natal Town and Regional Planning Commission Supplementary Report*, 20, Pietermaritzburg.

Kiggundu, L., 1986, "Distribution of rainfall erosivity in Swaziland," *Research paper* 22. University of Swaziland, Kwaluseni Campus. Swaziland.

Killick, D. J. B., 1963, "An account of the plant ecology of the Cathedral Peak area of the Natal Drakensberg," *Memoirs of the Botanical Survey of South Africa*, Vol. 32, Department of Agricultural Technical Service, Pretoria.

Leavesley, G. H., Lichty, R. W., Troutman, B. M. and Saindon, L. G., 1983, "Precipitation-runoff modeling system. User's manual," *U.S. Geological Survey Water-Resources Investigation Report* 83-4238, 207 p.

Linstrom, W., 1979, "1:250.000 Geological Series. Sheet 2928 Drakensberg," *Geological Survey*, Pretoria.

Märker, M., Flügel, W.-A. and Rodolfi, G., 1999, "Das Konzept der „Erosions Response Units" (ERU) und seine Anwendung am Beispiel des semi-ariden Mkomazi-Einzugsgebietes in der Provinz Kwazulu/Natal, Südafrika," *Tübinger Geowissenschaftliche Studien, Reihe D.: Geoökologie und Quartaerforschung. Angewandte Studien zu Massenverlagerungen,* pp. 231-241, Tübingen.

Murdoch, G., 1970, "Soils and Land Capability in Swaziland," *Swaziland Ministry of Agriculture*. Mbabane.

Mushala, H. M., 2000, "An investigation of the spatial distribution of soil erosion in the Mbuluzi river basin of Swaziland," *UNISWA Research Journal of Science & Techniques* 3 (2), pp. 32-37.