Second Edition

AIRPORT SYSTEMS

PLANNING, DESIGN, AND MANAGEMENT

Richard de Neufville ▪ Amedeo Odoni

with contributions by Peter Belobaba and Tom Reynolds

Mc Graw Hill Education

Airport Systems

About the Authors

Dr. Richard de Neufville is professor of Engineering Systems and professor of Civil and Environmental Engineering at MIT. He is known for his development of engineering systems analysis and many texts, most recently, *Flexibility in Engineering Design*. He has consulted and taught on airport planning "on all continents except Antarctica" for over 40 years.

Among his many honors are the Francis X. McKelvey award for Aviation, the FAA award for Excellence in Education (with Amedeo Odoni) and several other teaching awards, the Irwin Sizer award for the Most Significant Contribution to MIT Education, the Chevalier des Palmes Académiques (France), and an honorary doctorate from the Delft University of Technology.

Dr. Amedeo R. Odoni is professor of Aeronautics and Astronautics and professor of Civil and Environmental Engineering at MIT. He specializes in the use of operations research and other quantitative methods in planning, designing, operating, and evaluating airport and air traffic management systems. Over the years he has consulted at Amsterdam, Athens, Boston, Milan, Montreal, Munich, New Delhi, New York, Sydney, Stockholm, and many other airports, as well as several civil aviation authorities.

He is a member of the National Academy of Engineering, a fellow of INFORMS, and the recipient of an honorary doctorate from the Athens University of Economics and Business and of many awards for his research and teaching. He has served, among other positions, as codirector of MIT's Operations Research Center and of NEXTOR, the National Center of Excellence in Aviation Operations Research, established by the FAA in 1996.

Dr. Peter P. Belobaba is principal research scientist at MIT, where he teaches graduate courses on the Airline Industry and Airline Management. He is program manager of MIT's Global Airline Industry Program and director of the PODS Revenue Management Research Consortium. He has a Ph.D. in Flight Transportation Systems from MIT. He is a lead author and editor of the recently released book, *The Global Airline Industry*.

Dr. Belobaba has been involved in research related to airline economics, pricing, competition, and revenue management since 1985. He has worked as a consultant on the development and implementation of revenue management systems at over 40 airlines and other companies worldwide. He has also published articles in a variety of journals, including *Airline Business, Operations Research, Transportation Science, Journal of Revenue and Pricing Management*, and *Journal of Air Transport Management*.

Dr. Tom G. Reynolds specializes in advanced air transportation systems development and environmental mitigations at MIT Lincoln Laboratory, where he works in the Air Traffic Control Division. He has particular interest in the development of advanced technologies and operations for improving efficiency and mitigating environmental impacts of aviation, and has helped deploy many improvements at major international airports in the United Kingdom and the United States. He has a Ph.D. in Aerospace Systems from MIT, and has worked on the research staff at MIT, the University of Cambridge, and British Airways Engineering. He has won several national awards including the AIAA Orville & Wilbur Wright Graduate Award and was a U.K. Fulbright Scholar.

Airport Systems
Planning, Design, and Management

Richard de Neufville

Amedeo R. Odoni

with contributions by Peter Belobaba and Tom Reynolds

Second Edition

New York Chicago San Francisco
Lisbon London Madrid Mexico City
Milan New Delhi San Juan
Seoul Singapore Sydney Toronto

ISBN: 978-0-07-177059-0

MHID: 0-07-177059-3

e-Book conversion by Cenveo® Publisher Services

Version 1.0

The material in this eBook also appears in the print version of this title: ISBN: 978-0-07-177058-3, MHID: 0-07-177058-5.

To

Ginger, Julie, and Robert de Neufville

Eleni Mahaira-Odoni

Mary Belobaba

The Reynolds family

Contents

Preface

Welcome to the second edition of *Airport Systems Planning, Design, and Management*! Recognizing the widespread adoption and worldwide use of the first edition (translated into two non-Roman scripts, Greek and Mandarin), we have thoroughly rewritten the text to make it as useful as possible to our readers. This new version deals with the major shifts in the airport/airline industry that have occurred in the past decade. It contains significant new material and is now thoroughly up-to-date.

This book is about creating effective and efficient airports. To achieve this objective, professionals need to consider the whole problem, from the initial planning, through the design of the facilities, to the ultimate management and operation of the airport. The text uniquely integrates these phases, in contrast to other books that only deal with parts of the problem.

Specifically, the book

- Begins by setting the industry context that affects airport development
- Follows with chapters on the systems aspects of airport development, such as dealing with uncertainty in forecasts, the environment, and financing
- Then deals with the airside issues (runway, airfield capacity, and delays) and the landside (design of terminals and airport access)
- And, ends with reference material supporting the overall text (such as forecasting, flows and queues, etc.)

The text takes a worldwide perspective throughout. It thus serves both North American and international users.

What Are Its Distinctive Themes?

The book emphasizes helping the readers understand the issues. This is important. To deal with them effectively and efficiently, professionals need to appreciate how and why things work as they do. This is especially true because circumstances for airport development differ from place to place, and are constantly evolving. The book's approach contrasts with other books that focus on current rules and formulas.

The book also takes a systems approach. It recognizes that the different aspects of airport development affect each other. Designers should therefore not consider them in isolation. For example, efficient airport operations depend on thoughtful planning of the airfield and effective design of the terminal area. Likewise, good management of the facilities reduces the need for capacity and improves the planning process. The text uniquely provides an integrated approach to airport development.

Motivation for Revision

The airport industry has evolved considerably in the decade since the 2003 edition.

- The airline demand for airport services has changed as major airlines have consolidated and the low-cost carriers have become major factors in the industry.
- Environmental regulations and international rules have greatly shifted emphasis.
- Airport technology has changed—new types of aircraft, satellite-based air traffic control, security controls, and information technology serving passengers and bags.
- Techniques and models for planning, designing, and managing airports have advanced considerably.
- New research results are available.

Intended Audience

The book is for all those with a major interest in airport planning, design, and management. This includes owners and operators; architects and engineers; government officials; airlines, concessionaires, and other providers of airport services; travelers and shippers; and neighbors and communities, as well as members of the public. Readers need no specific experience or skills to use it. A serious interest in the topic is all that is required to make good use of the text. The authors recognize that most people become involved with airport planning, design, and management later in their careers and come from a broad range of professional backgrounds.

The book has proven to be useful worldwide. It stresses universally applicable concepts and approaches to airport problems. It refers to several different sets of international and national standards on the airside and the landside and points out both similarities and differences in current airport practices around the globe. The text draws heavily on worldwide experience to bring out the best available approaches to each issue.

The text assumes that readers need to deal with current issues in airport planning, design, and management. It focuses on the actual problems that arise, and on practical, effective ways of dealing with them. Theory and methodology appear only to the extent that they are relevant and useful. The authors have tried to illustrate theory and methods with appropriate examples wherever possible.

The text is suitable for students in planning and design curricula. The authors have used the material that has led up to this book since around 1980, in both their courses at MIT (the Massachusetts Institute of Technology) and professional short courses worldwide, "on every continent except Antarctica."

The Content

The book concentrates on significant commercial airports, those with more than about 1 million passengers a year. It only considers smaller airports or military bases when they provide a region with significant current or prospective capacity to handle airline traffic. Likewise, it does not deal with special facilities such as heliports or seaplane bases.

The text covers both the development and management aspects of airports. Systems design recognizes that the costs of building and operating a major facility such as an airport are comparable. Good planning and design thus makes sure that the physical configuration of a project facilitates operations and that the management procedures enable owners to avoid unnecessary capital costs.

The text discusses in detail each of the major development topics:

- Airport site characteristics
- The layout of runways, taxiways, and aircraft aprons
- Design of passenger buildings and their internal systems, including security
- Analysis of environmental impacts
- Planning for ground access to the airport

It also treats the operational and managerial issues of the following:

- Air traffic control
- Management of congestion and queues
- The determination of peak-hour traffic
- Environmental impacts
- Financing, pricing, and demand management

Competition increasingly provides the context for commercial airports. The success of any airport depends most importantly on its advantages compared to other airports, now and as they may be in the future. The text thus carefully describes competition between airports, both within and between metropolitan areas, as well as in the context of airline networks operating nationally, internationally, and globally. It also discusses how international trends in the industry might change the competitive picture.

Dynamic strategic planning is the approach used to bring these specific topics together. It is the modern method for designing complex systems over time. It builds upon the understanding that all forecasts are unreliable, uses the procedures of decision and options analysis of risky situations, and incorporates the economics of financing. The text covers these topics as needed. The overall object is to plan, design, and manage airports so that they can respond flexibly to the unknown, uncertain future conditions.

Format

The book should be easy to understand. It is free of unnecessary mathematical expressions or technical terms. Most of the material is easily accessible to the broad range of persons concerned with airport systems planning, design, and management: engineers and architects as well as managers who do not have a technical background. We have made the text easy to use by the many airport professionals who are neither engineers nor native speakers of English.

The book features numerous examples illustrating the application of the concepts and methods. It draws upon actual cases from the authors' worldwide experience. The emphasis throughout is on dealing effectively with real issues.

A reference section presents basic theory and, in some cases, background mathematics. Persons who do not need this complementary material can skip this section. Users can combine this reference material with chapters on specific issues to meet their need for information on a particular topic. As the following User's Guide describes, readers can tailor the material to their requirements.

Richard de Neufville	*Amedeo Odoni*
Peter Belobaba	*Tom Reynolds*

Acknowledgments

Over decades, many people helped the authors learn about airport systems planning, design, and management. Recognizing that we cannot possibly list them all, we particularly want to thank those who played important roles in helping us develop material for both this second and the 2003 first editions.

From government, our collaborators have included Larry Kiernan, Ashraf Jan, Richard Doucette, and Danielle Rinsler, U.S. Federal Aviation Administration; Zale Anis, Volpe National Transportation Systems Center; Jean-Marie Chevallier, Aéroports de Paris; Dr. Lloyd McCoomb, Greater Toronto Airport Authority; and Flavio Leo, Massport.

Numerous consultants and practicing professionals helped ensure that the work reflected actual conditions in the field. These include John Heimlich, Katherine Andrus, Tom Browne, Patricia Edwards, Russell Gold, Paul McGraw, and Robert Zoldos, formerly of Air Transport Association, now Airlines for America; Richard Marchi, Airports Council International; Regine Weston, Weston-Wong Associates and Arup; Christina Avgerinos, Ascendant; Dr. Philippe Bonnefoy, Booz Allen Hamilton; Dr. Husni Idris, Engility Corporation; Cliff King, Louis Berger Associates; William Woodhead, formerly of Kinhill Engineers, Australia; Harley Moore, Lea+Elliott; Dan Kasper, LECG; Robert Weinberg, Marketplace Development; William Swedish, MITRE; Shota Morita and his Airports Department, Pacific Consultants International, Tokyo; Steve Belin, formerly of Simat Helliesen and Eichner; Thomas Brown, United Airlines; and Martin Plazyk, VanDer-Lande Industries.

We also gratefully acknowledge the contributions of academic colleagues: Prof. Alex de Barros, University of Calgary; Prof. Rigas Doganis, Cranfield University, U.K.; Prof. John-Paul Clarke, Georgia Institute of Technology; Prof. Robert Caves and Dr. Ian Humphreys, Loughborough University, U.K.; Prof. Arnold Barnett, Prof. R. John Hansman, Prof. John Miller, and Prof. Ian Waitz, MIT; Dr. Christoph Wollersheim, MIT; Dr. Steven Bussolari and Dr. Hayley Davison Reynolds, MIT Lincoln Laboratory; and Prof. Andreas Schäfer, University College London, U.K.

We especially appreciate the assistance of the students who helped us prepare this second edition: Erika Kutscher, Hèctor Fornes Martinez, Kamala Shetty, Vincent Surges, and Alexander Wulz.

We gratefully acknowledge the FAA-sponsored Airport Cooperative Research Program. The ACRP has sponsored a great body of work since its beginnings in 2005. The fruits of

these efforts are freely available through the Transportation Research Board of the U.S. National Academy of Sciences. These represent many significant contributions to airport planning that are essential to being up-to-date in the field. This text draws extensively on these latest results, and we are privileged to be the first to integrate them in a textbook.

User's Guide

You can create your own book. Readers can tailor the material to their own needs. Persons interested in a specific topic can put together a self-contained set of chapters that will give them what they need to know about that subject. Architects interested in the design of passenger buildings, for example, can assemble an integrated guide to the subject. They can put together the chapter on that topic and the supporting chapters on the analysis of queues and peak-hour analysis. Users do not have to get involved in topics of no current concern and can concentrate on their immediate interests.

Readers can likewise tailor the material to their own skills or depth of interest. Many readers will use the book to get help on a specific project. They will initially want information relevant to only one topic, such as airport financing or airport access, and will be able to get it. The chapters on specific problems, the design of passenger buildings, for instance, are self-contained and provide the necessary guidelines in a way that anyone should be able to understand. Users who do not need the supporting reference material, either because it is not relevant to their job or they know it already, can simply skip it.

The text is modular, in short. Its chapters can be assembled in different ways for a variety of needs. This organization is possible because many of the methods used in airport systems planning are common to several different topics. An understanding of the behavior of flows and queues of traffic, for example, is necessary for the detailed design of both runways and passenger buildings. The reference sections dealing with specific methods fit in with several chapters that deal with specific problems.

How to Do It

To appreciate how to tailor the material to your own needs, it is useful to look at the organization of the material. The mode of use then becomes clear.

The text consists of two distinct blocks. As the table of contents indicates, the first block consists of substantive chapters devoted to specific topics in systems planning and management, airside and landside. The second block, Part 5, provides reference on methods of analysis such as forecasting and queuing theory. These reference materials provide in one place coherent discussions of procedures that apply to several of the substantive chapters.

Recommended Combinations

Each of the big blocks on systems planning, airside and landside, is a self-contained unit. Readers can approach them independently of the others. This arrangement should be useful to persons with responsibilities or interests especially in those fields. For example, managers and government officials might focus on particular topics: planners on systems planning, aviation and air traffic control specialists on the air-side, and architects and civil engineers on the landside.

All readers may be interested in Chaps. 1 through 3, which provide context on the future of the airport/airline industry and give an international perspective. They may then choose topics according to their interests. Referring to the following Menu, the authors suggest these packages for readers with broad interests:

- Systems Planning: the block in column A plus column B under Risk
- Airside: the block in column A plus column B under Variable Loads
- Landside: the block in column A plus column B under Detailed Design

Menu of Chapters

(A) Issues	(B) Reference
System Planning	**Risk**
Dynamic Strategic Planning	Data Validation
Multi-airport Systems	Forecasting
Aviation Environmental Impacts and Airport-Level Mitigations	
Organization and Financing	
User Charges	
Airside	**Variable Loads**
Airfield Design	Flows and Queues at Airports
Airfield Capacity	Peak-Day and Peak-Hour Analysis
Airfield Delay	

Demand Management

Air Traffic Management

Landside

Configuration of Passenger Buildings

Overall Design of Passenger Buildings

Detailed Design of Passenger Buildings

Ground Access and Distribution

Detailed Design

Forecasting

Flows and Queues at Airports

Peak-Day and Peak-Hour Analysis

PART I

Introduction

CHAPTER 1

The Future of the Airport and Airline Industry

Airport systems exist in the context of their major clients, the airlines. To build airport facilities that will perform effectively over the 30 to 50 years of their lifetime, it is necessary to appreciate this context. Understanding the state of the airport/airline industry in the early twenty-first century gives a perspective on its future. This is the starting point for a forward-looking text on airport systems planning.

Three trends dominate the airport/airline industry in the early twenty-first century:

1. *Long-term growth*, which has been about 4 percent per year worldwide. This implies a doubling of traffic about every 15 to 20 years and drives the demand for expansion and improvement. It also leads to the development of new airports, of multiple airport systems in metropolitan areas, and of niche airports serving leisure traffic or cargo.
2. *Organizational change*, as economic and political deregulation continues to spread worldwide. Economic deregulation creates opportunities for low-cost and integrated cargo airlines to grow, impels governments to privatize their airlines and airports, and leads traditional airlines to consolidate. Political deregulation, such as open-skies agreements, enables new markets, changes in traffic patterns, and increases competition. These ongoing changes in airport clients and their needs make for an uncertain, instable future. Airports will consequently have to plan flexibly so that they can adapt easily as required.
3. *Technical change*, most obviously in aircraft and air traffic control, but also contextually, particularly as regards information technology that continues to redefine the way we do business. These developments increase the efficiency and the capacity of airport facilities and processes. Airports need to adapt to these new opportunities as they occur.

Taken together, these trends are substantially changing the context, objectives, and criteria of excellence and efficiency for airport systems planning and design.

1.1 The Airport Industry in the Early Twenty-First Century

Airports and air transport continue their exciting long-term growth. The industry is large, innovative, and has excellent prospects. We need to appreciate this historical base before launching into the future. Moreover, the industry is in the midst of substantial organizational and technical changes that are redefining the practice of airport systems planning and design.

The industry is large. As of 2012, it involves about 2.5 billion airline passengers worldwide plus large amounts of cargo. Its annual revenues are more than U.S. $0.5 trillion (one million million dollars). The world airlines operate approximately 12,400 major jet aircraft, valued in the hundreds of billions of dollars. The annual investments in airport infrastructure come to about $10 billion a year. To put these figures in perspective, the industry moves the equivalent of well over a third of the world's population every year, and its revenues are close to 40 percent of the gross domestic product of the United States. By any measure, this is an important activity.

The industry is actively growing. From 1990 to 2012, the worldwide long-term growth rate in the number of airline passengers has been about 4 percent a year—averaging periods of stagnation and boom. During that period, global passenger traffic grew by 120 percent; it more than doubled. As of 2012, this growth was mostly occurring in Asia, where air transportation is becoming increasingly affordable to its large populations. In the first decade of the twenty-first century, annual passenger traffic grew at an average of 9 percent in Asia, 5 percent in Europe, and 1 percent in North America.

Airport planners thus routinely have to deal with the possibility of 25 to 100 percent increments in demand. This is because the planning horizon for large-scale infrastructure projects is normally between 10 and 15 years, because of the need to create the designs, assemble financing, and proceed successfully through political and environmental reviews.

The growth in air transport translates into major airport projects. About a dozen major programs for airport development, costing over a billion dollars each, are typically under way at any time. Table 1.1 illustrates the situation. Naturally, many smaller projects are ongoing simultaneously.

City	Project
Athens	New international airport
Atlanta	New runway, new international building
Bengaluru	New international airport
Bangkok/Suvarnabhumi	New international airport
Barcelona/El Prat	New terminal and new runway
Beijing	New terminal, runway, and rail link to city
Berlin/Schönefeld	Complete airport overhaul
Chicago/O'Hare	New runways, terminals
Dallas/Fort Worth	New automated people mover
Delhi	Complete airport overhaul
Dubai	New passenger terminal, cargo terminal
Hyderabad	New international airport
London/Heathrow	$7 billion Terminal 5, Terminal 2
Nagoya/Chubu	New international airport
New York/Kennedy	JetBlue terminal and railroad connection
Madrid/Barajas	New terminals and a runway
Paris/de Gaulle	New terminals
Seoul/Incheon	New international airport
Singapore	New Terminal 3 and low-cost terminal
Shanghai/Pudong	New international airport
Toronto/Pearson	Complete airport overhaul
Tokyo/Haneda	New runways, terminals

TABLE 1.1 Some Billion-Dollar-Plus Airport Projects of 2002–2012

Airline/airport traffic has been concentrated in the United States. It is the locus of close to half the worldwide air transportation and airport activity. U.S.-based airports and airlines dominate their competitors in size. In 2011, U.S.-based airlines accounted for 7 of the top 10 airlines (Table 1.2). Likewise, many of the busiest airports in the world in terms of the number of passengers have been in the United States. In 2011, U.S. airports occupied 7 of the 20 top spots (Table 1.3). The U.S. share of the world traffic has, however, been decreasing as traffic grows in Europe, the Middle East, and Asia. Its market share fell from about 40 percent in 1990 to around 30 percent in 2011.

Major Airlines	Associated Airlines	Aircraft
American	**American Eagle**	**861**
Delta		**762**
Southwest/Air Tran		**716**
United/Continental		**703**
FedEx	**Includes charters**	**654**
Lufthansa	Austrian, CityLine, Brussels, Swiss	539
UPS	**Includes charters**	**526**
Air France	KLM, Cityhopper	422
US Airways	**US Airways Express**	**382**
China Southern		377
British Airways	Iberia, Iberia Express	343

Sources: www.airfleets.net; www.fedex.com; www.pressroom.ups.com.

Table 1.2 U.S.-Based Airlines Were the Largest in the World in 2011 (ranked by size of fleet)

	Airport	Passengers (millions)	Movements (thousands)
1	**Atlanta**	**92**	**950**
2	Beijing/International	77	518
3	London/Heathrow	69	455
4	**Chicago/O'Hare**	**67**	**883**
5	Tokyo/Narita	62	343
6	**Los Angeles/International**	**62**	**667**
7	Paris/de Gaulle	61	500
8	**Dallas/Fort Worth**	**58**	**652**
9	Frankfurt/International	56	464
10	Hong Kong/Chek Lap Kok	53	316
11	**Denver**	**53**	**630**
12	Jakarta/Soekarno-Hatta	52	310
13	Dubai/International	47	307
14	Amsterdam/Schiphol	50	402
15	Madrid/Barajas	50	434
16	Bangkok/Suvarnabhumi	48	270
17	**New York/Kennedy**	**48**	**400**
18	Singapore	47	369
19	Guangzhou/Baiyun	45	329
20	**Las Vegas/McCarran**	**41**	**506**

Source: Airports Council International, 2012.

TABLE 1.3 Busiest Airports in the World in 2011 (ranked by number of passengers)

The United States has been a leader in the development of mass air transport. As of 2011, people in the United States on average took 2.3 trips by air every year. This rate was about triple that of Europe and 10 times that in the rest of the world. Historically, average fares in the United States were considerably less expensive than elsewhere.

The air transport industry in the United States faced the challenges of high volumes of traffic well ahead of the rest of the world. It has correspondingly led in the development of major innovations that continue to transform, commercial aviation and airport planning and design worldwide. Table 1.4 indicates some of them. These innovations, together with the trends discussed in the following sections, have been radically changing the concept of airport systems planning and design. Indeed, airport systems planning and design in the United States has differed significantly from that in the rest of the world. Therefore, to the extent that countries follow American examples, they will be introducing significant changes.

Innovation	Implications and Effects
Economic deregulation	Airlines can fly where they want and charge any fare. Spread to Canada, Australia, European Union, etc.
"Southwest" model of low-cost airline	Copied worldwide (Westjet in Canada, Ryanair in Europe, AirAsia in Asia, etc.)
U.S. "open-skies" policy	Deregulation of airline routes between the United States and over 100 other countries
Airline alliances	Coordination of flights, frequent flyer benefits, etc. (Oneworld, SkyTeam, Star Alliance)
Integrated air cargo services	Simplification of small cargo service, strong impulse on e-tail (FedEx, UPS, etc.)
Transfer hubs	Airline efficiency, higher flight frequency (Atlanta, Chicago/O'Hare, Dallas/Ft. Worth, Denver, etc.)
Midfield concourses	Rapid, efficient transfer of connecting passengers (airports above, also in England, Spain, Malaysia, etc.)
Automated people movers	Wide use at transfer hubs (Tampa, midfield concourses, between terminals)
Global positioning system	Reduces needs for ground navigation facilities, enables more direct and economical airline routings

TABLE 1.4 Organizational Innovations in Air Transport from the United States

Airlines in the United States have always been private. Elsewhere, however, governments usually owned and operated airlines and airports. It was only around the 1990s that Britain, the Netherlands, Germany, and Japan began to privatize their airlines, setting off a worldwide trend.

Airports in the United States generally operate in an implicit public-private partnership. Public entities own the land and are responsible for the runways and other airside facilities. Private companies design, build, and operate much of terminals, hangers, and other landside facilities. Most important, private sources provide much of the money for airport infrastructure.[1] Airports in the United States have therefore traditionally paid close attention to the returns on investments and ways to make the facilities pay. In this, the United States contrasts with other countries whose airports were almost all owned, designed, financed, built, and operated by government employees until the trend toward airport privatization began in the 1990s.

The preceding means that the context, objectives, and criteria of excellence for airport planning, management, and design are fundamentally changing. Rapid changes in the industry require strategic thinking and the flexibility to adapt to new circumstances. Increased commercialization and privatization of airports calls for an appreciation of the economic and financial aspects of airport operation. Narrow technical excellence is not sufficient to deliver good value for money for airports. Airport professionals need to create dynamic, strategic plans that incorporate flexible designs and enable airport operators to manage their risks.

The current environment for airport planning and design requires a systems approach. This contrasts with traditional airport engineering that has tended to focus narrowly on technical matters to the exclusion of issues such as costs and revenues, volatile traffic and risks, and operations and management. Government and international agencies have set fixed design standards that did not allow tradeoffs between cost and service. Textbooks followed the same vein. (See, e.g., FAA, 1988; IATA, 2004; ICAO, 1987; Horonjeff et al., 2010; Ashford et al., 2011.) Comprehensive systems planning and design has not been the norm.

In response to current needs, this text broadly considers the range of factors that shape the performance of the airport. It expands the concept of airport planning and design to include operations and long-term management through technical and economical measures. Correspondingly, it uses a wider range of tools for analyzing preferable solutions, as Chaps. 18 through 21 indicate. This systems approach should be broadly useful to all professionals actively associated with airports.

1.2 Long-Term Growth

Aviation passenger and cargo traffic grew remarkably over the last generation. Over good and bad years, passenger traffic worldwide increased at an average of about 4 percent from 1990 to 2010. This meant that air travel more than doubled during that time.

Growth rates differ significantly by geographic areas. In the United States, the growth rate in the number of enplanements from 2000 to 2010 fell to about 1 percent a year, whereas in the rest of the world this traffic grew about 150 percent over the decade. See Fig. 1.1.

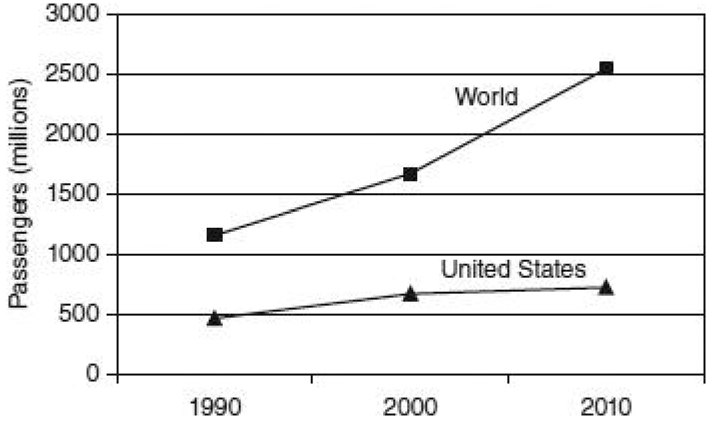


FIGURE 1.1 Growth in airline traffic worldwide. (*Sources*: International Air Transport Association, ICAO.)

Cheaper, safer air service has propelled the growth in aviation traffic. Most obviously, the real price of air travel has persistently fallen over the last decades. A steady rise in demand mirrored this long-term drop in prices, as basic economics expounds and Fig. 1.2 confirms. Meanwhile, flight safety has improved dramatically, as Fig. 1.3 shows. Passengers and cargo now enjoy cheaper, safer travel than they did a generation ago.

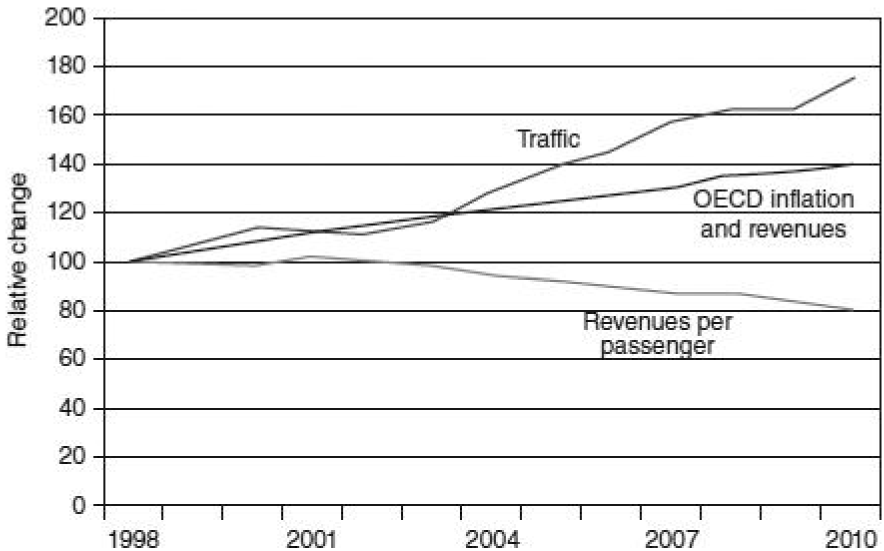


FIGURE 1.2 Traffic worldwide has grown rapidly as costs of airline traffic decreased. (*Source*: IATA World Air Transport Statistics.)

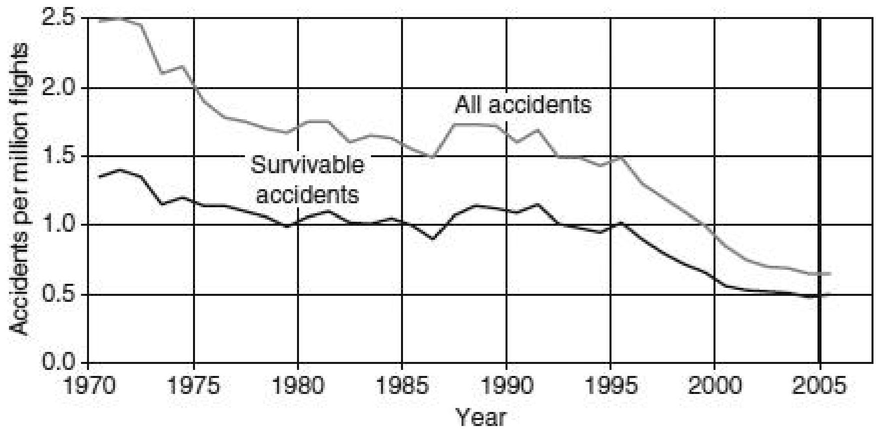


FIGURE 1.3 Long-term accident rates have dropped significantly. (*Source*: FAA, 2010.)

Forecasts of future traffic are questionable. Small differences in assumptions cumulate to enormous differences in consequences 25 years or more from now. Airport professionals should thus be tentative about traffic predictions. For example, slight deviations of plus or minus 1 percent from a long-term annual growth rate lead to substantially different forecasts. A 5 percent annual rate of growth compounded over 25 years gives a result about 140 percent greater, in terms of the starting amount, than a 3 percent annual rate of growth. Managers should place any estimate of long-range forecasts in a broad range of possibilities. Chapter 4 discusses this issue in detail.

Traffic will almost certainly continue to grow substantially. Most of the world rarely flies, and the market is far from saturated. Plausible increases in population and national wealth, and the tendency of members of younger generations to fly, will lead to more traffic. Increased globalization also impels long-distance travel for business and personal reasons, in general only realistically feasible by air. Even a modest growth rate of 3 percent a year doubles traffic in 25 years.

No one can count on steady growth, however. Trends may slow down or stop. Over the last decades, a series of major causes steadily reduced operating costs that in turn lowered airfares and drove the historical rise in air traffic. These were the following:

- Larger, more efficient aircraft
- Economic deregulation of the airlines, accompanied by competition
- Worldwide privatization of aviation and increased attention to costs

- The consequent competitive restraint on wages
- The introduction by airlines of differential pricing and revenue management systems that raised load factors
- Historically low fuel prices (when adjusted for inflation)

Some trends may reverse. For example, fuel prices might rise considerably. Business travel may give way to inexpensive video conferencing or other communications. Concerns about security may limit travel, particularly for short-haul flights. Persistent economic difficulties might stunt traffic growth. The likely scenario, however, is that aviation will register substantial overall increases. When applied to the existing market, even lower growth rates will lead to substantial growth.

Overall, it is reasonable to assume that by 2040 the level of passenger traffic could be up to two or even three times higher than that in 2012. For example, the number of enplaned passengers in the United States could be in the range of 1000 to 2000 million a year, compared to the 720 million a year flying in 2010. Airport planners should thus prepare for the possibility of substantial expansion. However, because the growth is speculative, they should not commit to building facilities until they can confirm this growth. In short, they need to manage their risks consciously, as Chaps. 4 and 19 indicate.

The composition of the total traffic may differ significantly from what it has been. Over time, air travel has diffused from the rich to the masses, from the early to the later developed nations. It has shifted from being a luxury good for the elite, to a necessary business need, to mass transportation, to international tourism. Airport planners may anticipate an extension of such patterns, both domestically in their own markets and internationally, from North America, Europe, and Japan to the rest of the world. Asia, the Middle East, and even Africa are sure to be increasingly important for the airport/airline industry.

Cargo traffic may continue to expand dramatically as companies reorganize their distribution systems around electronic commerce. As of 2012, in a development not widely perceived, the integrated package carriers UPS and FedEx were already among the largest airlines in the world in terms of aircraft operated (see Tables 1.2 and 1.6). To the extent that businesses continue to substitute web sites for brick-and-mortar stores, and to deliver products directly to customers rather than through local warehouses and in-store inventories, the integrated cargo carriers may grow rapidly. This traffic may be a driving force for many future airport developments.

1.3 Organizational Change

The organization of the airport/airline industry has been fundamentally changing over the last generation—and the process continues. Economic deregulation creates opportunities for low-cost and integrated cargo airlines to grow, impels governments to privatize their airlines, and leads traditional airlines to consolidate. Political deregulation, such as open-skies agreements, enables new markets, changes traffic patterns, and increases competition. This evolution greatly affects airport systems planning and design.

The U.S. economic deregulation of the airlines in 1978 catalyzed these organizational changes. Deregulation allowed U.S. domestic airlines to establish and drop routes as they please, charge whatever prices they wish—and do so at a moment's notice, without having to ask for government permission. This event led to rapid innovation in services, big increases in productivity, and significant fare drops. The example proved contagious, and similar deregulation of air travel has spread to major international markets, notably Australia, Canada, the European Union, Japan, and India. More recently, the United States has effectively been promoting political deregulation of the airlines through its "open-skies" agreements with other countries. These treaties eliminate governmental restrictions on airline destinations, frequencies, and fares, and allow wide access to each other's markets (except for domestic flights, known legally as *cabotage*). The result is that much of the air transport industry now operates in a context completely different from the one that prevailed until the late 1990s.

Low-Cost and Integrated Cargo Airlines

Deregulation enabled new low-cost passenger and integrated cargo airlines to flourish. They have become major and, in some markets, dominant airlines. Their innovative modes of operation are correspondingly changing airport design and operations.

Southwest represents the salient success of low-cost airlines. It became the leading carrier of domestic passengers in the United States (Table 1.5). It has been a role model for comparable low-cost carriers in other markets: WestJet in Canada, Ryanair and easyJet in Europe, and AirAsia in Southeast Asia.

Airline	Domestic Market Share (%)
Southwest-Air Tran	20.6
Delta	14.2
United-Continental	11.1
American	10.2

Source: US Bureau of Transportation Statistics, Schedule T1, 2011.

TABLE 1.5 Southwest Airlines Was the Dominant U.S. Domestic Carrier in 2011 (in terms of passengers carried)

Southwest established the standard for low-cost operations in many ways. It uses a standard fleet of aircraft to drive down training and maintenance costs and has flexible work rules that use personnel efficiently to do many tasks. To maximize the utilization of its fleet, it has historically tried to use uncongested airports with minimum delays, and to turn around aircraft in as little as 20 minutes at the gate. These operating policies directly affect airport planning. The low-cost airlines' push toward uncongested airports has favored the development of secondary airports in metropolitan areas, such as Dallas/Love Field, Miami/Fort Lauderdale, and London/Stansted (see Chap. 5). Their emphasis on quick turnaround times reduces the need for gates and terminal space. Moreover, low-cost airlines have led the way for the development of low-cost terminals internationally, as at Paris/de Gaulle and Singapore.

FedEx similarly is the prototype of integrated cargo airlines that provide door-to-door service between suppliers and customers. It integrates its fleet of aircraft with huge fleets of ground vehicles using highly automated, standardized facilities and advanced IT technology throughout. It provides a remarkably efficient service that is reconfiguring the distribution of goods both for manufacturers and consumers. Distributors increasing substitute the integrated cargo service for local warehouses or retail stores.

FedEx and UPS have grown to dominate the market for airfreight. In 2010, they carried two to four times more tons than nearest competitors (Table 1.6). Both are among the largest airlines in the world, in terms of number of aircraft. Moreover, they have been highly profitable and correspondingly have the capacity to finance the kind of airport facilities they need.

Airline	Tons Carried (thousands)
FedEx	6949
UPS	4509
Korean Air	1805
Emirates	1777
United-Continental	1760
Cathay Pacific	1579
China Airlines	1374

Source: IATA World Air Transport Statistics, http://www.iata.org/ps/index_products.asp.

Table 1.6 FedEx and UPS Dominated Their Competitors in 2010 (in terms of tons carried)

FedEx and UPS have also been responsible for the development of major cargo hub airports such as Memphis and Louisville in the United States, and numerous major distribution centers such as Los Angeles/Ontario, Guangzhou, and Paris/de Gaulle. They are responsible for many "cargo airports" insofar as cargo is the major component of the activity at the airport. In the age of integrated cargo carriers, cargo is no longer a peripheral activity secondary to passenger traffic; it can be a primary driver of airport development.

Privatization

Globally, governments are getting out of the aviation business. They are privatizing airlines and airports. Business management in a market economy is replacing government ownership in a regulated environment. This trend further changes airport planning, design, and management.

Except in the United States, the standard practice for most of the twentieth century was that government bodies owned and operated both airports and airlines. Whereas in the United States the airlines were private and local authorities normally ran the airports, almost everywhere else national ministries or their dependencies ran both airlines and airports. The airport/airline industry thus benefited from public subsidies and protection. Correspondingly, it operated in a highly regulated, political context, in which political interests often dominated economic or commercial rationales. This era, and the design and management mentalities that go with it, are fast disappearing. Worldwide, the airport/airline industry is converging toward standard American practice: airlines are private and airports operate under some form of public–private partnership.

Governments have been privatizing their national airlines (Table 1.7). Airlines therefore increasingly follow economic self-interest. They are finding it imperative to drop unprofitable routes and streamline their operations. With the understanding that the business involves economies of scale and scope, national airlines are also merging and disappearing (Table 1.8). These reorganizations affect their airports, most obviously by reducing operations at the hubs of the closed airlines.

Airline	Status
Air Canada	Completely
Air France	Partial
Alitalia	Completely
Australian	Completely
British Airways	Completely
Japan Airlines	Completely
Lan Chile	Completely
Lufthansa	Completely
Qantas	Completely
SAS	50% private

TABLE 1.7 Examples of Privatized Airlines

Airline	Status
Iberia	Merged with British
KLM	Merged with Air France
Sabena	Bankrupt, gone
Swiss	Merged with Lufthansa

TABLE 1.8 Examples of Consolidated National Airlines

Governments have likewise been privatizing airports. They do this either by creating some forms of public–private partnership similar to those prevailing in the United States,[2] or by creating independent companies that they then regulate as local monopolies (Table 1.9). In this environment, cost and economic performance are increasingly crucial criteria for good design, and they are radically changing the timing and nature of what airports decide to build.

Airports	Status
Argentina/major airports	Private company
Australia/major airports	50-year leases
Austria/Vienna	Private company
Denmark/Copenhagen	Private company
Germany/Frankfurt	Private company
Greece/Athens	Public/private company
India/Delhi	30-year lease
India/Bengalaru	Private company
Italy/Rome and Milan	Private companies
Malaysia	Public/private company
Mexico/regions	50-year leases
S. Africa/Johannesburg	Partially privatized
Switzerland/Zurich	Private company
U.K./Heathrow	Private company

Table 1.9 Examples of Privatized Airports

Globalization

Political deregulation is also opening up more airports to international, intercontinental airline services. "Open-skies" agreements permit airlines from each country to serve any destination in each other's country and are becoming increasingly widespread (Table 1.10). They permit, for example, American Airlines to serve Barcelona from any U.S. airport. Such freedom contrasts with previous conventions that limited foreign airlines to a few gateway airports in each country. These agreements enable more convenient nonstop intercontinental service for secondary cities in different countries, as Chap. 2 describes. This has an impact on airports: it favors the use of smaller long-distance aircraft such as the Boeing 787 launched in 2011 (and reduces the need for the much larger Airbus A-380).

Between United States and	As of
Belgium	March 1995
France	October 2001
United Kingdom	April 2007
Chile	October 1997
Uruguay	October 2004
Malaysia	June 1997
Indonesia	July 2004

TABLE 1.10 Examples of Open-Skies Agreements as of 2012

In parallel, airlines have formed three global alliances. These groupings enable the airlines to coordinate their schedules and practices and thus provide services that are more convenient to customers. They have become a notable force in the airport/airline industry, as they account for about half the passenger traffic worldwide (Table 1.11). They present a challenge to airports managers; they demand common airport locations and services, and do so with great bargaining power.

Feature	Alliance		
	Star	Oneworld	SkyTeam
Major Airlines	United, Air Canada, Air China, ANA, Air New Zealand, LOT, Lufthansa, SAS, Swiss, Singapore, South African Airways, TAP, Thai, Turkish, US Airways	American, British/Iberia, Cathay, Japan, Lan, Qantas, Qatar	Air France/KLM, Alitalia, Delta, China Airlines, China Eastern, China Southern, Etihad, Korean, Saudi, Aeroflot, Aeromexico
Smaller Airlines	Asiana, Austrian, BMI, Brussels, Egyptair Adria, Croatia, Spanair, EVA, Shenzen, Copa, AviancaTaca	Finnair, Malev, S7, Royal Jordanian	Air Europa, Czech, Kenya, Middle East Airways, TAROM, Vietnam, Xiamen
Market Share %	30	19	19

TABLE 1.11 Global Airline Alliances in 2012

To a lesser extent, major airport companies have partnered with lesser airports worldwide to provide a range of management services (Table 1.12). These arrangements have brought leading practices to less developed facilities. So far, such arrangements are relatively insignificant overall.

Lead Airport Group	Arrangements with
Abertis (Spain)	Airports in the United Kingdom, Sweden, Mexico, United States, Colombia, Bolivia, and Chile
Ferrovial (Spain)	BAA (U.K. airports)
Frankfurt/International	Athens, Antalya (Turkey), Lima (Peru)
GMR group (India)	Delhi, Hyderabad, Istanbul/Sabiha (Turkey), Malé (Maldives)
Paris	26 Airports worldwide: North Central Mexico, Egypt, Cambodia, Liege (Belgium), Algiers, Amman (Jordan), Jeddah (Saudi Arabia), Mauritius, Conakry (Guinea)
Airports Company South Africa	10 airports in South Africa
Vantage (Vancouver Airport)	Operating agreements in South America (Bahamas, Chile, Dominican Republic, Jamaica), Cyprus, and in the United Kingdom (Doncaster, Liverpool, and Durham Tees Valley)
Zurich	Bengalaru

TABLE 1.12 Examples of International Airport Consortia as of 2012

1.4 Technological Change

The information age is leading to major revisions in the concept of airport facilities. Other developments, such as satellite-based global positioning systems (GPS) and the large Airbus A-380, will continue to affect airports by increasing capacity and demands, but they do not imply major conceptual revisions. In addition, the steady introduction of people movers into airport designs will continue to enable the use of midfield and remote terminals that reduce the great costs and delays of aircraft movements on the ground (see Chap. 14).

Information technology is affecting airports principally in two ways:

1. Electronic processing of passengers and bags, both before the passenger arrives at the airport and after, which affects the design of terminals

2. E-tail, the phenomenon of Internet ordering of products, which is fueling the impressive growth of integrated cargo airlines such as FedEx and UPS, as already discussed

Electronic processing of passengers speeds up the process, reduces queues, and thus greatly lessens the need for large spaces for checking-in passengers or for moving them through border controls. As of 2012, the possibility of printing boarding passes away from the airport and the use of check-in kiosks at the airport had already greatly reduced the number of airline agents and counters required for check-in (Fig. 1.4). As the industry moves toward the use of "boarding passes" on personal mobile devices, further steps and delays will drop out of the process. Similarly, the electronic clearance of passengers through border controls will speed up the processing of many travelers, reducing queues and the need for cavernous arrival halls. Already, the initial uses of Global Entry in the United States, resident cards in Singapore, and Privium in the Netherlands provide attractive, speedy service to travelers. As electronic processing speeds up service, airports can anticipate the possibility that they will need much less space per person for ticketing and check-in than they now do (see Chap. 16).

FIGURE 1.4 Kiosks reduce check-in space for Lufthansa counter at Berlin/Tegel.

Similarly, airports should anticipate the potential for electronic identification of bags to expedite the processing of bags. Continuing work on radio-frequency identification devices (RFIDs) is promising and has potential.

1.5 Implications for Airports Systems Planning and Design

Taken together, the trends in the airport/airline industry are substantially changing the context, objectives, and criteria of excellence for airport planning and design. Airport professionals now need more than narrow technical skills. They must be responsive to a range of economic and management issues.

The context is increasingly commercial and economic. Planners and designers are no longer designing primarily for administrators according to standard norms. They must respond to a broad range of business interests, such as the airlines, the airport operators, and concessionaires of all sorts. Through these immediate clients, they will have to cater to their customers. This means that airport planners and designers will have to think in terms of profitability, revenues, and service to users.

The objectives consequently focus more on performance than on monuments. Airports will build more low-cost, efficient terminals. Value for money, good service, and functionality will become dominant considerations. Architectural significance and grand visions will be important but may become secondary considerations. In general, airport planning and design will become more democratic, more in tune with everyday needs, and less directive or technocratic.

The criteria of excellence will correspondingly focus on cost-effectiveness, value for money, efficiency both technical and economic, and profitability. Airport planners and designers will have to factor these considerations into the purely technical analyses of traditional airport engineering. This requires skills not usually part of engineering or architectural training. It calls for an understanding of economic and financial analyses. It extends beyond construction to operations and the management of risk. In short, it calls for a systems perspective.

A systems approach will be the basis for proper future planning and design of airports. Airport professionals will recognize that they will have to consider technical, economic, and social issues jointly as part of a larger system evolving over time to meet varying loads and demands. This text presents the essential elements of how to do this.

Exercises

1.1. Select an airport or region for a case study. Obtain data on growth of airport traffic and airline operations. What are the trends over the last 10 years? How do they compare with international or regional trends? Discuss how future traffic might evolve for your case.

1.2. Select an airport and find out how the continuing reorganization of the airline industry has affected operations over the previous 10 years? Have airline clients changed? Have they needed new facilities or required relocation? Describe the overall evolution.

1.3. Estimate the growth rate for integrated cargo carriers by comparing current statistics with those in Table 1.6. Use the web to obtain company reports on major carriers to document recent interesting developments. Explain and discuss your view on how you see this activity developing in your region.

References

ACI, Airports Council International (2011) "World Airport Traffic Report for 2010," Montreal, Canada.

Ashford, N., Mumayiz, S. A., and Wright, P. (2011) *Airport Engineering: Planning, Design and Development of 21st Century Airports*, 4th ed., Wiley, New York.

FAA, Federal Aviation Administration (1988) *Planning and Design Guidelines for Airport Terminal Facilities*, Advisory Circular 150/5360-13, U.S. Government Printing Office, Washington, DC.

FAA, Federal Aviation Administration (2010) *Trends in Accidents and Fatalities in Large Transport Aircraft*, DOT/FAA/AR-10/16, National Technical Information Service.

Horonjeff, R., McKelvey, F. X., Sproule, W., and Young, S. (2010) *Planning and Design of Airports*, 5th ed., McGraw-Hill, New York.

IATA, International Air Transport Association (2004) *Airport Development Reference Manual*, 9th ed., IATA, Montreal, Canada.

ICAO, International Civil Aviation Organization (1987) *Airport Planning Manual, Part 1: Master Planning*, 2d ed., Doc 9184, ICAO, Montreal, Canada.

[1]Major airports in the United States raise capital to build passenger buildings, hangars, garages, and the like through bonds offered to private investors or through fees charged to passengers [the Passenger Facility Charge (PFC)]. The U.S. government, through its Federal Aviation Administration, pays a share of the cost for runways, air traffic control, and safety measures. The government contributions are most significant at smaller airports but less important at established major airports.

[2]At most airports in the United States, a majority of the management functions—such as finance, design, construction, and much of management—are provided by private companies, although public agencies own the land. See Chap. 3.

CHAPTER 2

The Evolving Airline Industry: Impacts on Airports

For decades, government regulation of airline fares and services in many cases constrained the growth of passenger traffic. The U.S. Airline Deregulation Act of 1978 represented a major turning point for commercial air transportation, and this kind of relaxation or elimination of economic regulation of airline markets has since spread to most regions of the world. This liberalization allowed increased competition and has led to a dramatic transformation of airline characteristics, as well as their operating and commercial practices.

Most of the changes due to deregulation started in the United States and spread rapidly throughout most regions of the world. With increased competition, air travelers have seen dramatically lower airfares (in real terms) as well as changes to route networks and service quality. The removal of barriers to entry allowed innovative new entrant airlines with lower cost structures to offer consumers new options for air travel at lower prices. At the same time, established airlines have experienced increased profit volatility and, in some cases, bankruptcy and liquidation.

Increased competition has driven fundamental changes in airline fleets, routes, schedules, and operations, all affecting basic airline economics, operating costs, and productivity. These changes affect many different facets of airport operations. An understanding of how these recent changes and the expected future evolution of airlines can affect airports is essential for airport systems planners.

This chapter summarizes the most important trends in airline planning and business practices that have emerged with increased liberalization, and discusses their implications for airports. It examines the trends in fleet composition, network structure, and scheduling that can have a direct impact on airport planners and operators. It also discusses airline operational variability and its effects on operational requirements at airports. It then relates these changes in airline business practices to reductions in airline operating costs and improved productivity. The chapter concludes with a summary of the most important airline industry trends that will continue to affect airports in the future.

2.1 Trends in Airline Fleets

An airline's fleet is described by the total number of aircraft and the specific types of aircraft that it operates. Each aircraft type has different technical and performance characteristics, most commonly defined by its *range* and *size*. The "range" of an aircraft is the maximum distance it can fly without stopping for additional fuel, while still carrying a reasonable payload of passengers and/or cargo. The "size" of an aircraft can be represented by its seating or cargo capacity, as indicators of the amount of payload that it can carry. Other important technical and performance characteristics of each aircraft type include a variety of factors related to both airline operational and airport physical constraints. For example, each type has maximum takeoff and landing weights that determine minimum runway length requirements and, in turn, the feasible airports for operating the aircraft. Similarly, limitations on taxiways and gate space and even ground equipment at different airports can impose constraints on the airline's choice of aircraft type.

Published prices for a narrow-body 150-seat aircraft that is typically used for short- to medium-haul flights range from U.S. $60 to 80 million. The list price of the largest long-range wide-body aircraft, the Airbus A380 that can seat up to 600 passengers, is over U.S. $350 million (Airbus, 2012). However, airlines typically pay significantly less than the published list prices because of intense sales competition and price discounting by the aircraft manufacturers.

The fleet planning process requires airlines to make long-term strategic decisions that will affect their network structures and ability to operate specific routes for many years, even decades. These investments in aircraft can affect airline balance sheets for 10 to 15 years through depreciation costs as well as long-term debt and interest expenses. The decision to acquire specific aircraft types can have an even longer impact on an airline's operations, as some commercial aircraft more than 30 years old are still in use today.

Environmental concerns and regulations are having a growing impact on airline fleet decisions. The noise impact of commercial jet aircraft is a major issue for airports and the communities that surround them. Many airports now have regulations and/or curfews that limit or prevent the operation of older aircraft types with engines that exceed specified noise levels (see Chap. 6). Similarly, there is a growing trend toward imposition of air pollution regulations designed to limit aircraft emissions around airports. At the start of 2012, the European Union imposed an "emissions trading scheme" (ETS) intended to limit the carbon emissions of airlines operating into and out of European airports. These environmental regulations provide further incentives to airlines to update their aging fleets with newer-technology aircraft that are both quieter and cleaner in terms of emissions, but which have substantially higher ownership costs.

As a general rule, the largest aircraft types operate on routes with the longest flight distances. This relationship has less to do with technical or performance issues (such as fuel capacity) than with the realities of airline frequency competition. All else equal, larger aircraft have lower operating costs per mile and per seat for any given flight distance. Irrespective of distance, it would make economic sense for airlines to operate fewer frequencies with larger aircraft to increase passenger loads on each flight and reduce costs—both total operating costs (due to fewer flights) and unit costs per seat (due to fixed costs being spread over more seats per flight). On competitive routes, however, frequency share is the primary determinant of airline market share, particularly on short-haul routes where more flights improve the convenience of air travel relative to other modes. Frequency share is especially important in the competition for time-sensitive business travelers who pay higher fares than leisure travelers.

Figures 2.1 and 2.2 summarize the typical seating capacity and range characteristics of different commercial jet aircraft types available to airlines in 2012. The presentation distinguishes between single-aisle or "narrow-body" aircraft with approximately 200 seats and fewer (see Fig. 2.1) and two-aisle "wide-body" aircraft typically with more than 200 seats (see Fig. 2.2). The positive relationship between aircraft size and range is apparent in both figures, although the strength of this relationship has weakened significantly over the past several decades. The principal aircraft manufacturers have substantially increased the number of aircraft types available, giving airlines a greater choice of aircraft with different range and capacity combinations.

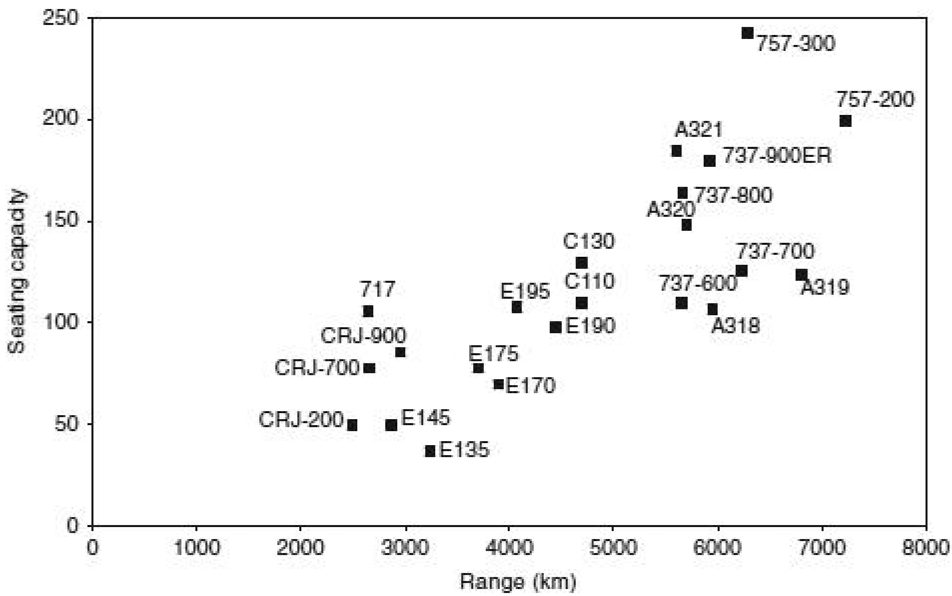


Figure 2.1 Narrow-body commercial jet aircraft. (*Sources*: Manufacturer web sites—www.airbus.com, www.boeing.com, www.embraer.com, www.bombardier.com.)

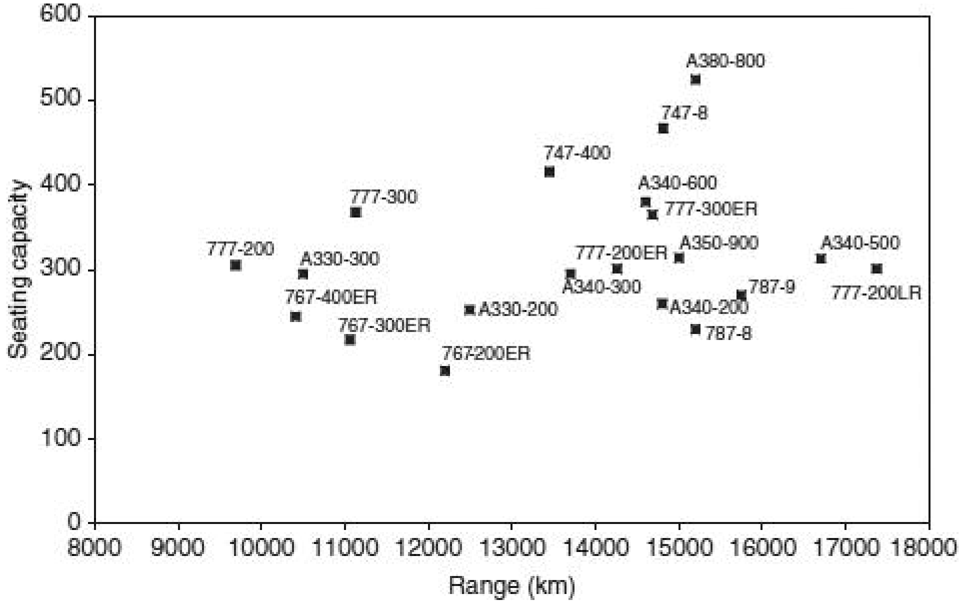


FIGURE 2.2 Wide-body commercial jet aircraft. (*Sources*: Manufacturer web sites—www.airbus.com, www.boeing.com, www.embraer.com, www.bombardier.com.)

The smallest narrow-body passenger jet aircraft shown in Fig. 2.1 include the 35- to 50-seat "regional jets" developed by Bombardier and Embraer in the 1990s. At the upper end of the narrow-body spectrum are Boeing and Airbus products with 170 to 200 seats and a maximum range of 6000 to 7000 km.

A notable trend is the general increase in the range capabilities of relatively small aircraft. Several aircraft types with 120 to 130 seats can operate nonstop flights over 6000 km (e.g., B737-700 and A319). These smaller aircraft can serve transcontinental routes in North America as well as medium-haul international routes such as Amsterdam-Amman, for example. The development of smaller aircraft with longer ranges enables airlines to provide nonstop flights on routes with relatively low demands. It also allows them to increase the frequency of flights on competitive medium-haul routes that previously were limited to much larger aircraft types.

Small regional jets with 35 to 50 seats were introduced in the mid-1990s and their use grew rapidly, especially in North America and Europe. These small jets allowed airlines to offer the speeds and passenger comfort of much larger jet aircraft on short-haul routes, in many cases replacing slower and noisier turboprop aircraft. They also enabled airlines to offer more frequent departures on competitive short- to medium-distance routes. Perhaps the most important driver of the success of these regional jets, however, was their appeal to large U.S. and European airlines with unionized pilots. Pilot union contract "scope clauses" required airlines to employ well-paid unionized pilots for any jet aircraft with over a certain number of seats, typically 70. With the development of 35- to 50-seat regional jets, airlines were able to hire lower-paid pilots to fly these smaller aircraft.

These impacts of small regional jets have been most apparent in U.S. domestic operations. As Fig. 2.3 shows, regional jets were introduced in 1997, and the number operated by U.S. carriers on domestic flights grew to 1500 by 2006. The vast majority of these aircraft were EMB135, EMB145, and CRJ-100 and CRJ-200 aircraft, all with 50 seats or fewer. Contrary to conventional wisdom, these regional jets were used by hub airlines and their commuter partners primarily to increase the frequency of service from the hub to small spoke cities, *not* to over-fly the hubs with new nonstop services (Mozdzanowska, 2004).

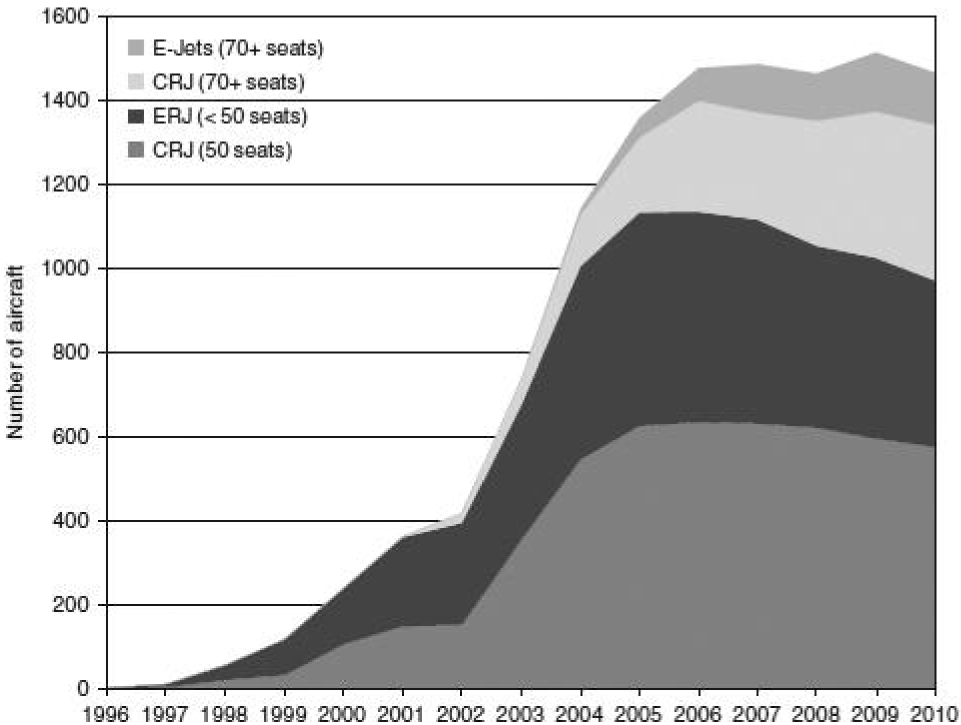


FIGURE 2.3 Regional jets operated by U.S. airlines. (*Courtesy*: A. Wulz, MIT. *Source*: US DOT Form 41.)

The growth of the small regional jet fleet has slowed since 2000, as surging fuel prices began to put the economics of 50-seat regional jets into question. Many of the pilot contracts that allowed airlines to fly smaller regional jets with nonunion pilots also came due for renegotiation, further reducing their economic appeal to airlines. Figure 2.3 shows that, after 2005, the growth of U.S. domestic regional jet operations slowed dramatically and that the growth that occurred was limited to 70- and 90-seat regional jets with lower unit costs per seat. With even more dramatic increases in fuel prices starting in 2008, some airlines replaced 50-seat regional jets with newer 70-seat turboprop aircraft that consume less fuel per seat-kilometer.

These trends led to the more recent emergence of a new set of aircraft. As the airlines shifted away from small regional jets, Embraer led the development of a new category of aircraft with capacity and range characteristics in between the early regional jets and the larger narrow-body offerings of Boeing and Airbus. Shown on Fig. 2.1, the Embraer 170/175/190/195 series filled a previous gap in terms of both seats (75–100) and range (~ 4000 km). Bombardier also announced plans for its "C-series" aircraft, with slightly higher capacities and increased range capabilities.

The capacity and range characteristics of large wide-body jet aircraft are plotted in Fig. 2.2. On this graph, the positive correlation between seating capacity and maximum range is much less apparent than in the case of narrow bodies. The capacities of many of the new long-range aircraft have decreased over the past decades, allowing airlines to serve relatively low-demand long-haul international routes nonstop. In addition, a medium-size/medium-range category of new aircraft types has emerged, as airlines find new "missions" for aircraft with intermediate combinations of range and capacity. Both the Boeing 787 and the Airbus 350 are new aircraft types that provide excellent examples of this trend.

First delivered in late 2011, the 787 is a relatively small (230-seat) aircraft with a very long range of over 15,000 km. Some have referred to the 787 as a "game changer" for airlines hoping to expand their networks by adding routes previously thought not to be sustainable given low demand and/or not feasible given their long distances. Two of the earliest routes for the 787 provide examples of how such an airplane will be used—Japan Air Lines started the first nonstop flights between Boston and Tokyo in 2012, while United has announced plans for the first non-stop service between Denver and Tokyo.

Note that, in both cases, the airlines use the 787 to add service from their existing hubs to new destinations rather than providing "point-to-point" nonstop services. Although not strictly "point-to-point," these new flights could well divert traffic from established airline hubs—the Boston-Tokyo flight will carry passengers that previously connected via Chicago, for example, while the Denver-Tokyo service will affect the volume of traffic connecting at San Francisco. As large airlines continue to reinforce their own hubs with more nonstop services to smaller connecting spoke cities, passengers can bypass other existing hubs. This is of particular concern to European network carriers who see the buildup of large connecting hubs in the Middle East as a threat to the traffic at their European hubs.

An important exception to the general trend of smaller wide-body aircraft is the Airbus A380 aircraft with 500 to 600 seats and a 15,000-km maximum range. This aircraft has been in service since 2009 and is operated by over half a dozen international airlines on long-haul routes where demand is high and frequency competition is not a major factor. For example, Air France replaced two smaller wide-body flights with one daily A380 flight between Paris and Montreal, reducing its unit costs on the route with little risk of losing market share. Other A380 operators have also assigned the aircraft to the heaviest routes into their connecting hubs—Frankfurt-New York for Lufthansa and Singapore-London for Singapore Airlines are two examples.

The largest operator of this largest wide-body aircraft type, Emirates, provides another case study of how airlines will use the A380 and how it might ultimately change global airline competition. In 2012, Emirates operated over 20 A380 aircraft and had about another 80 on order. The airline is based at its single connecting hub in Dubai, and virtually all of its flights operate to and from this airport. As Dubai has relatively small local demand for travel, Emirates depends heavily on connecting passengers that neither originate nor terminate their trips in Dubai. Emirates thus uses the A380 and other wide-body aircraft to carry mostly connecting traffic into and out of Dubai. For example, an industry report indicated that only 10 percent of the average passenger load on an Emirates A380 flight from Toronto to Dubai is actually destined to Dubai with most passengers connecting to dozens of destinations beyond Dubai. As another example, Emirates in 2012 operated daily A380 flights between Manchester, England, and Dubai, a nonstop route that few would have predicted could support such a large aircraft.

With the increased diversity of available commercial aircraft, airlines in different regions of the world have adopted different fleet and network strategies reflected in the average size of their aircraft. As Fig. 2.4 shows, the global average size of commercial jet aircraft is 136 seats, but it varies substantially among airlines from different parts of the world. The emphasis of Middle East and Far East airlines on the operation of long-haul services with the largest wide-body aircraft gives them substantially larger average aircraft sizes, at 199 and 172 seats, respectively. On the other hand, U.S. airlines have an average aircraft size that is 40 percent smaller, at 119 seats, as many short- to medium-haul domestic routes depend heavily on frequency competition for market share. The average aircraft size is even smaller in regions such as Central America and Canada, where both frequency competition and lower levels of demand for air travel lead to the use of smaller aircraft.

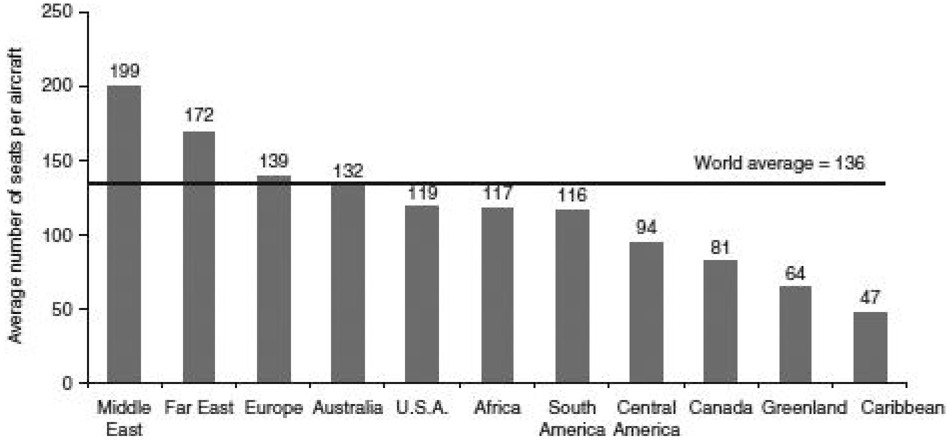


FIGURE 2.4 Average number of seats per aircraft. (*Courtesy*: K. Shetty, MIT. *Source*: Official Airline Guide, October 2010.)

Looking ahead, the future fleet composition of the world's airlines will reflect the trends described previously. Airlines will continue to use small new-generation narrow-body aircraft to provide increased frequency of flights on competitive short-haul routes, and to operate a variety of wide-body aircraft types of different sizes to expand airline networks primarily through further growth of existing connecting hubs. There has been little evidence to date of a widespread shift to nonstop point-to-point services, with the exception of some new entrant low-cost carriers (LCCs).

Figure 2.5 illustrates the worldwide large jet aircraft (excluding regional jets) order backlog at the end of 2011, categorized both by aircraft type and world region. Asia-Pacific airlines have the most aircraft on order, a result of more rapid air travel demand growth in that region as well as a quicker recovery from the effects of economic recession in 2008–2009. Contributing to the number of these orders is the continued rapid growth of LCCs such as AirAsia, airlines that still see tremendous untapped potential for low-fare air travel in the region. North American and European airlines rank second and third, respectively, in terms of total aircraft orders despite their relatively larger size of fleets and networks. A decade of poor profitability, exacerbated by much deeper impacts of recession and fuel prices have kept these more established airline groups from renewing and expanding their fleets as quickly. Worth noting is the volume of aircraft on order by Middle East airlines—Emirates, Etihad, Qatar, and others all have very aggressive growth plans.

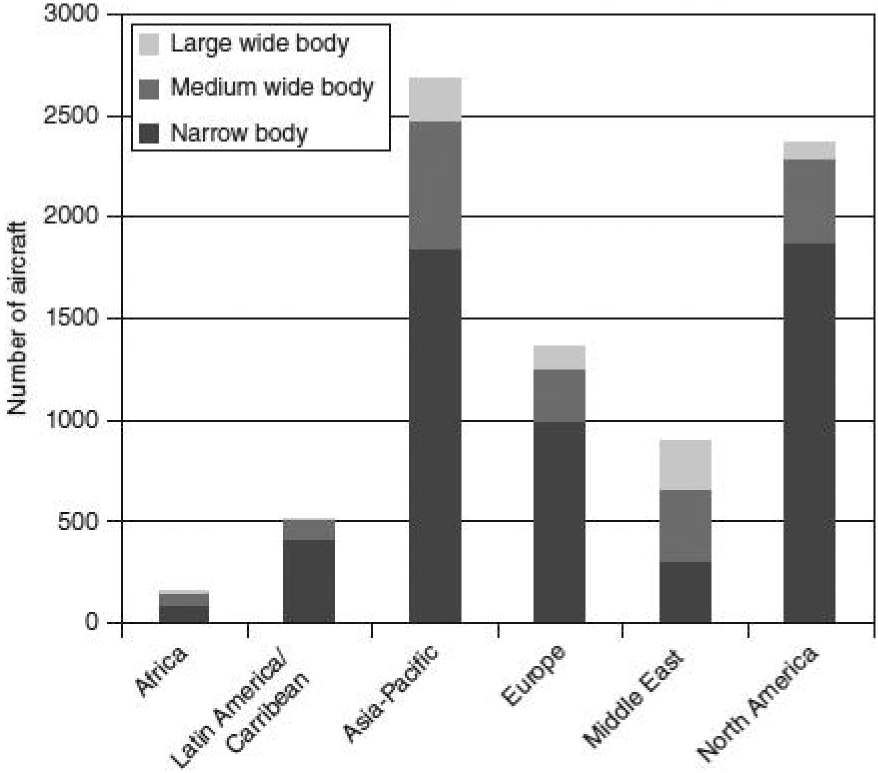


FIGURE 2.5 Large commercial jet order backlog, December 2011. (*Courtesy*: V. Surges, MIT. *Source*: Manufacturer web sites—www.airbus.com, www.boeing.com.)

Although Fig. 2.5 shows some differences in the aircraft type composition of orders by region, the overall picture is that airlines in every region will continue to acquire aircraft of all types. There is no apparent trend toward larger or smaller aircraft, but it is true that Asia-Pacific and Middle East carriers have a greater proportion of wide-body aircraft on order. This reflects both the geographical realities of these two regions (most flights are international and longer distance) and the aggressive expansion plans of the largest carriers operating in these regions. In contrast, North American and European airlines continue to require a greater proportion of smaller narrow-body aircraft to serve shorter-haul routes where frequency competition is more important.

Airline decisions, based on their network structures and competitive scheduling practices, determine the aircraft types that serve any individual airport. Large international carriers will focus on wide-body aircraft, but they will also need the smallest regional jets to provide connecting feed on short-haul routes. New entrant LCCs initially focused on 150-seat narrow-body aircraft, but there is recent evidence of shifts in both directions—Air

Asia uses larger aircraft for some international routes in Asia, whereas JetBlue in the United States and Azul in Brazil have acquired smaller 100-seat aircraft for lower demand routes.

For airports, the diversity of airline fleet characteristics and the absence of universal trends in aircraft types used by airlines simply mean that tremendous flexibility will continue to be paramount. From the smallest regional jets to the largest A380, different aircraft types can have significant airport implications in terms of gate configurations, runway and taxiway requirements, as well as terminal waiting lounge and passenger processing facilities. With the increased pressures of airline competition, volatility of airline profitability and growing movement toward consolidation in the global airline industry, airports will have to accommodate a range of aircraft sizes at any given time and will also have to respond to changes in the fleet characteristics of their airline tenants, sometimes with little advance notice.

2.2 Airline Network Structures

The dominant network structure for the vast majority of the world's largest airlines is the "hub-and-spoke" model, in contrast to the simpler "point-to-point" operations of some smaller new entrant carriers. The large hub airlines depend on connecting passenger traffic to increase loads and revenues on flights into and out of their hub airports. Some airlines with relatively low local market demand at their hubs, such as KLM (Amsterdam hub), Singapore Airlines (Singapore hub), and the rapidly growing Emirates (Dubai hub), could not have grown to their current size without focusing to a large extent on connecting passengers that transit their hubs (also known as "sixth freedom traffic").

Hub-and-spoke network structures allow airlines to serve many origin-destination (O-D) markets with fewer flights, requiring fewer aircraft departures that generate fewer "available seat-kilometers" (ASKs) at lower total operating costs than a complete point-to-point route network. Consider a hypothetical connecting hub network with 10 flights into and 10 flights out of a single "connecting bank" at a hub airport, as shown in Fig. 2.6. A "connecting bank" refers to pattern of operations in which many aircraft arrive within a short period of time at the hub airport, passengers and baggage transfer between connecting flights, and the aircraft then depart with the connecting passengers and baggage on board. Connecting banks can last from approximately 1 hour at smaller domestic hub airports to 2 to 3 hours or longer at larger international hubs.

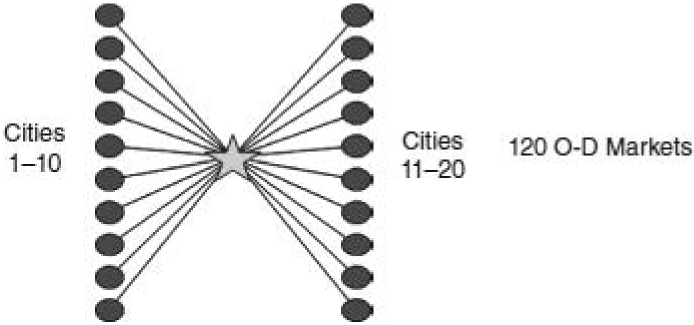


FIGURE 2.6 Hypothetical connecting hub network.

In this small example of a connecting bank, each flight leg arriving or departing the hub provides simultaneous service to 11 O-D markets—one "local" market between the hub and the spoke, plus 10 additional "connecting" markets. This airline thus provides service to a total of 120 O-D markets with only 20 flight legs and as few as 10 aircraft that traverse the hub. In contrast, a complete "point-to-point" network providing nonstop service to each market in this example would require 120 flight legs and 50 or more aircraft, depending on scheduling patterns and aircraft rotation requirements.

By consolidating traffic from many different O-D markets on each flight leg into and out of the hub airport, the airline can provide connecting service to low-demand O-D markets that cannot support nonstop flights. Consolidation of O-D market demands further allows the hub airline to provide an increased frequency of connecting departures, by offering multiple connecting banks per day at its hub airport. This increased departure frequency further increases the airline's revenues and contributes to higher market shares relative to its competitors.

Hub networks require substantially fewer flights and aircraft (as well as flight crew and other airline staff) to serve a large number of O-D markets, as compared to complete point-to-point networks. The concentration of its operations at a large hub airport also provides the hub airline with additional operational and cost advantages—economies of scale in terms of its aircraft maintenance operations, catering facilities, and airport ground handling services, for example. Hub operations also give the airline more opportunities for real-time "swapping" of aircraft in response to mechanical or weather delays and cancellations, given the large number of aircraft that converge at the hub during a connecting bank.

Hub operations also create incremental costs for the airline. Longer aircraft ground or "turn" times associated with connecting hubs can reduce aircraft and crew utilization compared to point-to-point networks. Whereas a point-to-point LCC can turn a narrow-body aircraft in 20 to 30 minutes, a large hub airline will keep the same type of aircraft (as well as its pilots and flight attendants) on the ground at the hub for 60 minutes or more, to accommodate connecting passengers and baggage. Increased turn times reduce the output of each aircraft (ASKs) over which fixed costs can be spread, leading to higher unit costs.

A large hub operation can also result in uneven use of airport resources (such as airport gates and runway capacity), and of airline resources and personnel. Surges of arrivals and departures during connecting banks require high levels of ground service and gate staffing, while leaving these human resources underutilized during off-peak periods. The number of scheduled departures and arrivals during connecting banks can exceed the airport's runway capacity, leading to flight delays in peak periods and unused capacity in off-peak periods. Operationally, weather delays at the hub airport can have severe impacts on the ability of passengers to connect successfully at the hub according to plan. Missed passenger and baggage connections in turn increase operating costs for the airline.

From a route planning perspective, a hub-and-spoke network structure affects how airlines evaluate the economics of new services. New routes to smaller spoke cities become easier to justify in an established hub network. In the hypothetical hub network of Fig. 2.6, the airline might require only five passengers per flight out of a new spoke city to each of 10 connecting destinations (in addition to the spoke-to-hub "local" demand of, say, 25) to make the operation of that flight with a 100-seat aircraft profitable. Even if the local O-D market demand is too small to justify the new service on its own, the new connecting passengers carried by the flight can make an incremental contribution to the airline's total network revenue that exceeds the operating costs of the new service.

Despite repeated forecasts of more point-to-point flights, the development of bigger and stronger hubs has continued in all regions of the world, especially during slow economic times and/or periods of high fuel costs. During the financial crisis of 2008, the largest U.S. and European airlines responded to the drop in demand and spiking fuel prices by eliminating virtually all flights that did not originate or terminate at their hubs.

The reliance of U.S. airlines on hub operations is very high and increasing in recent years. As Fig. 2.7 shows, well over 90 percent of all U.S. domestic flights in 2010 originated or terminated at major connecting hub airports for all of the large legacy airlines—American, Delta/Northwest, United/Continental, and US Airways. For United/Continental in particular the proportion of hub flights exceeded 99 percent in 2010. This reliance on hub operations is not limited to U.S. legacy airlines. Low-cost carriers AirTran, Frontier, and JetBlue all operate over 80 percent of their domestic flights through their own connecting hubs. The only exception is Southwest, which pioneered the point-to-point style replicated by other LCCs around the world, but even it now operates over 50 percent of its flights into a connecting hub. As air transportation markets mature, the opportunity for LCCs to profitably serve point-to-point routes without any connecting traffic support diminishes. Although we have not yet seen this same level of saturation of LCC services in other regions of the world, the U.S. experience is nonetheless instructive.

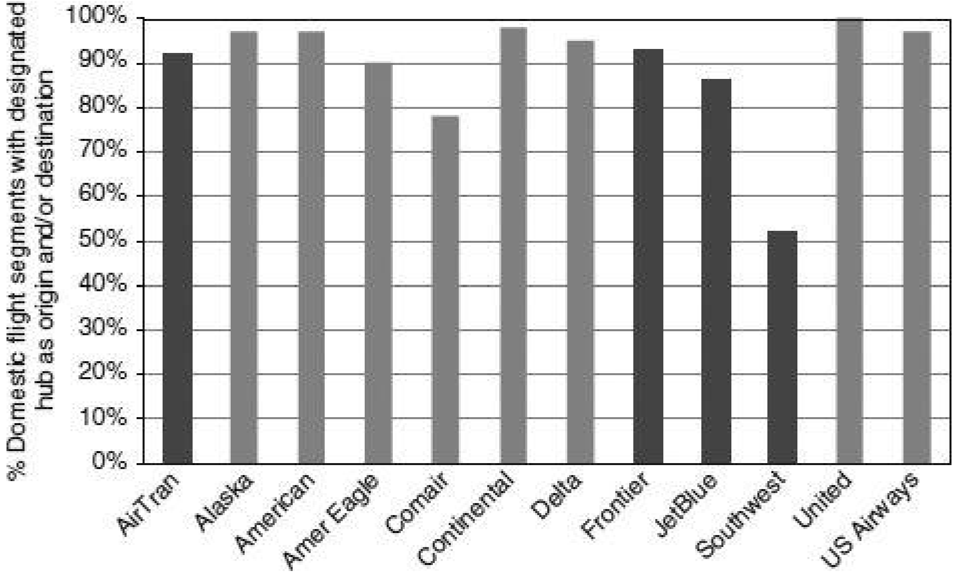


FIGURE 2.7 Proportion of hub flights operated by U.S. airlines, 2010. (*Source*: Belobaba et al., 2011.)

For the vast majority of world airlines, the economic advantages of hub network operations have consistently outweighed their operational costs. There is little reason to expect the dominant hub-and-spoke network model to falter. There are undoubtedly still many routes in the world that can support new nonstop service by an LCC focused exclusively on serving local point-to-point traffic. However, as air travel markets mature and LCC costs rise, these opportunities will inevitably become scarcer. As has occurred in North America, LCCs in other world regions will have to consider some form of connecting hub operation to contribute incremental traffic and revenues to sustain their growth plans and profitability.

Several global airline industry trends reinforce the reliance on the connecting hub model among non-LCC airlines. These include the increasing liberalization of international routes, growing global alliances, as well as the development of new longer-range aircraft with smaller capacities, described in the previous section.

"Open-skies" bilateral agreements between countries remove most of the regulatory constraints on the scheduling and pricing of international services. They effectively allow all airlines of either country to operate flights between any two points in the countries involved. They allow airlines to fly what once were thought to be relatively low-demand international nonstop routes from their hubs (e.g., Salt Lake City-Paris/de Gaulle by Delta, Frankfurt-Phoenix by Lufthansa, and Dubai-Hamburg by Emirates). The growth of global airline alliances has encouraged these new international services, with one or both end points being major hubs for one of the partners in the alliance. For example, Salt Lake City is a Delta hub and Paris de Gaulle is an Air France hub, and both carriers are partners in the SkyTeam alliance (see Chap. 1). In addition, the increased range capabilities of smaller international aircraft like the Boeing 767 and 787, and the Airbus A330 and A350 mean that airlines can offer these nonstop flights with a lower risk of not filling seats.

For airports, these trends in network evolution and airline route planning suggest that a proactive approach to attracting new airlines and new routes could be beneficial. Airports must, however, understand the changing business models and network characteristics of airlines with differing values and objectives, in order to offer them attractive proposals. An LCC with primarily point-to-point operations will be most interested in finding new airport destinations with large traffic catchment areas, which can offer lower user fees and improved operational reliability than competing airports (e.g., in terms of short turnaround times and lack of congestion). A large network carrier considering a new service from an airport to its hub will be more interested in the potential for its flight to capture an adequate amount of connecting traffic from the new spoke city via its hub, above and beyond the local market potential.

2.3 Airline Scheduling and Fleet Assignment Optimization

Airlines must develop feasible and profitable schedules of operations for their aircraft and crews, given decisions about fleet and network structure. The process of developing airline schedules typically begins a year or more before the flight departure date. It then continues right up until the departure time of the flight, as last-minute schedule changes or flight cancellations might be required to deal with unexpected or "irregular" operations. The airline schedule development process involves four interrelated decisions:

1. *Frequency*: How many flights per day should be operated on each route in the network?
2. *Timetable development*: What will be the departure and arrival times of each flight?
3. *Fleet assignment*: What aircraft type will be used for each flight departure?
4. *Aircraft rotations*: How will available aircraft be routed over the airline's network?

This section describes each of these decisions briefly, to provide a basis for understanding the impacts of airline schedules on the operations of both airlines and airports.

The choice of what frequency to operate on a specific route depends on both competitive and economic considerations. Greater frequency of departures on a route improves the "schedule coverage" of an airline, that is, the proportion of desired passenger departure times that can be accommodated by the airline's flight departure times. Greater schedule coverage is particularly important for time-sensitive business travelers. More frequent flights improve the convenience of air travel for passengers and reward the airline with higher traffic, revenues, and increased market share at the expense of its competitors.

In competitive markets, airline frequency share is the most important factor that determines each airline's market share of total demand, assuming that both prices and on-board service quality tend to be similar among competing carriers. The shorter the distance involved, the more important frequency share is, given that actual flight times represent a relatively small proportion of the passengers' total travel time. In these markets, it is common for competing airlines to operate smaller capacity aircraft with higher operating costs per seat and per seat-kilometer, trading off higher costs against the revenue benefits of higher market share.

The objective of "load consolidation" also affects airline frequency decisions on a route. Consolidating passenger traffic from multiple O-D markets onto one aircraft can allow that airline to operate higher frequency on the route (increasing its market share) and/or larger aircraft (reducing its unit operating costs). This ability to consolidate loads is a fundamental reason for the economic success of airline hubs.

Given a chosen frequency of departures on a route, the process of timetable development determines the specific departure and arrival times of each flight. All else equal, peak departure times (early morning and late afternoon) are most attractive both to business travelers willing to pay higher fares and to many leisure travelers as well. However, developing a timetable of flight departures requires airline schedulers to make tradeoffs between aircraft utilization (block hours per day) and schedule convenience for the passengers. See Example 2.1.

Example 2.1 A peak-hour departure at 17.00 from airport A and arrival at airport B at 19.30 is more likely to be attractive to business passengers and therefore profitable for the airline. Once that aircraft arrives at airport B, a minimum "turnaround" time is required to prepare it for the next flight. Turnaround times vary by aircraft type and the characteristics of the flights involved—a narrow-body aircraft in domestic service might require as little as 30 minutes, whereas a wide-body international aircraft takes 2 hours or more to prepare it for the next departure.

Assuming a 45-minute "turnaround" time, the aircraft arriving at airport B at 19.30 could be scheduled to depart again as soon as 20.15. However, a late evening departure might not be attractive to many passengers, so the airline scheduler must decide whether to operate the return flight at 20.15 with fewer passengers and less revenue or hold the aircraft (and potentially its crew) on the ground overnight until the next peak departure time, say at 07.30 the following day. In the latter case, the aircraft and its crew will be idle for 12 hours, reducing aircraft and crew utilization and increasing costs.

Typically, most airlines develop timetables to maximize aircraft utilization. They keep "turnaround" times to a minimum and the aircraft and crews flying as much as possible to reduce unit costs. This approach can lead to off-peak flights with relatively low load factors, which might be necessary to maintain competitive frequency share and to position aircraft for peak flights at other cities. It can also leave little buffer time for maintenance and weather delays, if adequate slack is not built into the timetable.

A variety of factors can constrain timetable development. Airline hubs with fixed connecting banks require that flights arrive from spoke cities and depart from the hub at predetermined times. At large hubs, the use of connecting banks creates surges of aircraft and passenger activity that require relatively high airport capacities, in terms of both airside elements (runways, taxiways, and gates) and terminal facilities. The most successful connecting hub airports have been able to expand their capacities (e.g., Atlanta) to facilitate connections, in contrast to more constrained airports that are close to saturation (e.g., Tokyo/Narita and London/Heathrow).

Time zone differences also limit feasible departure and arrival times, especially on long-haul routes. For example, flights from eastern cities in North America to Europe typically do not depart before 16.00, as passengers do not want to arrive at their European destination much before 06.00 local time. Regulatory constraints, such as airport arrival and departure slot times, and noise curfews can further limit the scheduling flexibility for an airline. Finally, crew scheduling and aircraft maintenance requirements can also impose significant constraints on timetable development.

Fleet assignment is the problem of allocating the specific aircraft type to be flown on each flight leg, given a network of routes, a set of flight departure times and available aircraft types from the airline's existing fleet. The objective of fleet assignment is to minimize the combined costs of "spill" (rejected demand and lost revenue) and aircraft operating costs. Spill occurs when the aircraft assigned to a flight departure is too small and potential demand and revenues are lost to the airline. Airlines can reduce (or eliminate) spill by assigning a large enough aircraft to accommodate all possible peak-day demands for the flight in question. However, larger aircraft have higher operating costs and will fly with many empty seats on most nonpeak days.

Many airlines use fleet assignment software tools based on large-scale mathematical network optimization methods. These assign aircraft to maximize expected profitability, subject to constraints such as minimum ground times, maintenance requirements, and number of aircraft by type available in the airline's fleet. Aircraft routing models are used to assign specific aircraft "tail numbers" to each flight, creating rotations that satisfy aircraft maintenance requirements and maintain a balance of inbound and outbound flights at each airport. The use of these optimization tools has allowed airlines to achieve higher aircraft utilization rates and reduce total aircraft and crew costs.

The final product of the airline schedule development process is a detailed plan of how the airline will operate on a given date in the future. It includes aircraft schedules, crew assignments, and a large number of other operational details. This plan will have been optimized to reduce costs, increase revenues, and maximize profit under what are typically assumed to be favorable operating conditions. Unfortunately, almost every day presents a variety of unexpected and unplanned events that force any airline to deviate from its optimized schedule.

Dealing with "irregular operations" requires airlines to revise their planned schedule right up until the flight departs or is cancelled. A cancelled flight can seriously disrupt aircraft rotations, crew schedules, and maintenance plans, not to mention passenger trips. Under conditions of disruptions and/or flight cancellations, the primary objective for the airline is to return to normal operations as quickly as possible. In this effort to get the airline "back on plan" with respect to the planned timetable, flight cancellations or aircraft rerouting sometimes take precedence over passenger convenience. The next section describes the characteristics of airline operations occurring at the airport that can contribute to these deviations from the planned schedule.

2.4 Airline Operations at the Airport

Much of the uncertainty and volatility affecting airline operations stems from activities that occur at the departure and/or arrival airports for any given flight. Although there are factors that contribute to variability in the actual airborne times of each flight on a given route, operational issues involving the handling of passengers, baggage, and aircraft on the ground can significantly impact the ability of the airline to operate its planned schedule on time. These operational issues can affect the operations of the airport itself, in terms of congestion and delays on the airside, as well as gate utilization and passenger flows on the landside. This section provides a brief overview of airline ground operations at the airport and describes the variability of different stages of the planned flight schedule that can lead to irregularities and delays.

At each airport served by an airline, it is the responsibility of the airline's ground operations staff, sometimes called "station control," to implement the schedule plan without compromising safety, subject to company goals for on-time performance and overall flight completion. They must coordinate aircraft and crew operations, and process the airline's passengers, baggage, and cargo subject to a large variety of operational constraints, for example:

- No matter how large the airline's operation at an airport, there is a limit on the number of gates available to the airline, many with constraints on the size of aircraft they can accommodate.
- There might also be flow limitations on the capacity of the airport's taxiways and runways, which can be affected by weather and resulting field conditions.

- The availability of equipment and ground crew resources, operated by both the airline and the airport.
- Air traffic control (ATC) congestion and delays, both at the airport and those that affect flights en route to or from the airport.

If the airport is a connecting hub for the airline, station control also tries to ensure that passengers and baggage make their connecting flights. They must trade off the costs and benefits of holding a flight at the gate beyond its scheduled departure time when one or more incoming flights have been disrupted and arrive later than planned. Delaying the departing flight to accommodate passenger and baggage connections improves the passengers' travel experience and reduces the costs to the airline associated with passenger rebooking, including meals and hotel accommodations, and delayed baggage delivery to the destination. On the other hand, holding the flight adds to airline operating costs and can result in further down-line delays for the next scheduled departures of the same aircraft and/or crews.

The "turnaround" activities associated with each arriving and departing aircraft at an airport are central to keeping the airline operations running smoothly and are major factors in delays that can affect airport operations as well. These activities include deplaning passengers and baggage, cleaning and catering the aircraft, performing required security checks, refueling the aircraft, and then boarding the passengers for the next departure. Completion of the flight turnaround process within the scheduled turn times is a critical factor in the airline's ability to operate its planned schedule with minimal delays. Although a significant portion of airline delays are associated with schedule disruptions away from the gate (e.g., taxiway queues awaiting takeoff or airborne delays resulting from en route congestion), many other flight delays are encountered and absorbed at the gate, which affects the overall operation of the airport.

Most airlines incorporate "buffer" time into the planned schedule for each flight to take into account what can be substantial variability in actual flight block times due to various delays. Figure 2.8 shows the distribution of actual block times for a large sample of flights that operated on the New York/Newark (EWR) to Los Angeles/International (LAX) nonstop route. The median block time is 363 minutes (or 6:03) for this 2450-mile flight, while the dispersion is quite large. Actual block times ranged from about 310 to almost 500 minutes. (Note that EWR can be a very congested airport subject to substantial air traffic delays, which explain the long delays apparent in this distribution.) If an airline published a block time of 363 minutes, 50 percent of its actual flights would arrive after their scheduled arrival time (but, of course, 50 percent would also arrive early!). Increasing the planned block time reduces the probability of late arrivals but at the same time increases the likelihood of early arrivals at the destination airport. Early arrivals not only lead to reduced utilization of aircraft and crews, they introduce their own form of “irregular operation” for the airline, as gates, ground crews, and baggage handlers might not be available at the time of the early arrival.

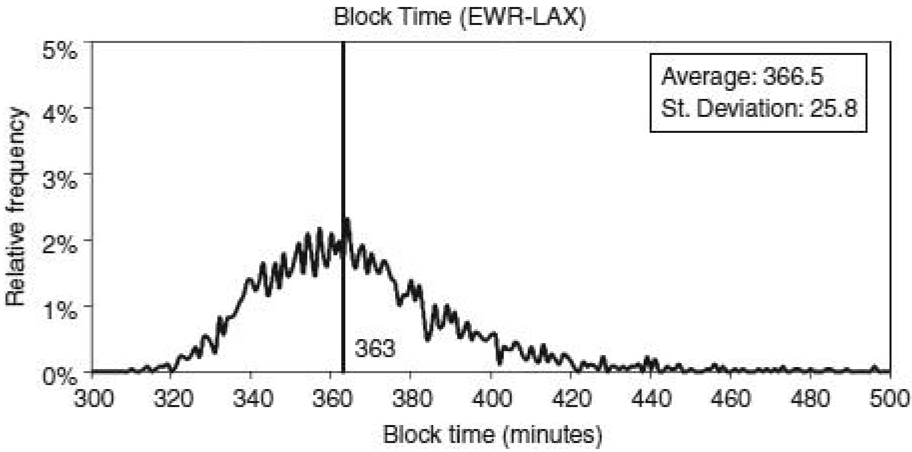


FIGURE 2.8 Distribution of actual block times. (*Source*: Skaltsas, 2011.)

Decisions made by airline schedulers regarding planned departure, arrival, and block times can significantly impact the operations of the airline at an airport, and thus the operations of the airport itself. For any given flight schedule, the operational variability for an actual flight on a given date can result in flight delays, missed passenger and baggage connections, irregular operations, and even flight cancellations for the airline. For the airport, this variability can create problems in terms of gate availability for other scheduled flights. Figure 2.9 shows the distribution of gate delay times for the same sample of EWR-LAX flights. Although some flights departed the gate as much as 10 minutes before scheduled departure time, the median gate delay time was 5 minutes and the average was more than 17 minutes. These gate delays and their variability, in particular, make it more difficult for both the airline and the airport operator to optimize the utilization of gates, ground crews, and related scarce resources.

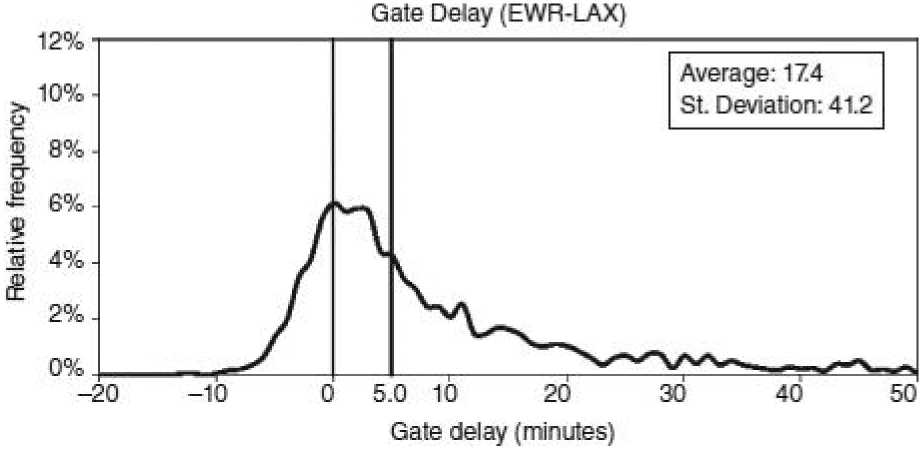


FIGURE 2.9 Distribution of gate delay times. (*Source*: Skaltsas, 2011.)

Whether the delay occurs at the departure gate or during any of the subsequent stages of the flight's operation, a late arrival at the destination airport is an inconvenience for passengers. For a flight into an airline's hub airport, an arrival delay of even 30 to 60 minutes can translate into a major disruption of travel plans if it means that passengers cannot make their connecting flights. At best, a missed connection delays the passenger's arrival at their destination by several hours, depending on when the next flight from the hub to the passenger's destination is scheduled. However, with increasing airline load factors, reaccommodation of disrupted passengers on later flights becomes more difficult due to a lack of available seats.

Delayed flights complicate operations for both the airline and the airport operator. Disrupted passengers end up spending substantially more time (and perhaps more of their money) at the airport. They inevitably end up with a negative perception of their overall travel experience that they typically attribute to both the airport and the airline (or in some cases, incorrectly to the airport alone). For airports wishing to develop and improve customer perceptions of their services, it is therefore important to identify ways in which airport operators can collaborate with airline ground operations to reduce the negative impacts of flight delays, not only on operational factors but also on the passengers themselves.

2.5 Airline Operating Costs and Productivity

With the deregulation of airline markets and competition from new entrant LCCs with significantly lower cost structures, airlines focused their attention on cost management. The evolution of fleets, routes, scheduling and operating practices described in previous sections has contributed to substantial changes to airline cost structures and improvements in airline productivity measures. This section illustrates these trends with operating cost data for U.S. airlines and provides supporting evidence of similar trends for non-U.S. airlines to the extent data are available. It first discusses changes in the shares of different operating cost components, followed by an overview of unit cost trends and a comparison of network legacy carrier (NLC) and LCC unit costs and their convergence in the recent past. As before, it considers the implications of these trends for both airlines and airports.

Airline operating costs can be allocated to different categories in a variety of ways. A common approach is that of "administrative" cost categorization. This is often used in airline financial statements, which identify expenses for labor, fuel, capital, materials, and various services used to produce the airline's output. Figure 2.10 shows these costs as reported by U.S. airlines for 2011. The two largest cost categories are fuel and labor, which represented approximately 32 percent and 27 percent, respectively, of total U.S. airline operating expenses.

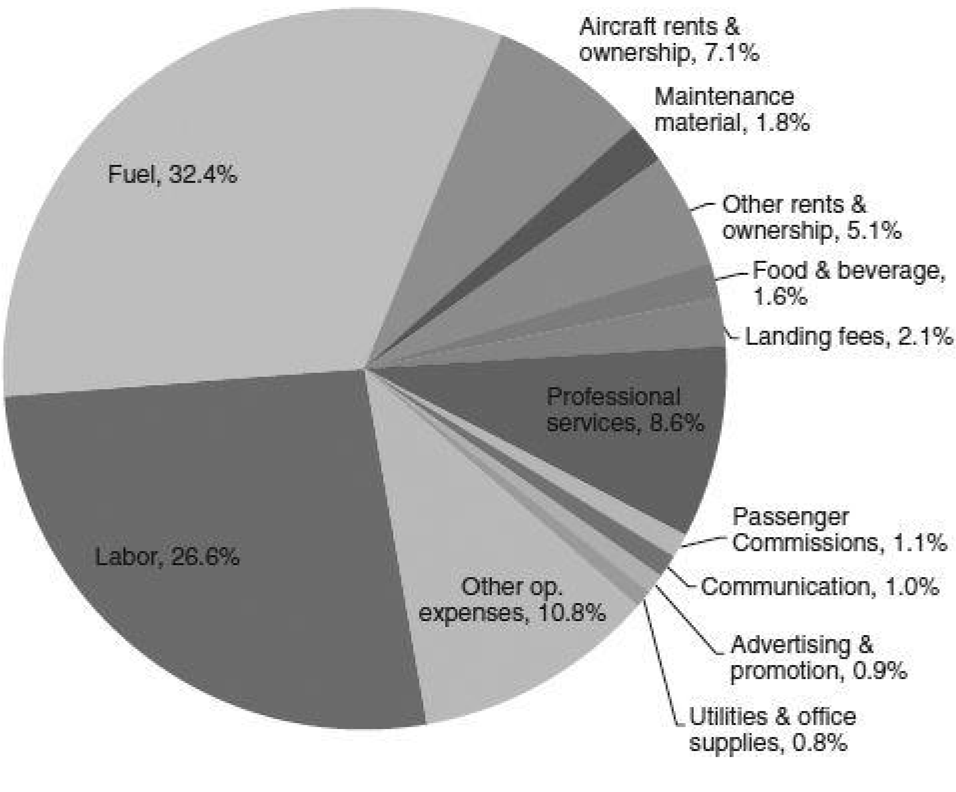


FIGURE 2.10 U.S. airline operating cost breakdown, 2011. (*Source*: Air Transport Association, 2012.)

Historically, fuel accounted for a smaller portion of total operating expenses. Fuel costs accounted for a record high 36 percent of total airline operating expenses in 2008, compared to about 30 percent during the first fuel crisis after deregulation in 1980. Its proportion of total operating expenses again increased to 32 percent in 2011, and future fuel price increases could push this proportion higher. Labor costs, on the other hand, have declined in terms of absolute and relative contribution to total operating expenses, especially since the bankruptcies and subsequent restructuring by U.S. NLCs in the early 2000s. The share of total operating expenses related to labor decreased from 42 percent in 1978 to 27 percent in 2011. Nonlabor costs have fluctuated as a proportion of total expenses, as the role of many of the smaller contributors to nonlabor costs has changed considerably as well: Outside maintenance, nonaircraft ownership, and aircraft rental costs have increased, while the once significant commissions (travel agency payments) category has all but disappeared. Airport landing fees account for about 2 percent of total operating costs for U.S. airlines, but the airlines continue to express concerns about the growth of this expense category.

Airlines around the world have experienced similar changes in the share of major cost categories. They have experienced significant increases in the share of fuel costs and concomitant reductions in the shares of other cost categories, notably reductions in labor expenses. Table 2.1 compares cost component shares across airlines from different regions of the world. The increase in fuel cost shares between 2001 and 2008 is particularly striking, as the fuel cost component climbed to 32.3 percent of total airline operating costs from only 13.6 percent in 2001. The share of labor costs decreased in all regions (due in part to the increases in fuel cost share) over this same time, but the decrease in the labor cost shares of North American carriers is the greatest of all regions, dropping from the highest of all regions to approximately the world average. Asia-Pacific airlines maintained the lowest labor cost share of the regions shown.

	North America		Europe		Asia Pacific	
	2001 (%)	2008 (%)	2001 (%)	2008 (%)	2001 (%)	2008 (%)
Labor	36	22	27	25	17	15
Fuel	13	34	12	25	16	37
Aircraft (Rentals/ ownership)	12	8	10	8	14	12
Other	39	36	51	42	53	36

Source: International Air Transport Association, 2010.

Table 2.1 Share of Major Operating Cost Components by Region

Unit cost is the ratio of the airline's total operating expenses to ASKs produced [or ATKs (available ton-kilometers) in the case of cargo airlines]. For passenger airlines, unit cost is also known as "CASK," meaning "cost per ASK" [the equivalent measure in miles is "CASM" (cost per available seat-mile)]. The relationships between unit costs and airline size, average aircraft capacity and average stage length are all expected to be negative, in theory at least. That is, a large airline is expected to see some economies of scale (reduction in unit costs with increased output), as its fixed costs are spread over a larger output of ASKs. A larger-capacity aircraft is also expected to show some economies of aircraft size, as the fixed costs are spread over more seats for any given flight, resulting in lower costs per seat. Likewise, longer stage lengths mean that the relatively fixed costs of ground servicing, for example, can be spread over more ASKs produced by each flight.

The differences in total unit costs among airlines around the world are reflected in the comparison of 25 of the largest passenger airlines (ranked in terms of revenue passenger-kilometers carried), shown in Fig. 2.11. This graph plots the 2010 total unit costs against average stage length, as reported in the Airline Business financial database (Airline Business, 2011). In general, there exists an inverse relationship between total unit cost and average stage length, which is apparent in Fig. 2.11. Several airlines with the longest stage lengths—Emirates, Cathay Pacific, EVA, and Thai Airways—report relatively low unit costs, although Singapore Airlines' unit cost appears to be somewhat higher given its long stage length. On the other hand, large network airlines with shorter stage lengths report significantly higher unit costs. Valid comparisons of unit costs across airlines require taking average stage length into account. Looking at Fig. 2.11, we cannot conclude that Delta has a cost disadvantage compared to Emirates, given the large difference in their average stage lengths. However, two airlines with similar stage lengths can be compared; for example, American's unit cost is indeed higher than Delta's. The true outliers are the airlines with short trip lengths *and* lower unit costs—specifically the LCCs—Southwest, easyJet, and JetBlue, and particularly Ryanair.

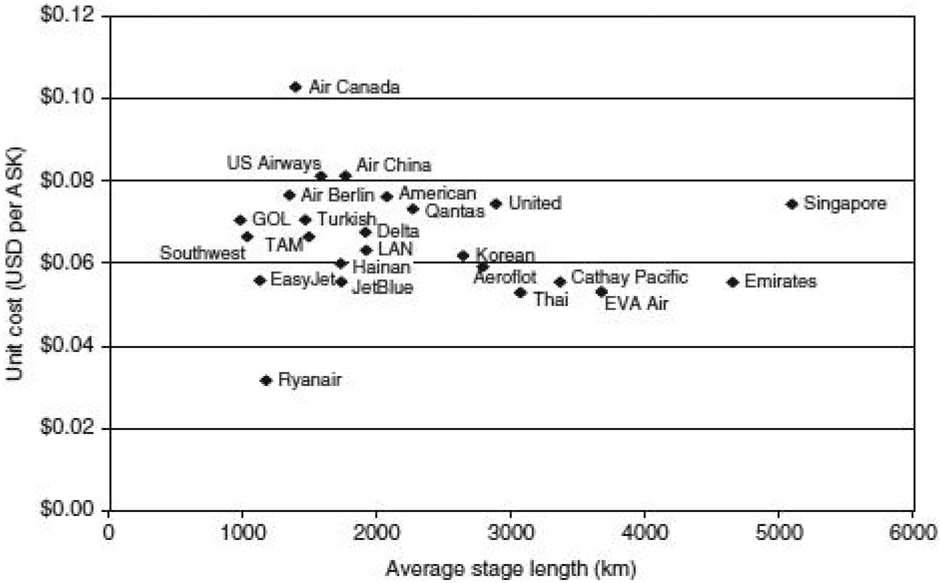


FIGURE 2.11 Unit costs versus stage length, 2010. (*Source*: Airline Business Database, 2011.)

A further comparison of airline unit costs focuses on three broad categories: fuel, labor, and nonlabor expenses. Fuel expenses are most straightforward to categorize, whereas labor costs include total salaries, benefits, and other costs paid to employees. Nonlabor costs include all other operating expenses. This last category includes cost items that represent the "structural" costs of the airline over which management can exert influence and are therefore an indication of how airline network and product strategies affect "controllable costs" not related to fuel or labor inputs.

The average unit cost of U.S. passenger airlines, expressed in constant dollars per ASM, has declined almost 40 percent since deregulation. In 1979, it cost the average U.S. airline an inflation adjusted[1] 18.3¢ to produce one ASM; that unit cost dropped to 11.2¢ in 2009 (Belobaba et al., 2011). Figure 2.12 shows the corresponding inflation-adjusted unit costs for the fuel, labor, and nonlabor categories. A substantial part of the reduction in real nonlabor unit costs since the mid-1990s can be attributed to dramatic cuts in airline distribution costs—first with the elimination of travel agency commissions in the late 1990s, followed by the use of Internet and related technologies for ticket distribution since 2000. Overall, nonlabor costs have decreased by 25 percent in real terms since 1978.

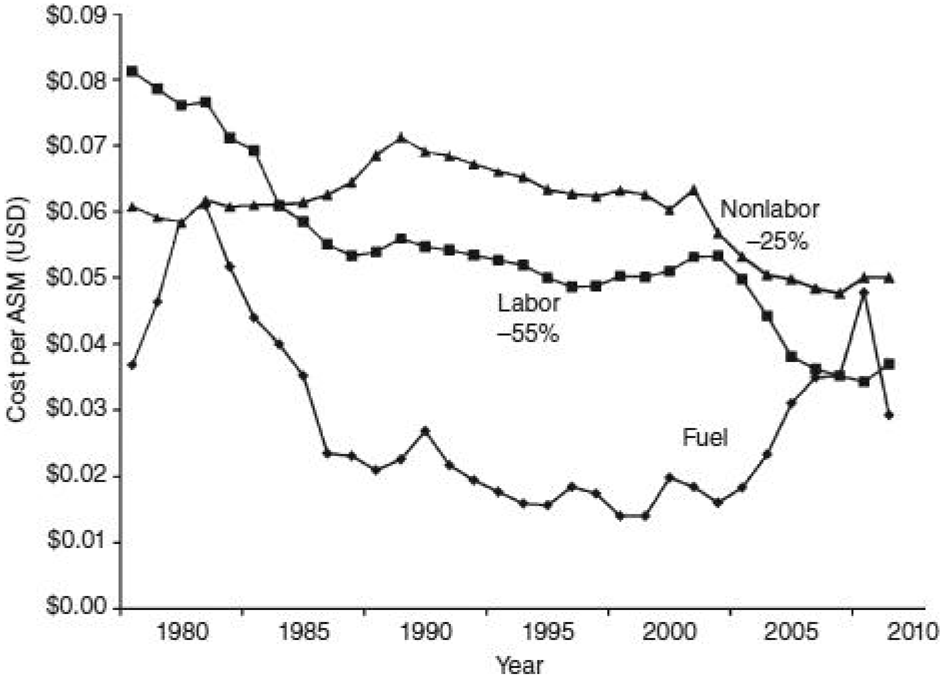


FIGURE 2.12 U.S. airline real unit costs by category. (*Source*: Belobaba et al., 2011.)

A dramatic drop in U.S. passenger airline labor unit costs occurred between 2002 and 2006, as NLCs went through bankruptcies, layoffs, and restructuring of labor contracts. In cumulative terms, the average real labor unit cost for U.S. airlines has decreased by 55 percent since deregulation in 1978. Together, the labor and nonlabor operating cost categories combined (excluding fuel) have seen a 40 percent decrease in real unit costs since then. Fuel unit costs expressed in inflation-adjusted terms, on the other hand, have exhibited much greater volatility than the other two cost categories. Very high fuel unit costs in the early 1980s exceeded the recent peak in 2008 in real terms, but much of the period from the late 1980s through the early 2000s was characterized by fairly low and stable real fuel unit costs. Fuel unit costs began to surge in 2005 and peaked in 2008.

Figure 2.13 compares inflation-adjusted NLC and LCC unit costs in the United States from 1990 to 2009. Over most of this period, the NLC group reported total unit costs approximately 2¢ (USD) per ASM higher than the LCC aggregate. In percentage terms, the unit costs of the two airline groups have been converging. LCCs still had a clear unit cost advantage in 2009, but their unit costs relative to NLCs were about 20 percent lower in 2009 compared to 30 percent lower in 2001. The nonlabor unit cost gap between the groups has remained approximately 1¢ per ASM. This reflects the airline's structural costs that are driven by a variety of factors such as network structure, fleet type, and outsourcing activity to name a few. NLCs have certain structural costs (hub operations, international flights, lounges, and other services) that result in this inherent and consistent nonlabor unit cost gap of about 1¢ per ASM. (Tsoukalas et al., 2008)

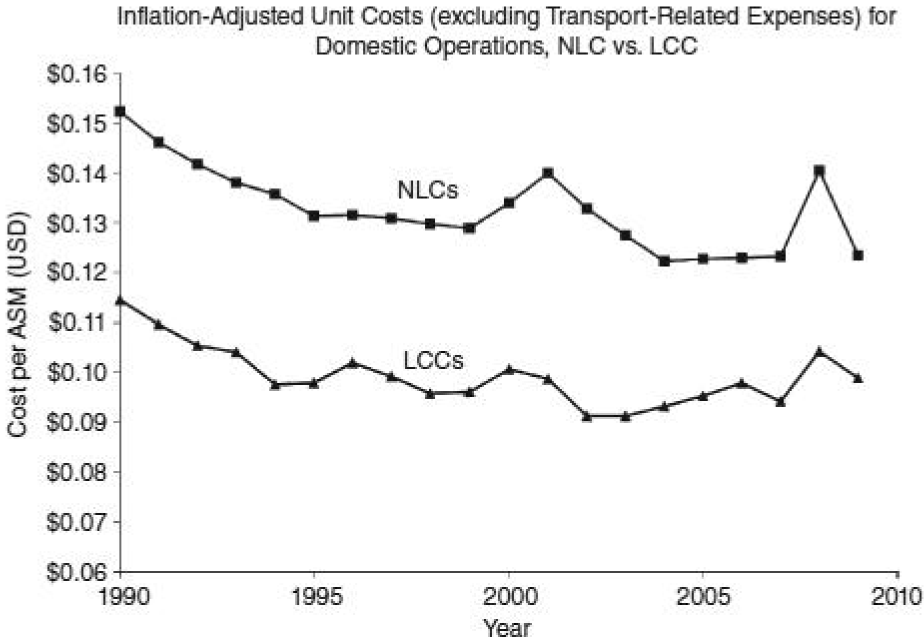


FIGURE 2.13 Inflation-adjusted unit costs, NLCs versus LCCs. (*Source*: Belobaba et al., 2011.)

Labor unit costs dropped dramatically after 2000, as Fig. 2.14 shows. Whereas labor costs in real terms gradually declined for LCCs, NLCs experienced a dramatic downturn as several of the largest carriers were able to renegotiate labor contracts and reduce workforces after filing Chap. 11 bankruptcy. This decline in labor unit costs substantially narrowed the historic gap between NLCs and LCCs. NLC real labor unit costs dropped by 40 percent between 2002 and 2007, increasing thereafter as many of the labor contracts of the early 2000s were up for renegotiation by 2008.

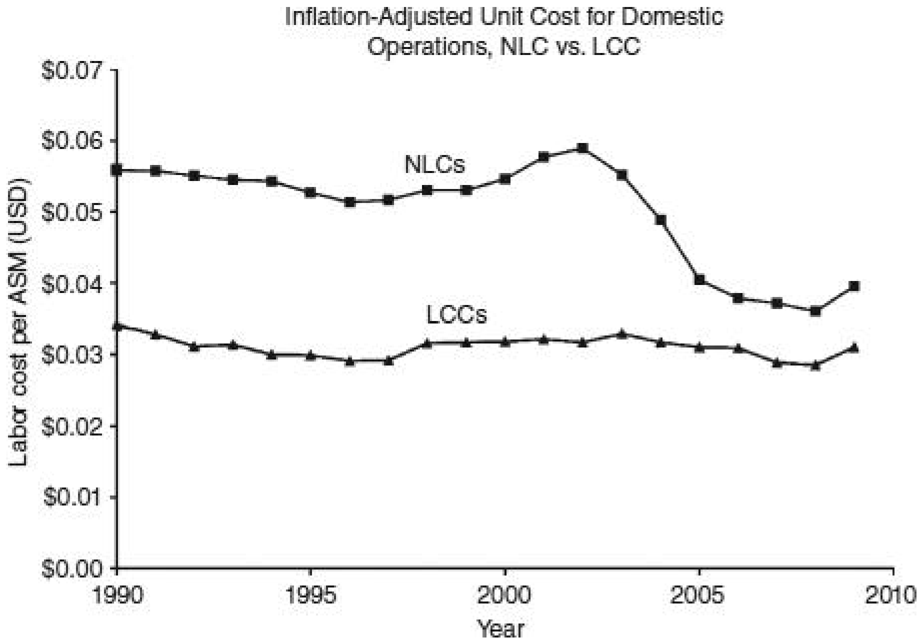


FIGURE 2.14 Inflation-adjusted labor unit costs, NLCs versus LCCs. (*Source*: Belobaba et al., 2011.)

The increased competition afforded by deregulation in the United States, especially the introduction and rapid growth of LCCs in U.S. domestic markets, has been the driving force behind these dramatic reductions in airline unit costs. These cost improvements were in large part passed on to consumers in the form of lower airfares. A continuing downward trend in average fares led to significant increases in the demand for air travel—in 2010, average inflation-adjusted U.S. airline fares remained at about 50 percent of their prederegulation levels (Belobaba et al., 2011). However, the downward trend in average fares has begun to level off, given the realities of higher fuel costs and the increasing difficulty faced by airlines hoping to achieve additional cost reductions. There is simply not much room for further declines in average fares, which raises questions about the continued rapid growth of air travel demand in the United States. Although there are still opportunities to reduce fares and stimulate demand in other world regions, the ability of airlines to manage costs will determine the extent to which this growth ultimately occurs.

Given that rapid growth of LCC competition was largely responsible for the cost reductions of U.S. legacy carriers, it is reasonable to expect that some form of cost restructuring will come to legacy carriers elsewhere. In particular, European legacy carriers have yet to undergo the same cost cutting pains as their U.S. counterparts but will face increasing pressure to do so in light of both increasing LCC presence within Europe and the growth of Middle East mega-carriers attempting to divert connecting traffic away from European hubs. As shown earlier in Fig. 2.11, the long-haul, large aircraft operations of the largest Middle East and Asian carriers give them a significant unit cost advantage.

In addition to reducing unit operating costs, airlines in the United States and around the world have improved aircraft and labor productivity through more efficient fleet utilization, hub network operations, and schedule optimization. In many cases, legacy airlines have adopted some of the innovative operational practices of their LCC competitors to further improve both cost efficiency and productivity. For example, legacy airlines have put an emphasis on reducing aircraft turn times at their hubs, in some cases moving to "continuous" banks instead of fixed connecting banks to increase aircraft productivity. They have renegotiated work rules with their unionized personnel to allow more cross-utilization of labor to perform different tasks—flight attendants who help clean up the aircraft cabin or assist in the boarding process at the gate are but two examples.

These strategies have led to substantial increases in aircraft and labor productivity for NLCs and LCCs alike, particularly since 2000. LCCs consistently have been able to achieve higher aircraft utilization, with more point-to-point flights and shorter turnaround times, whereas connecting hubs, international services, and even time zone constraints act to limit NLC utilization rates. For U.S. airlines, aircraft utilization has increased dramatically since deregulation, by an average of 20 percent through 2007 before decreasing with the more recent financial and fuel cost challenges. These improvements in aircraft utilization by both NLCs and LCCs translate directly into lower unit costs, as fixed operating costs are spread over more block-hours per day and, in turn, increased production of output (ASKs).

Employee productivity of the LCC group of U.S. airlines remains about 10 percent higher than that of NLCs, even as both groups have increased ASMs per employee by more than 35 percent in recent years (Belobaba et al., 2009). Figure 2.15 shows the evolution of labor productivity, expressed as ASMs per FTE (full-time equivalent employee). After peaking at almost 550,000 in 2000, U.S. airline employment plummeted by over 30 percent by 2010, driven largely by the NLC labor force cuts even as LCC employment continued to grow. NLCs thus achieved increases in labor productivity through substantial reductions in their workforce as well as through the relaxation of restrictive work rules in union contracts. Both NLCs and LCCs have also been able to increase employee productivity by replacing humans with technology—for making reservations, buying tickets, and checking in.

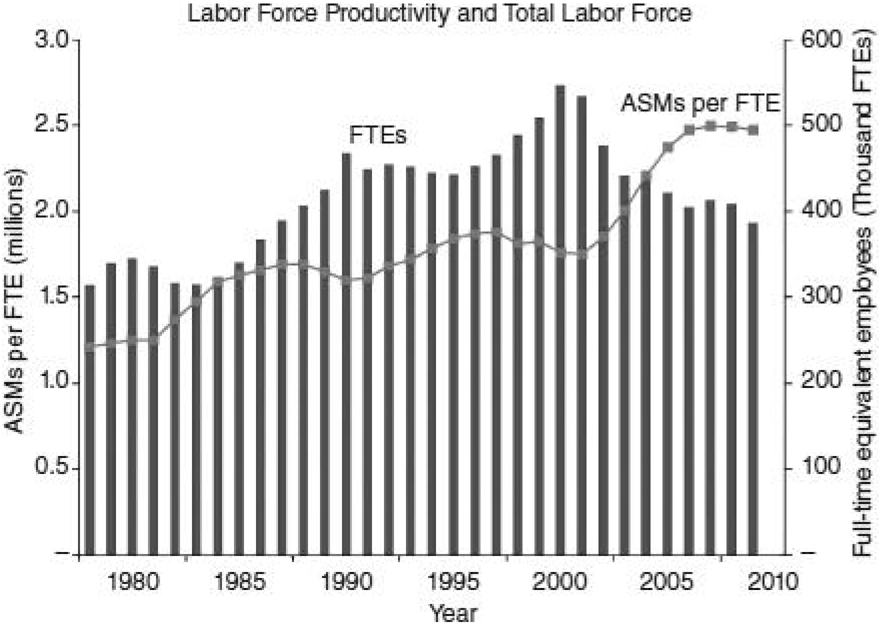


FIGURE 2.15 Labor productivity of U.S. airlines. (*Source*: Belobaba et al., 2011.)

Differences in airline operating costs, specifically unit costs, are largely driven by differences in the productivity of inputs, specifically, aircraft and labor. With the rapid growth in the United States and around the world of LCCs as effective competitors, the focus of established airlines on reduced unit costs and improved employee and aircraft productivity has become critical to profitability. With LCCs still experiencing rapid growth in many regions of the world, and with the continued growth of Middle East and Asian hub network airlines with structural unit cost advantages, all airlines will continue to emphasize cost cutting and productivity gains in virtually all of their management strategies.

The pressure on airports by airlines hoping to lower costs and increase productivity will intensify as global airline competition intensifies. LCCs considering new service to an airport are becoming more aggressive in demanding not only cost concessions from airports in the form of lower landing fees and rental rates, but guarantees of minimum traffic and/or revenue generation in many cases. Expanding mega-carriers with large connecting hubs will pursue cost concessions from their hub airport operator in light of the increasing volumes of traffic they are creating at their hubs. At the same time, as they consider the possibility of introducing services to new spoke airports to feed their connecting hubs, these mega-carriers will look for offers of lower landing fees and/or facilities costs from airports competing for the new services. In addition, all airlines will continue to resist increases in airport fees, and perhaps begin to adjust their networks and schedules in an effort to reduce the impacts of airport charges on their operating costs.

2.6 Summary

This chapter explored the effects of over 30 years of airline market deregulation and liberalization on airlines, and how the evolution of airline strategies and operating practices impact airports and their operators. The discussion considered the changing business models of different types of airlines, looking at trends in fleet composition, network structures, scheduling practices, as well as the unwavering emphasis by airline managers on cost reductions and productivity improvements. This final section explores the implications of these airline industry trends for the future, with a focus on the airport perspective.

Perhaps the most important trend affecting both the airline industry and airports over the past several decades has been the emergence and dramatic growth of new entrant LCC airlines. Although liberalization unleashed competition among existing airlines, it was the innovation of LCCs with dramatically different product offerings and lower cost structures that forced (or will force) all legacy airlines to adapt or perish. Many elements of the so-called "LCC business model" have been adopted by legacy airlines in an effort to remain competitive with LCCs. Scheduling faster aircraft turnaround times and increasing cross-utilization of airline employees at airports, for example, have helped legacy carriers improve aircraft and labor productivity. Reducing in-flight services and meals has also allowed legacy airlines to reduce unit costs—a strategy that at the same time provides the airport with increased opportunities to sell food to passengers before departure.

Just as legacy airlines have learned from new entrant LCCs, many LCCs are adopting more traditional product and service strategies to expand their market shares of the less price-sensitive demand segment. What were once "no-frills" airlines have begun to enhance their on-board services with live television and Internet options, as well as separate sections with extra legroom and, in some cases, even a business class cabin. Starting in the United States and spreading to LCCs around the world, there is much evidence of a convergence of LCC and NLC business models, and to some extent a convergence of unit costs. Whereas legacy airlines restructure to reduce costs, LCCs typically see their cost structures increase as they mature—new employees achieve seniority and demand increased compensation, and once-new aircraft require increasing maintenance expenditures. The larger an LCC becomes, the more likely it is to move to some form of hub network, with multiple aircraft types and more traditional product offerings.

LCCs are still growing strong in many parts of the world. Especially in emerging economies like Brazil, India, and China, there remains a tremendous potential for LCC growth and the stimulation of new air travel demand with lower fares. However, actual air travel growth could be constrained by the inability of current and planned airport capacity to handle it. In addition, the possibility of increasing fuel costs makes it more difficult for even an LCC to operate profitably at the extremely low fare levels typically associated with such airlines.

LCCs will continue to play a role in the evolution of the global airline industry, but it is increasingly clear that they will not dominate air travel markets to the extent that some had predicted. The growth potential for LCCs is limited both by the number of point-to-point markets that can support such services and by the reality of increasing costs as LCCs mature. At the same time, existing legacy airlines facing LCC competition will not simply stand by and watch their market shares erode. In the United States and around the world, legacy airlines have responded by adopting some of the business practices of LCCs to match low fares for price-sensitive travelers, while further enhancing their premium services to retain the high-fare price-insensitive market segment.

For airports, there is therefore no single airline business model that is expected to become predominant. Both LCCs and legacy network airlines will continue to compete and coexist at airports. Despite differences in their business models, both groups will have similar concerns. Improved productivity and lower costs will be paramount, and airline managements will continue to scrutinize airport fees and the cost of terminal facilities. The lessons from airport experiments with less elaborate “no frills” terminals are of interest not only to maturing LCC/hybrid airlines but to established network carriers as well. Overall, airports can expect that airlines will be less willing to pay for “signature architect” terminal buildings at airports.

The second most important trend for airline industry evolution is the rapid growth of the emerging global mega-carriers. These airlines from the Middle East as well as Asia have very aggressive growth plans, designed to capitalize on their structural unit cost advantages. Carriers such as Emirates, Etihad, Qatar, and Turkish have already made a mark on many airports and the traffic flows of existing legacy airlines. It is not at all clear that all of these emerging global airlines will be able to realize their ambitious growth plans, given the realities of economic cycles and a limit to the total volume of air travel demand at fares that can cover increasing operating costs. As mentioned, much of their connecting hub traffic will have to come from competition with existing U.S. and especially European legacy airlines.

As they expand their hubs, the mega-carriers are all looking to serve new spoke cities, and the competition between airports hoping to attract them will intensify. Airports will face increasing demands for a variety of fee reductions and/or revenue guarantees from these airlines. Just as LCCs have successfully pitted competing airports serving the same catchment area against each other in the quest for cost concessions, these new global carriers will also promote the traffic benefits to the airport of starting new international spoke-to-hub flights. However, the airport needs of these full-service international airlines are very different from those of LCCs. They require gates that can accommodate large wide-body aircraft, premium-class check-in and lounge facilities, and expanded immigration and customs processing.

Finally, the recent evolution and future prospects for NLCs will have an important impact on both industry structure and airports. Legacy airlines have been forced to achieve cost reductions and productivity gains, with the goal of making them cost competitive with LCCs in domestic markets and with the new global mega-carriers in international markets. However, the financial difficulties of U.S. legacy airlines have prevented them from renewing fleets and investing in their product offerings, whereas European legacy airlines have yet to go through the painful process of restructuring to become more cost competitive.

Looking ahead, the consolidation of U.S. airlines into perhaps three international network carriers and two to three very large domestic LCCs is expected to help them become more competitive on a global scale. European legacy carriers face a more significant set of challenges, with a greater need for cost and productivity improvement and a greater threat to their network flows from the emerging global mega-carriers. The continued participation in and expansion of global alliances and joint ventures is an extension of what is likely to be an inevitable process of consolidation, and these alliances will be important to established NLCs hoping to retain market shares while reducing unit costs in the face of these threats.

For airports, the lesson is that flexibility is critical in virtually every facet of airport systems planning and operations. The absence of a dominant business model or singular driving force in the airline industry, when combined with the tremendous volatility of factors ranging from fuel prices to airline profits and even the future existence of individual carriers, sets the stage for continued and unpredictable change. Furthermore, airlines can alter their business models, fleets, and routes within several months, if necessary to adapt to changing conditions. Airports, on the other hand, routinely must plan for capital investments in facilities that will remain in place for decades. Airports must therefore plan, design, and manage flexibly.

Exercises

2.1. Consider the way A380 and the B787 aircraft are penetrating airline fleets. What are the current total sales of these aircraft? Which airlines are using them? Pick two airlines that use one or the other: on what routes do they fly these aircraft? What do you conclude about their future role?

2.2. Examine the network of one of the Gulf-based carriers using their web site. What are their routes? With what frequencies do they fly these routes? To what extent does their service complete with other hub airports? (Think of it this way: if you were in one of the cities served by their hub, what alternative routes would you have? Through which alternative hubs might you fly?)

2.3. Choose one or two airlines operating at your local airport. Describe their pattern of service for some destinations: at what times and how often do they serve them? What are the scheduled turnaround times? By visiting the airport and observing actual arrivals and departures, determine how closely their actual operations match the schedule. What do you conclude from these observations?

2.4. Choose one or two low-cost airlines serving your airport or region. Describe their history: how have they competed with the established airlines? What are the comparative fares? What routes do they serve? What kind of facilities do they use at a local airport? How do you think these airlines might affect the future facilities and operations at this airport?

2.5. Consider the development of a major airport in your region: what airlines have started to serve this airport over the past decade? How have their requirements impacted the need for and use of airport facilities? To what extent have changes in the airline industry (fare levels, new technology, nature of competition) impacted the development of the airport and its facilities?

References

Air Transport Association (2012) "US Passenger Airline Cost Index." www.airlines.org.

Airbus (2012) "Aircraft List Prices 2012," http://www.airbus.com/presscentre/pressreleases/press-release-detail/detail/new-airbus-aircraft-list-pricesfor-2012/, May 15.

Airline Business (2011) "Scheduled Passenger Traffic Rankings 2010," Annual Surveys, Airline Business Database, http://www.flightglobal.com/staticpages/premiumdata.html.

Belobaba, P., Hernandez, K., Jenkins, J., Powell, R., and Swelbar, W. (2011) "Productivity Trends in the U.S. Passenger Airline Industry 1978–2010," Transportation@MIT, MIT, Cambridge, MA.

Belobaba, P., Odoni, A., and Barnhart, C. (eds.) (2009) *The Global Airline Industry*, John Wiley and Sons, Hoboken, NJ.

International Air Transport Association (2010) "Airline Fuel and Labour Cost Share," IATA Economic Briefing, February, www.iata.org.

Mozdzanowska, A. (2004) "Evaluation of Regional Jet Operating Patterns in the Continental United States," Master of Science Thesis, MIT Department of Aeronautics and Astronautics, Cambridge, MA.

Skaltsas, G. (2011) “Analysis of Airline Schedule Padding on U.S. Domestic Routes,” Master of Science Thesis, MIT Department of Civil and Environmental Engineering, Cambridge, MA.

Tsoukalas, G., Belobaba, P., and Swelbar, W. (2008) “Cost Convergence in the US Airline Industry: An Analysis of Unit Costs 1995–2006,” *Journal of Air Transport Management, 14*, pp. 179–187.

---

[1] Costs and unit costs are inflation adjusted to 2010 dollars.

CHAPTER 3

International Differences

Many aspects of airport planning, design, and management differ substantially across the world. Sometimes these differences represent relative levels of adoption of innovative procedures or new technologies. In significant instances, however, these differences represent deep-seated cultural perspectives. Various cultural contexts, countries or regions, have different norms about relative competencies and obligations of the several stakeholders in airport operations, specifically concerning the following:

- The role of central political power compared to that of the regions
- The permissible and desirable level of participation of private business
- The relative importance of technical experts and managers
- The criteria for excellent performance
- The rights and capabilities of workers

Practices common in one region may be socially unacceptable in another. The variety of norms has two immediate consequences:

- There is no single right answer—the concept of excellence depends on the context.
- And thus, the "best practice" of one region may not be transferable to another.

Airport professionals should recognize that technical solutions may depend on social values.[1] Consequently, global organizations should be careful about how they propose to transfer their practices from one region to another. Complementarily, airport operators need to be careful about how they import "best practices" from other countries. In short, airport professionals need to recognize that plans, designs, and operational practices often embody social and cultural assumptions. Therefore, they need to be sure that their proposals suit the local context and future.

3.1 Introduction

Air transport is a global business with remarkably similar international standards. Airlines use aircraft from a few dominant manufacturers: Airbus, Boeing, and Embraer. Manufacturers design aircraft to virtually the same standards. International airlines carry passengers, baggage, and freight that have similar characteristics. The air transport industry places almost identical requirements on airports and airport managers worldwide.

Two agencies define many of the international requirements for airport design and operation. These are the International Civil Aviation Organization (ICAO) and the United States Federal Aviation Administration (FAA). The ICAO is a United Nations agency that promulgates internationally accepted standards for airports (ICAO, 2004, 2005, 2006). The FAA also sets standards (e.g., FAA, 2001, 2010) and often pioneers the norms that ICAO later follows. See, for example, the discussion of aircraft categories and runway separations in Chap. 10. The FAA has a dominant role because the United States constitutes the largest single market for aviation and has devoted the most money and research to establishing standards. Moreover, because aircraft manufacturers want to sell into the big North American market, they make sure that their aircraft meet the FAA standards.

International standards apply most strictly to matters concerned with the safety of aircraft in the air. International practice is almost identical for all elements of the airport that concern flight: runway markings and lighting, navigation equipment, and zones to be kept clear of obstructions around the airport. National differences concerning safety are small.

However, national differences are great when it comes to land-side features of airport planning, design, and management. Although the air transport industry places almost identical physical loads on airports and airport managers, many nations develop or adopt their own distinctive solutions to these requirements. The loads on the system are similar, but the technical solutions are not. National social values mediate the translation from the technical specification of the problem to the facilities and services that meet this specification. For example, American and European designers meet the requirement to position aircraft at aircraft gates in strikingly different ways, as Sec. 3.2 indicates. In general, the practice of airport planning, design, and management differs considerably among countries.

National differences in airport practice need special attention. Airport professionals might falsely assume that, because the requirements are similar worldwide, the solutions should be also. The truth is otherwise. What is done in different countries, and indeed what should be done to meet local social requirements, often differs greatly from what might appear to be good practice elsewhere. This chapter presents this issue and offers guidelines for airport professionals on how to cope with this reality.

3.2 Some Physical Differences

Some obvious physical differences in the design and operation of airports across the world motivate the discussion in this chapter. These illustrate how some seemingly tangential social assumptions and practices can have important consequences in terms of airport design, cost, and efficiency. They provide tangible evidence of the social construction of technology, the way cultural assumptions shape the seemingly technical solutions to design problems. These examples indicate how this phenomenon occurs regarding governmental and managerial practices. Readers can easily verify these examples visually by looking at different sites.

Check-in Facilities

A passenger approaching a check-in counter in North America will normally encounter an agent standing in an open passageway running between the counter and the parallel baggage conveyor belt. During the check-in process, this agent and others are likely to move up and down the passageway as they sort out issues with colleagues. Finally, somebody will pick up the passenger's bags and place them on the conveyor (Fig. 3.1). This is the normal practice met by about half the airline travelers in the world.

FIGURE 3.1 Typical U.S. check-in arrangement: agents stand and move bags. (*Source*: Ben Mutzabaugh, USA Today.)

In much of Europe and elsewhere, by contrast, passengers checking-in will typically meet an agent sitting down. The agent will not lift the bags. These will move on a small belt between the passenger and the main baggage conveyor belt. These small belts serving each agent cut across the space behind the check-in counters and effectively prevent agents from moving directly to other colleagues along the check-in desks (Fig. 3.2). This arrangement is the alternative standard for check-in facilities.

FIGURE 3.2 Typical European check-in arrangement: agents sit and do not lift bags. (*Source*: Munich Airport.)

Both approaches represent "best practice" in their own context. On the face of it, the North American arrangement is more cost-effective. It requires less equipment and permits agents to move around efficiently as needed. The European pattern, however, conforms to their concept of a humane work environment. Clerks and agents are entitled to sit down on the job, whether they work in airports or super markets. This social norm is widely established and unlikely to change in the near future. What people think of as the best solution is not a purely technical matter; the judgment rests on social and cultural assumptions.

The issue for designers arises when it comes to designing check-in facilities outside North America or Western Europe. Which tradition should they adopt? This is not merely a technical choice; it is a social judgment.

Aircraft Contact Stands

Aircraft “contact” stands are those that close to the passenger buildings. They contrast with the “remote” stands far from the buildings. Passengers normally board aircraft at contact stands through some kind of permanent aerobridge. To get on aircraft parked remotely, they must take some kind of bus.

The design of contact stands and the passenger building differs fundamentally between North America and Europe. Specifically, the connections between the passenger building and the aircraft tend to be very different. In America, the usual arrangement is that aircraft in the contact stands are right next to the passenger building. The aircraft nose may be as close as 10 m, as Fig. 3.3 indicates. In this arrangement, the telescopic, movable aerobridges connect directly between the passenger building and the aircraft.[2]

FIGURE 3.3 Aircraft connecting directly to passenger building. (*Source:* Aeroporto de Congonhas.)

In Europe, however, aircraft at the “contact” stands typically park relatively far from the passenger building. The nose of the aircraft may easily be 25 to 40 m away. The system of aerobridges can correspondingly be up to 70 m long. Airport operators in Europe, and in many countries worldwide, expect that vehicles operating on the airfield will normally not intersect the paths of aircraft. Ground vehicles will circulate on two-lane roads laid out at the face of each passenger building. Moreover, the design will typically provide parking spaces between this roadway and the building. Consequently, aircraft gates in airport buildings in Europe feature both 15- to 20-m bridges for passengers to cross over the road and substantial piers that support these bridges and connect with the movable aerobridges. The view in Fig. 3.4 illustrates this pattern.

FIGURE 3.4 Typical European contact position: long distance between building and aircraft featuring a long bridge over road and parking spaces. (*Source*: London/Heathrow terminal 5.)

The difference in design of the contact stands can have enormous implications for their cost and the efficient use of extremely valuable airfield space. The European design most obviously requires a much greater investment for each gate, because of the cost of the fixed bridge over the roadway and the pier to support it. Moreover, this additional construction is only a fraction of the added expense. The greater cost comes from the inefficient use of space and the extra cost of making the passenger buildings longer. Example 3.1 illustrates the different implications of the two approaches to designing contact stands for aircraft.

What accounts for this expensive difference in practice? Simply put, Americans expect that the drivers of apron vehicles will drive safely and coordinate their movements with the control tower as necessary. The rate of accidents between aircraft and vehicles driving on the apron appears to be close to zero and is not an issue of real concern. European and other airport operators also daily transport thousands of passengers across aircraft taxiways to remote stands, as they do at Paris/de Gaulle, Milan/Linate and Lisbon, and other airports that use remote parking. Yet European and other operators insist on having separate roadways for apron vehicles. Apparently, they do not have confidence in their employees or control systems. In general terms, this major design difference is another example of the way that social attitudes and assumptions shape technology.

Example 3.1 Effect of Different Standards for Aircraft Contact Stands Consider a passenger building with six finger piers, such as has existed at Miami/International and Amsterdam/Schiphol.

The practice of laying out roadways on both sides widens the effective width of each pier by about 40 m, compared to the North American practice. Applying this standard to six finger piers lengthens the complex by about

240 m. This arrangement makes the central building connecting the piers much longer. It increases walking distances. It makes the building more expensive. The cost of passenger buildings can easily be $2000/m^2$. If the main passenger building is 25 m wide and has two floors, the extra cost of the complex due to the 240-m extension is about $24 million in construction costs—to which must be added operating costs, including cleaning, climate control, and maintenance.

An arrangement with roadways in front of the building also limits the capacity of the airfield apron. The width required for each finger pier with roadways is 40 m. Assuming that aircraft are about 70 m long and taxiways are 80 m wide, as for a Boeing 747 or an Airbus 380, the total width for a finger pier with roadways is about 300 m. That is about 15 percent more space per aircraft than for a finger pier without roadways. This corresponding capacity reduction of around 15 percent may be critical at some airports.

Table 3.1 summarizes the specific differences in design practice these cases demonstrate. As these examples suggest, differences in national or regional practices can have significant consequences on both land-side and air-side operations. The examples are tangible instances of a general phenomenon. As the next sections discuss, the social and cultural differences between regions also lead to significant differences in airport planning objectives, procedures, and criteria. These may fundamentally affect the nature of airport operations in different contexts.

Facility	Design Element	United States	Mostly Elsewhere
Check-in facilities	Workstations	No seats: Agents stand, move around, lift bags	Seats provided: Agents sit, do not lift baggage
	Baggage handling	Agents lift bags: Area between check-in counters is unobstructed	Small belts to main belt: Agents cannot move easily along counter area
International departures and arrivals	Passport and customs controls	No exit control. All arrivals must clear border controls	Exit controls. Transit passengers do not enter country
Aircraft contact stands	Aircraft: Distance from passenger building	About 10 m: Essentially right at building, no road along building	25–40 m: Space for a road and parking between aircraft and building
	Apron vehicles: Circulation	Across open apron: No dedicated road	On airside road: Between building and aircraft
	Apron vehicles: Parking	At face of building, or in space around aircraft	Special parking areas, often at face of building

TABLE 3.1 Some Physical Differences in Design of Airport Passenger Buildings between United States and Other Traditions

3.3 Some Useful Distinctions

Countries and regions differ. Their ingrained habits and concepts are not the same. Some of these distinctions have significant practical implications for the planning, design, and operation of airports. This section presents the dimensions of distinction that seem to drive the most important consequences.

The important dimensions that characterize national differences are not necessarily permanent. They reflect patterns of thinking and behavior that people have learned. They may thus change over time as the result of either evolutionary or cataclysmic events. In Britain, for example, government became less centralized in the last quarter of the twentieth century, as regions such as Scotland and Wales developed their own legislatures. In some cases, a transition may be abrupt. The highly centralized Soviet Union, for example, rapidly dissolved into the Confederation of Independent States and a multitude of autonomous regions. Despite the possibility of change, however, national characteristics are deep-seated. Airport managers can assume for working purposes that local social assumptions and patterns of behavior are a fact of life.

Experts in comparative government and politics agree that there are important differences in national attitudes and values. They also know that it is difficult to describe the full complexity of national patterns satisfactorily. The literature on the topic is controversial. Specialists offer alternative, often-conflicting interpretations.[3] In this context, the following discussion modestly suggests some of the important considerations. Its purpose is to alert airport practitioners to important national considerations to take into account and to stimulate them to think about how these considerations affect their practice.

For airport professionals, the important national differences are those that affect the environment in which they operate. These differences concern who, what, and how things are done. Specifically, airport operators need to understand the following:

- Who makes decisions—central authorities or pluralistic stakeholders?
- What is the decision-making process—is it directed autocratically or negotiated among interest groups? Who has the right to participate in this process?
- Which values and goals are most important—economic benefits? Social values? Regional prestige?
- What are the criteria for excellence—high profits? Good service? Beauty?

Two dimensions of national differences seem most important for airport planning, design, and management. These dimensions usefully define most of the answers to the questions of who, what, and why. These concern the diversity in the decision-making process and performance criteria.

Diversity in the decision-making process reflects the number and types of stakeholders who strongly influence decisions. In some contexts there is effectively no diversity. This is the case when central authorities or personalities are the final arbiters. In other contexts there is great diversity, as multiple levels of government and numerous stakeholders negotiate resolutions to any issue. Countries also differ in the kinds of goals they promote and the criteria they apply. In some cases, decision makers define goals quite specifically and numerically, either in technical or in economic terms. In other cases there never is any clear definition of goals or objectives.

These two dimensions correlate with each other to some extent. Centralized, directive governments have the ability to impose criteria on the decision-making process. Pluralistic decision-making processes that negotiate developments will not be able to maintain, let alone impose, consistent numerical criteria of performance. Centralized decision-making processes are therefore more likely to be able to impose performance criteria—although they do not have to do so.

National Differences in Diversity of Decision Making

Countries that have had salient roles in the development of airport systems differ greatly in the way they expect decisions to be made. Several have strong traditions of central direction and control. Others are pluralistic and feature decentralized decision making.

In the United States, decisions about airports are highly decentralized. The central national institutions have little influence on specific designs—surprisingly so for persons from outside North America.[4] Under the Constitution of the United States, the power to make most major decisions—those concerning airports in particular—is in the hands of the states. Moreover, the state constitutions frequently leave decisions about such matters to local communities. Most frequently, local airport authorities and cities are responsible for developing plans and airport proposals. The U.S. FAA can support, encourage, and confirm local decisions but cannot impose its will.

In the United States, all major stakeholders are entitled and expected to have an active voice in decisions about airports. Airlines, for example, frequently operate their own passenger buildings. They also participate actively in the design of these facilities. As Chap. 7 indicates, airline tenants at many U.S. airports are guarantors of the revenue bonds and thus effectively have veto power over major investments on airports. Airlines in the United States often control what is built, how it is designed, and when it is implemented. Additionally, local communities and interest groups expect to participate actively in the decision-making process. Local stakeholders concerned with the airport may have specific rights to intervene. For example, the board of directors for Massport, the independent state agency responsible for operating Boston/Logan airport, by law includes representatives of local communities, citizens groups, and labor unions. Decisions about airport planning, design, and management in the United States are negotiated among the many stakeholders. These generalizations about the United States have exceptions, in view of the enormous diversity among the 50 states. However, diversity in the decision-making process is a fact in the United States.

To illustrate the decentralization and diversity of authority on airport activities in the United States, consider Massport's proposal to develop a new short runway for Boston/Logan airport. Airport planners in the FAA widely encouraged this proposal to add capacity. Nominally, the FAA administrator has the authority to approve this plan. In practice, however, elected officials from the area effectively have the power to block such plans. Local members of Congress have done so for many years, by threatening to block portions of the FAA's budget. In the United States, these kinds of planning issues are resolved through intense negotiations between various local authorities, the airlines, the FAA, as well as numerous advocacy groups represented by lobbyists. The tradition in the United States is that essentially all stakeholders in an issue have the right to participate in their resolution.

France, by contrast, has a history of central direction. The central government announces decisions and implements them. The public elects the government, of course. The government also pays attention to public needs, and has prepared and is beginning to implement mechanisms for compensation and remediation of damages to people and the environment (Faburel, 2001). However, the public has been neither expected nor is entitled to participate in the decision-making process itself (Block, 1975). Thus, in the 1970s, the French government located the Paris/de Gaulle airport, established development zones, built the facility, and directed specific airlines to relocate from the other airport, Paris/Orly. Around 2000 they replicated the process to develop two new parallel runways at Paris/de Gaulle. These developments went forward without significant public hearings or effective protest. French authorities expect to be able to act decisively in the best interests of the public.

Historically, many countries have had traditions of centralized national power, both overall and specifically regarding airport planning. Typically, countries have had national ministries responsible for the design, construction, and operation of airports. Since the turn of the twenty-first century, however, the degree of centralization of airport planning and management in major aviation markets has lessened considerably. Australia and Canada virtually eliminated their federal airport agencies. These countries devolved responsibilities for airports to companies and local authorities in the traditionally autonomous states and provinces. Britain transformed its governmental British Airports Authority into a company and made local airports—such as Manchester and Birmingham—operate as independent companies. Mexico devolved power from the central government to independent regional companies. This evolution has created more autonomous airport authorities and increased the diversity of airport operators.

Germany, Switzerland, and Italy have traditionally been decentralized. Germany is a federal system of Länder. Thus independent groups own and operate the major German airports (Berlin, Frankfurt, Munich, Düsseldorf, Hannover, etc.). Switzerland is also a federation of notably autonomous cantons. Independent companies operate the Zürich and Geneva airports. The situation is comparable in Italy.

Outside the United States, the increased diversity in airport operators has not translated into increased diversity in the local decision-making processes. The power to plan and design airport facilities is still typically in the hands of the airport operator. Airlines typically have little say in the definition of airport investments. For example, British Airways had essentially no part in the design of the $7 billion Terminal 5 at London/Heathrow that was designed for it. Likewise, in the mid-1990s the Frankfurt/International airport designed and built a billon-dollar passenger building for the German national airline, totally unsuited to its hubbing operation. In that case, however, Lufthansa managed not to occupy the building. Local constituencies likewise generally do not have a deciding role in airport decision making. Environmental and other groups may be heard or consulted, but they do not decide.

Britain has established the rule that an extensive public inquiry must be held for important issues. The investigation into the construction of the second runway for Manchester took about 5 years. The inquiry into the T5 passenger building at London/Heathrow took longer and reputedly cost over £81 million (about $125 million) (Thorpe, 2001). This practice is totally different from the procedures in France. In Britain, the interest groups have the right to express themselves and delay planning. Legally, however, they have no power. The Minister of State for the central government decides such issues authoritatively.

A quote from the British political philosopher Edmund Burke illustrates the fundamental differences in perspective on decision-making processes between countries. In this case, he was contrasting the centralized, unitary British view with the pluralistic, negotiated practice of the United States.

> Parliament is not a congress of ambassadors from different and hostile interests… Parliament is a deliberative assembly of one nation, with one interest, that of the whole, where not local purposes, not local prejudices ought to guide, but the general good… (Burke, 1774)

Overt negotiations among stakeholders to determine airport development are rare outside North America. In many contexts, such negotiations would be taboo. A common sentiment is that the duty of government is to govern, and if they cannot do so, they should resign. The case involving the second parallel runway at Tokyo/Narita illustrates the point. In this situation, several farming families did not wish to sell their land to make way for the construction of this facility. For more than 40 years, a handful of people prevented the completion of a major addition to a significant national asset, yet the national, societal, and political conventions prevented the authorities from negotiating or adjudicating any kind of compromise that would allow the nation to proceed.

Table 3.2 summarizes these national differences in assumptions about who gets to decide how airports should be planned, designed, and managed. As Sec. 3.4 indicates, these dissimilar perspectives can influence airport development fundamentally.

		Diversity Among Airport Operators		
		Little	**Some**	**Extensive**
Diversity in decision-making process	Airports decide	France, Japan	Australia, Mexico	Germany, Italy
	Stakeholders negotiate			United States

TABLE 3.2 Some National Differences in Diversity of Decision Making and Operators

National Differences in Performance Criteria

Countries that have led the development of airport systems also differ greatly in the way they define the objectives for airport planning. In some contexts, the objectives are loosely defined. In others, they may be quite specific. Because performance criteria shape the products of design, these differences have significant consequences for how nations develop airports.

The nature of the performance criteria depends on who defines them. It is therefore relevant to look at the kind of people who run airport planning agencies and operators. The differences between countries can be striking. Some countries recruit elite engineers into careers of management of public works and airports in particular. Other countries prefer generalists or economists rather than specialist engineers. Some countries have no particular pattern at all. These national patterns mark the practice of airport planning and management for these countries.

In the United States, there is no visible career pattern for the recruitment of airport executives. Leaders in the field are lawyers, managers, engineers, former military officers, and other professionals. They tend to enter airport planning from some other industry. Typically, they have established themselves in a related field, been working for one of the stakeholders in the airport business, and then become involved in airport planning. Recent leaders of Massport, the operator of Boston/Logan airport, have included a lawyer who had been a special assistant to the mayor of Boston, a former local Congressman, an activist for local groups concerned about noise, and an aeronautical engineer who had become a prominent entrepreneur. Such people bring a wide range of perspectives and norms for good performance to airport planning.

Broad performance criteria define airport planning in the United States. These emerge from distinct negotiations among interested parties. The FAA publishes standards for the airside of the airport, based on their mission to promote safety. These appear in their Advisory Circulars and are readily available on the web and in print. However, the FAA does not establish these norms by itself. It works them out through close discussions with industry groups such as the Airlines for America (former Air Transport Association of airlines), the Airports Council International (airports), the General Aviation Manufacturers Association, the National Association of State Airport Operators, and so on.

Criteria for economic performance of airports in the United States come from somewhere else entirely. These emerge from the groups that supply the funds, notably the airlines that pay the fees and the investment bankers that loan the money.[5] These standards are informal and negotiable. The consensus is loose but has important implications that imprint a distinctive mark on airport planning and management in the United States. Briefly stated, it is that American airports should be

- Operated as businesses with transparent public accounts (to reassure investors and guarantee repayment of the loans)
- Run as a public service and that are not supposed to make profits beyond what is needed to maintain the business
- Charging airlines fees as low as possible consistent with good service and attractive facilities.

France, by contrast, recruits its leaders for airport planning and management from its most talented engineers. Specifically, it usually obtained the future leaders for its national airports company, the Aéroports de Paris, largely from its most selective national engineering school, the École Polytechnique. It inducts about 30 of its best graduates each year and places them in a quasimilitary organization, the Corps des Ponts et Chaussées (of Bridges and Roads). These persons all share the same background, the same analytic and engineering approach, the same esprit de corps (see Suleiman, 1974). Similarly, it recruits lower-level engineers from less demanding national schools, such as that of the Travaux Publics de l'État (State Public Works). As can be expected, these professionals establish analytic and precise performance criteria. Although some say this tradition is in decline, its legacy persists.

Britain traditionally prefers generalists to specialists. A common view is that specialists become too involved in their field and cannot be trusted to have a sufficiently broad national perspective. British education at the elite universities such as Oxford and Cambridge stresses liberal subjects such as political economy, classical literature, and history. The British government has likewise recruited its future leaders from among such people (see Hennessy, 1989). These professionals take a broad, pragmatic view of decision making.

An anecdote captures the difference in approach between the technical and generalist approach perspective on airport planning. It involves British and French airport managers when the British Airport Authority (BAA) introduced peak-hour pricing at their London airports. This practice charges higher prices during peak hours. It thus reduces the peak demands, and the capacity and capital expenditures that the airport operator has to provide. It is an important means of increasing economic efficiency (see Chap. 12). It is theoretically possible to define the best peak-hour prices analytically from the equation of the demand for services. The BAA recognized, however, that the complexity of airline operations made any estimation of the demand speculative. Pragmatically, it chose to introduce a flat charge on each peak-hour operation. They intended to establish the principle of the charge, and then to raise or lower the peak-hour price until it achieved the intended effect. At this point, a French team came to London to learn from the British experience. They asked to see the equations and calculations the BAA used to determine the charge. The BAA said they could not present these, as they did not have any. When the French left the BAA, they exploded in anger at the "uncooperative, untrustworthy British," who refused to share with them. Despite the author's attempts to explain the situation, they would not believe that the British had such a different perspective. Being engineers steeped in equations, they did not appreciate the possibility of such a different outlook. Yet such differences are real and do complicate the international understanding of airport planning.

British authorities have established detailed performance criteria for airports. As a result of their privatization of airports, they needed to regulate airport companies to protect the public against monopoly pricing and excessive charges. Because companies can abuse a monopoly position by lowering standards of service, it is not sufficient to regulate prices. The British regulatory authorities therefore established complex performance criteria in many different areas (see Graham, 2005). These standards have constrained U.K. airport operators to emphasize the services specified in the regulatory criteria and have shaped the way British airport operators do business.

The important aspect to retain is that different national traditions emphasize distinct aspects of performance. The criteria prevalent in the United States are pragmatic and value the distinct interests of the important consumer groups, such as airlines and passengers. The French and Japanese traditions give more weight to technical factors. The British approach typically favors economy (Table 3.3). These perspectives strongly influence how airport planners and managers from these traditions build and operate airport systems, as the next section indicates. World travelers will recognize the differences from experience.

		Diversity among Performance Criteria		
		Economic	**Mixed**	**Technical**
Diversity among decision makers	Career specialist			France, Japan
	Widely recruited generalists	United Kingdom	United States	

TABLE 3.3 Some National Differences in Decision Makers and Performance Criteria

3.4 Implications for Practice

National differences in concepts of decision making about public projects and of performance criteria lead to significant general and practical implications. In the increasingly global practice of airport planning, these effects are becoming more significant for practitioners. European airport companies operating in the Americas, for example, need to think carefully about how they might modify their practice to suit the local context. North American consultants advising on airport development and operations overseas likewise need to tailor their suggestions to local realities.

General Implications

There is no single right answer. This is the fundamental takeaway for the reader of this chapter. Since the concept of excellence depends on the context, the best practice in one region may be impractical or otherwise unsuitable in another.

Although analysts may agree on the operational characteristics of different designs, when they disagree on the relative importance of these factors, they may not agree on which is best. For example, it is clear that the North American design for check-in facilities requires less capital investment, and that the European practice makes work easier for the check-in agent. Which design is better depends on how the airport operator and the local society value these features.

The related implication is that "best practices" of one region may not be readily transferable to another. "Best practices" in some places may appear to be poor or unsuitable elsewhere. Even when foreign best practices appear superior to local procedures, they may be sufficiently counter to national norms to make them impractical. For example, it would be difficult to introduce the standard U.S. design for check-in facilities in France or Britain, even if an airline or airport wanted to do so. Such a change would require extensive negotiations with workers and changes in work rules.

Finally, national differences in performance criteria limit the usefulness of international "benchmarking" of airports. "Benchmarking" is the practice of comparing performance at various sites in specific industries. The objective is to identify the sites that perform best in various categories. These then becomes a benchmarks, that is, standards for the rest of the industry. Benchmarking can identify sites that perform better than others overall and that might be taken as models. It can also identify sites that perform poorly overall, and that might need management attention. However, these individual measures of performance are difficult to translate into any internationally meaningful overall measure of performance. Any weighting of the categories represents notions of relative value that will not represent the priorities of all countries.

The fact that technical solutions depend on social values means that global organizations should be careful about how they transfer practices from one region to another. The experience of Amsterdam Airport Schiphol (AAS) illustrates the issue. In the late 1990s, AAS committed to design, construct, and operate an international passenger building at New York/Kennedy. They brought with them their excellent reputation and expertise from running an attractive facility at Amsterdam. They proceeded to design the New York facility along the same lines as Amsterdam. In particular, they planned to cater to the variety of foreign airlines that each had a relatively small presence at New York/Kennedy. However, AAS apparently did not understand the power of airlines in the United States to make their own arrangements. In Europe, airlines rarely have much influence on the design of passenger buildings. In the case of New York/Kennedy, airlines disrupted the plans of AAS in two ways. First, significant airlines left the International facility, either to a new building they built themselves (as Air France and its associates did in Terminal One) or to the buildings operated by their alliance partners, such as American Airlines. This phase left AAS with a big investment and insufficient tenants. Second, Delta Air Lines then agreed to take over much of the AAS building, provided it was redesigned. Delta wanted a standard U.S. configuration, favoring transfer operations and placing commercial activities near the departure gates, beyond security—exactly opposite to normal practice in Amsterdam. In short, AAS suffered when they tried to transfer excellent Dutch practice to New York. The case illustrates how the dependence of technical solutions on social values means that airport operators need to be careful how they import "best practices" from other countries.

Specific Implications

The differences in decision-making processes and criteria of performance translate into specific differences in how airport operators develop their facilities different countries. These concern the following:

- Artifacts—what they construct
- Type of service—the features they stress
- Operations—how they manage their properties

This section mentions some salient examples. Later chapters explore details.

In the United States, stakeholders in airport operations participate extensively in the decision-making process. The result is that the design of the airport reflects their concerns. For example, airlines like to minimize the time their aircraft have to taxi. As Chap. 14 explains, efficient designs can save the airlines hundreds of millions of dollars a year. Therefore, when airlines have a strong voice in the design of airports, as they do in the United States, they insist on designs that facilitate easy movement. These stagger the runways, so that landings end and takeoffs start near the passenger buildings. They also pave over large areas and thus eliminate restrictive taxiways that require aircraft to make many turns. For example, U.S. airport operators typically pave over the entire space between finger piers. Elsewhere, at Amsterdam/Schiphol, for example, large portions of this space may be left unpaved or are set aside for lights and is otherwise unavailable for aircraft maneuvers. The comparison of Atlanta and Kuala Lumpur/International illustrates this phenomenon. Both airports feature parallel runways on either side of passenger buildings. However, the paths the aircraft follow are much more direct and operationally less expensive at Atlanta (Figs. 3.5 and 3.6).

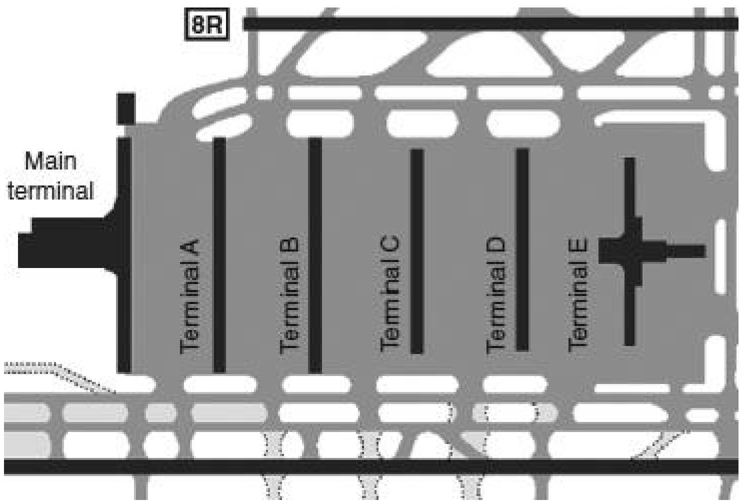


FIGURE 3.5 Typical U.S. apron layout in front of midfield passenger building: aircraft can access gate with a minimum of turns. (*Source*: Atlanta/Hartsfield Airport.)

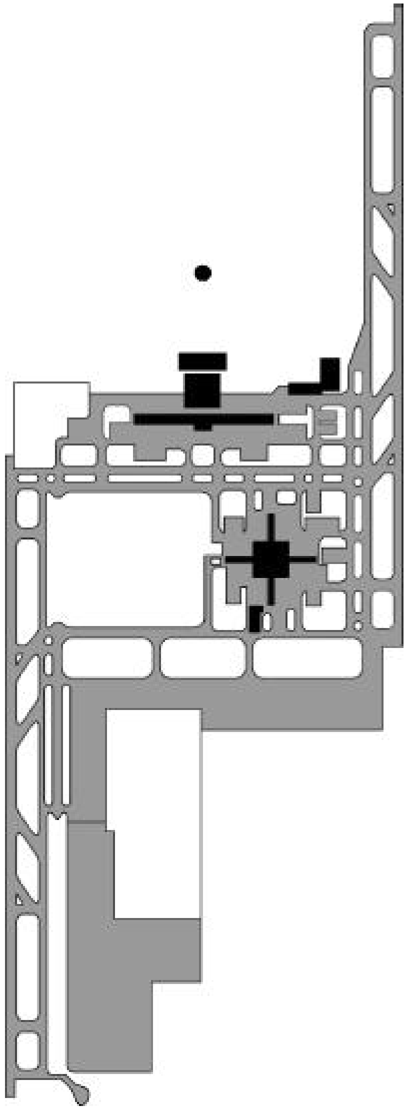


FIGURE 3.6 Taxiway layouts at Kuala Lumpur/International: aircraft require many turns to access gates. (*Source*: AeroStratos Pte Ltd, Singapore.)

Similarly, airport operators in the United States tend to cater to the individual desires of passengers. Specifically, they provide extensive parking facilities at affordable prices and promote easy access for automobiles. This conveniently enables individuals to proceed from home to airport directly door to door. In countries with centralized decision making, however, airport operators favor collective means of airport access (see Coogan, 2008).

They channel travelers into patterns that require combinations of travel by taxis and trains, which are inherently less convenient for individuals although they may be beneficial to the area as a whole. See Chap. 17 on airport access.

National differences in the concepts of excellence also influence the types of service airports offer. The French emphasis on technical excellence, for example, leads them to develop state-of-the-art innovations. The way they have integrated their high-speed rail system, the TGV, into Paris/de Gaulle and Lyon airports illustrates this phenomenon. Moreover, the existence of the TGV itself demonstrates the power of the central government to impose technical excellence for the national cause of public transport. The managing technical elite considers the airports to be an opportunity to develop and showcase all kinds of innovations. These have included unique developments such as variable-speed moving sidewalks, baggage belts lifting vertically through several stories, and check-in facilities after passenger control. They deliberately seek to place themselves in the role of technological leaders.

In Britain, the emphasis is on economy and return on investment. Naturally, this leads to less service and elegance. A popular British author described Terminal 4 at London/Heathrow in the following terms:

> Long, slow-moving lines stretch from the check-in desks nearly to the opposite wall of the concourse, crosshatched by two longer lines converging upon the narrow gate that leads to Passport Control, the Security gates, and the Departures Lounge. The queuing passengers shift their weight from one foot to another, or lean on the handles of their heaped baggage trolleys, or squat on the suitcases…[He looks] up at the low, steel-gray ceiling, where all the buildings' ducts and conduits are exposed…which makes [him] feel as if he is working in a hotel basement or the engine-room of a battle-ship. (Lodge, 1992, p. 3)

Centralized decision making also leads to operational procedures quite different from those prevailing in regions where decision-making power is distributed. In this respect, practice in the North American half of the airports market contrasts with that in the rest of the world. In Europe, Japan, and elsewhere, planning processes are directive and indicate what will happen. In the United States, on the other hand, plans are merely suggestive, as Chap. 4 indicates. As Chap. 12 indicates, operators of busy European and Asian airports typically manage their airspace through formal allocations of the "slots" for aircraft arrivals and departures. They also frequently charge high minimum fees on aircraft operations to discourage or effectively ban smaller aircraft from the congested airports. Such procedures are rare in North America. The airlines and operators of small aircraft have rights to operate pretty much when they choose, just as drivers are free to get in their cars and drive.

The pluralistic nature of the United States is evident throughout the operation of the airport itself. In North America, it is usual to have dozens of independent contractors managing various bits of the airport. Airlines typically handle their own baggage and check-in operations, often even their own passenger buildings. Competitive national corporations routinely manage the parking facilities, often several at the same airport. Independent contractors usually do the cleaning and operate security. Architecture and engineering firms carry out the design and construction management for the airport. In practice, most U.S. airports are highly privatized in that private companies run most of their operations. The situation has been vastly different in the rest of the world. The pattern elsewhere has been that the airport operator has provided all services. In Europe, antimonopoly directives now require airports in the European Community to permit competitive services, but the major airports typically offer and provide the whole range of operational services. A comparison between Boston/Logan and Frankfurt/International illustrates the difference. Both airports are about the same size and have around 15,000 to 17,000 workers on the airport. At Frankfurt/International, most of these employees work for the airport operator, whereas at Boston/Logan, only about 800 work for Massport. Table 3.4 summarizes the range of these particular distinctions.

Area of Practice	Common practice in	
	North America	Rest of World
Facility construction	Generous airfield paving to facilitate aircraft ground operations	Restricted amount of paving for taxiways and aircraft aprons
	Emphasis on private cars, automobile access, parking	Emphasis on collective transportation, rail access
Planning	Suggestive	Directive
Operations	Airlines usually schedule freely as they wish	Airports allocate landing and takeoff slots
	No discriminatory pricing; all users have access	Peak-hour pricing, small aircraft often excluded
	Airport operator has small staff; most services contracted out	Airport operator is a big employer; airport offers most services

TABLE 3.4 Some Distinctions in Airport Planning and Design between North America and the Rest of the World

Exercises

3.1. For some airport of interest to you, use the web and other references to identify the managers and their professional backgrounds. What kind of professional formation do they share, if any? How would you characterize this group? What does this imply for their decision making? If time allows, repeat this process for a foreign airport and compare the results.

3.2. Examine the decision-making process concerning the development of a major runway in the United States and Britain. Compare the process for the possible third runway for London/Heathrow with that of an American example, such as Boston/Logan, Chicago/O'Hare, Miami/International, San Francisco/International, or St. Louis/Lambert. Who do you think made the important decisions in these cases? What was the power of the several stakeholders? What do you conclude about airport planning processes in the United States?

3.3. Repeat the previous exercise for the development of a major new passenger building in the United States, for example, the International Buildings at New York/Kennedy or San Francisco/International, or the American Airlines buildings at Miami/International or Chicago/O'Hare. If time and interest allows, compare your conclusions with the results of the previous exercise.

References

Bijker, W., Hughes, T. P., and Pinch, T. (1987) *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, MIT Press, Cambridge, MA.

Block, J. (1975) "Planning the European Environs—A European Viewpoint," *Proceedings ASCE Conference on International Air Transportation*, San Francisco, March, pp. 191–204.

Burke, E. (1774) "Speech to the Electors of Bristol," in *Works of the Right Honorable Edmund Burke*, vol. 1, Henry Bohn, London, U.K., 1854, pp. 446–448. Also available in *The Founders' Constitution*, P. Kurland and R. Lerner (eds.), 1(13), University of Chicago Press, Chicago, IL.

Coogan, M. (2008) *Ground Access to Major Airports by Public Transportation*, Airport Cooperative Research Program Report 4, Transportation Research Board, Washington, DC.

FAA, Federal Aviation Administration (2001) *Application for U.S. Airworthiness Certificate*, Advisory Circular 21-12b, U.S. Government Printing Office, Washington, DC.

FAA, Federal Aviation Administration (2010) *Standards for Airport Markings*, Advisory Circular 150/5340-1K, U.S. Government Printing Office, Washington, DC.

Faburel, G. (2001) *Le bruit des avions—evaluation du coût social*, Presses de l'École Nationale des Ponts et Chaussées, Paris.

Graham, A. (2005) "Airport Benchmarking: A Review of the Current Situation," *Benchmarking: An International Journal, 12*(2) pp. 99–111.

Hennessy, P. (1989) *Whitehall*, Secker and Warburg, London, U.K.

ICAO, International Civil Aviation Organization (2004) *Aerodrome Design Manual, Part 3, Pavements*, Doc 9157P3, 3d ed., Montreal, Canada.

ICAO, International Civil Aviation Organization (2004) *Aerodrome Design Manual, Part 4, Visual Aids*, Doc 9157P4, 4th ed., Montreal, Canada.

ICAO, International Civil Aviation Organization (2005) *Aerodrome Design Manual, Part 2, Taxiways, Aprons and Holding Bays*, Doc 9157P2, 4th ed., Montreal, Canada.

ICAO, International Civil Aviation Organization (2006) *Aerodrome Design Manual, Part 1, Runways*, Doc 9157, 3d ed., Montreal, Canada.

Lodge, D. (1992) *Paradise Lost*, Penguin Books, New York.

Suleiman, E. (1974*) Politics, Power and Bureaucracy in France: The Administrative Elite*, Princeton University Press, Princeton, NJ.

Thorpe, K. (2001) "The Heathrow Terminal 5 Inquiry: An Inquiry Secretary's Perspective," http://www.planning-inspectorate.gov.uk/pi_journal/thorpe _art2.htm.

van der Horst, H. (1996) *The Low Sky: Understanding the Dutch*, Scriptum Books, Schiedan, the Netherlands.

van Wolferen, K. (1989) *The Enigma of Japanese Power: People and Politics in a Stateless Nation*, Knopf, New York.

---

[1]Social scientists call this the "social construction of technology." Bijker et al., 1987.

[2]Exceptions to this rule exist. For example, the Terminal E at Boston/Logan has a road along its face, as do the midfield concourses at Denver/International. As for much of airport planning, design, and management in North America, local airport operators can and do adopt their own practices.

[3]Readers wanting to explore these issues may want to start with some of the following texts: Hennessy (1989)—Britain; Suleiman (1974)—France; van Wolferen (1989)—Japan; van der Horst (1996)—the Netherlands.

[4]To illustrate the force of this pervasive aspect of political life in the United States, consider the question of education. Although the United States has a national Department of Education (corresponding to a Ministry of Education in other countries), this institution has almost no impact on how schools are run or on curricula. By tradition in the United States, local communities run and almost entirely pay for schools.

[5]To put the role of bankers and airlines in perspective, consider the construction of Denver/International. In this case, bankers raised about $3.2 billion in loans, secured mostly by airlines. Thus, United Airlines agreed to pay about $200 million/year for its midfield concourse. The FAA paid only about $800 million of the total capital cost of the new airport.

PART II

Systems Planning Design and Management

CHAPTER 4

Dynamic Strategic Planning

Dynamic strategic planning is the approach recommended for airport development. It recognizes the reality that the airport/aviation industry is highly uncertain; that we do not and cannot know what the future will bring. Airport planners, designers, and managers therefore need to consider many different possibilities. Dynamic strategic planning enables airport professionals to think through these contingencies. It leads to a flexible development strategy that positions airports to minimize risks, take advantage of opportunities as they arise, and thus maximize expected value.

Dynamic strategic planning adapts the traditional process of master planning to the current era. Conventional master planning is now inadequate. It became the standard approach over a generation ago, when governments both strongly regulated the airport/airline industry and controlled most of the airlines of the world. In that distant past, things changed slowly: new airlines were infrequent, low-cost airlines were rare, route patterns were stable, and airports operated in a known environment. That was then. Now, the airport/airline industry is constantly changing (Chap. 1). Airport planning needs to keep up with this evolution. Dynamic strategic planning is now appropriate.

The forecast is "always wrong." Managers and planners must face this fundamental reality in this era of innovation and competition in the airport/airline industry. Airlines unexpectedly form alliances, merge, and change their routes and services; passengers and shippers reorient their patterns. Such variations make forecasts of levels and types of traffic unreliable. Airport professionals must assume that the future reality can easily be different from what seems most likely at present.

Dynamic strategic planning leads planners to anticipate the range of possible futures and scenarios of operation—instead of merely a single forecast. It then analyzes how alternative developments would perform under the several scenarios. This information gives decision makers the reasonable basis for selecting the initial developments that lead to the preferred profile of risks and benefits. Dynamic strategic planning positions the airport to maximize its expected future performance by taking advantage of good opportunities and avoiding unnecessary developments. Overall, it builds appropriate flexibility into the design to facilitate smooth, effective transitions to new situations.

Doing dynamic strategic planning is like playing chess well. One first thinks ahead many moves. Then one makes an initial move to establish a position that enables a good response to threats and opportunities that might arise. As the situation advances, one rethinks the moves and proceeds as then seems appropriate, based on the reality of what is actually happening rather than one's original speculation of what might happen. The game plan alters move by move, period by period. As applied to airports, this means that one develops facilities with the flexibility to expand or change functions as seems best, adjusting developments period by period according to how the future unfolds. Sections 4.4 and 4.5 show how to plan for uncertainty.

4.1 Planning Concepts

The concept of planning needs explanation. It means different things in different contexts, to planning professionals and to airport planners in particular. Specific words and phrases, such as "plan," "master planning," and "strategic planning," have acquired meanings that are not obvious. Persons who have not been intimately involved in these practices or are not aware of local differences may get confused. It is therefore useful to identify the meaning of the several words for planning in the context of airport systems.

Plans

Professionals from different contexts do not share a common understanding about what the concept of planning implies. All agree that planning involves the preparation of a response to some possible future events and that a plan is a conceptual roadmap of what one could do. They disagree, however, between two contrasting perspectives:

- Is a plan a directive blueprint from top authorities that specifies what is to happen?
- Or, is a plan a collection of local suggestions of what airports might like to do, all of which are debatable or negotiable?

In many contexts, planning is a top-down, directive activity. Elite groups, typically government officials, prepare plans for important sectors of the economy, such as airports. They then transmit these plans to subordinates for development. This process prevails not only in autocratic but also in democratic countries. In Japan, for example, the responsible national ministry systematically identified and developed a sequence of major national projects, such as the island airports of Hiroshima, Osaka/Kansai, and Nagoya/Chubu, and of regional airports for each prefecture.

In the United States and some other countries, planning is a bottom-up, visionary activity. Local authorities prepare their own plans. This practice is common in countries that have strong regional governments, such as the provinces in Canada, the Länder in Germany, and the states and cities of the United States. In the United States, for example, local airports prepare lists of possible projects according to what they see as best for them, without consulting other airports and often in direct competition with them. Every 2 years the U.S. Federal Aviation Administration (FAA) collects these uncoordinated local plans and presents a revised National Plan of Integrated Airport Systems (NPIAS). This document is far from directive:

> Because the NPIAS is an aggregation of airport capital projects identified through the local planning process, rather than a spending plan, *no attempt is made to prioritize the projects the included development or evaluate whether the benefits of a specific development project would exceed its costs*. [Italics added] (U.S. Secretary of Transportation, 2010)

Such bottom-up "plans" are in no sense guides as to what will happen, and certainly do not dictate any specific allocation of money. These wishful local plans are very different from directive national plans. Readers should keep this difference in mind whenever they read or listen to international colleagues.

Master Plans

Master plans have a very specific meaning in the context of airport planning. As stated by the International Civil Aviation Organization (ICAO):

> An airport master plan presents the planner's conception of the *ultimate development of a specific airport*. [Italics added] (ICAO, 1987, pp. 1–2)

This definition is widely accepted internationally. ICAO is part of the United Nations, and representatives of the member states developed and agreed to it.

A master plan focuses on an architectural/engineering development at a single airport. Note that it involves three essential notions:

- Ultimate vision, that is, a current view of the possible long-term future, for example 20 years ahead
- Development, that is, the buildings, runways, and other physical facilities—not operational concepts or management issues
- Specific airports, not to a regional or national aviation system

The master plan is thus tightly constricted compared to national plans that governments have prepared and implemented in India, Japan, and elsewhere.

Traditional practice develops airport master plans in a linear process. The ICAO, the International Air Transport Association (IATA, an airline group), and the U.S. FAA provide the most commonly used guidelines (FAA, 2004 and 2005; IATA, 2004; ICAO, 1987). They cover both master plans for individual airports, and "integrated airport system planning." These several guidelines are fundamentally the same, although they differ in detail.

The key elements of this process are the following:

- Inventorying existing conditions
- Forecasting future traffic
- Determining facility requirements
- Developing several alternatives for comparative analysis
- Selecting the most acceptable alternative as the master plan

This master planning process is fundamentally flawed. It assumes that planners should only consider a single forecast. This is both unrealistic and irresponsible:

- It is unrealistic in this era of innovation and competition. The future is full of surprises, as experienced airport professionals know well from experience. Sec. 4.3 makes this point in detail. Traffic can develop in many ways.
- It is irresponsible because, by focusing on the most likely or preferred forecast, it neglects risks. It does not provide appropriate insurance against them. It is as if a business based its planning on the most likely forecast that its facilities would not suffer a fire, and consequently neither provided fire suppression systems nor bought insurance.

Master plans rapidly become obsolete. Airport operators frequently have to junk the ultimate, 20-year vision of the master plan after only a few years. Sometimes it is "dead on arrival" due to its inflexibility. Not too long ago, for example, the board of directors of one of the top airports in the United States voted to "accept" a master plan that had been 5 years in the making (they had to do this, so that they could legally pay the consultants). Then, as the next item of business, this same board voted a contract for a new planning process, because they already knew the approved master plan was out of date!

Despite the deficiencies of these documents, airport operators will continue to have to prepare master plans. This is because the national and international funding agencies expect to see these kinds of plans. In the United States, for example, the general rule is that airports can only get funds from the federal government for projects that are in the NPIAS.

Furthermore, projects only get into the NPIAS if they are included in an approved master plan.

The challenge for airport planners is to improve the master planning process, so that it can deal realistically and responsibly with the future. In principle, this is not difficult, because it is easy to modify the process from a technical point of view. In practice, old habits die hard, and it may take time for standard processes to evolve. Meanwhile, forward-thinking airport operators should be able to implement better planning procedures, such as dynamic strategic planning.

Strategic Plans

In the field of management, *strategic planning* refers to a disciplined process for analyzing the current situation of a business activity, and identifying the vision of how that entity should position itself regarding its customers and competitors (e.g., Porter 1985; Hax and Majluf, 1996). This process has fallen out of favor (Mintzberg, 1994; Hax, 1997). In large part, this is because corporate strategic planning in practice evolved into large, expensive, and burdensome processes. These efforts were like master planning: they tried to predict various future states and design corporate responses to these predictions. As the forecasts so often turned out to be wrong, the resulting strategic plans became obsolete, just as the airport master plans do. According to a leading proponent of strategic planning in business, "The criticism of strategic planning was well deserved. Strategic planning in most companies has not contributed to strategic thinking" (Porter, 1987).

Yet airport managers and operators need to think strategically. They need to examine the range of future possibilities, position their organizations to respond flexibly to the events that occur, and in fact react appropriately as the future becomes clear. Sections 4.4 and 4.5 describe the suitable dynamic strategic planning in detail.

Good strategic thinking for an airport must consider the context of the airport/airline industry. Events far beyond the airport boundaries affect the development and consequently the planning for airports. Decisions made in faraway airline boardrooms can drastically upset airport developments. For example, US Airways' decision to consolidate its operations in Philadelphia turned Pittsburgh's airport into a "ghost town." Traffic at this former hub dropped from 19.8 million annual passengers in 2001 to only 8.0 million in 2010. Similarly, when American Airlines bought TWA and closed its hub in St. Louis, traffic there dropped from 30.6 million annual passengers to 12.4 million in just 4 years. Likewise, the decisions by Emirates to expand aggressively, and of Dubai to provide enormous new airport facilities, are causing shifts in traffic flows affecting major European hubs. Airports must look beyond their boundaries when they do their planning. Airport planners need to take a systems view that looks at the larger airport/airline industry.

4.2 Systems Perspective

We need to recognize that airports are part of a system of transportation. They do not exist just by or for themselves. We must avoid the planning box that focuses narrowly on single airports. This section considers the concept of airport systems and addresses the operational questions: Who plans the development of airport systems? Who will be planning them in the future?

Airport Systems

An airport is part of an airport/airline system. It is not independent. Each is a part of one or more networks connecting other airports. These networks and systems can be either geographic or operational.

Geographically, for example, one can think of the following:

- *Regional networks* linking smaller airports with a regional or national center, as commuter aircraft feed traffic from all over the Southeast United States into Atlanta, or Argentine airports connect with Buenos Aires
- *Metropolitan multi-airport systems* serving a single metropolitan area, as Oakland, San Jose, and San Francisco/International serve the Bay Area, and the de Gaulle and Orly airports serve Paris (see Chap. 5 for detail)
- *National networks* linking the major cities of a country, as major airlines do for large countries such as the United States, Germany, and Japan
- International and intercontinental networks, connecting countries with each other

Alternatively, one can think of networks and airport systems defined functionally, by the type of traffic or the carrier:

- *Integrated cargo networks*, such as those constituted by major cargo integrators such as UPS or FedEx, which give traffic and meaning to airports such as Louisville, Kentucky, and Los Angeles/Ontario, which otherwise would have little to do with each other
- *Low-cost* networks, served by airlines such as Southwest in the United States, Ryanair in Europe, or AirAsia in Southeast Asia, which serve secondary airports such as Boston/Providence and Miami/Fort Lauderdale, or Frankfurt/Hahn and London/Stansted

In general, an airport is part of several systems of airports simultaneously. Memphis, Tennessee, for example, is both the major hub for the FedEx system of airports and part of a regional feeder system. London/Stansted is part of both a “low-cost” system of airports and the London multi-airport system. As a rule, it is not possible to divide airport systems into self-contained subsystems or modules, as a car can be divided into the chassis, the engine block, and the drive train. Airport systems overlap. In practice, they do not have a precise definition in terms of the aviation and air transport network.

National governments classify airports in a variety of ways. However, these categories do not necessarily define the systems meaningfully. In the United States the FAA organizes airports by the relative number of passengers: it defines a “medium hub” as an airport with at least 0.25 percent, but less than 1 percent of the annual passenger boardings in the United States. This definition has little consequence for planning and development. In Japan, the governmental distinguishes between “international” airports and others, and this label has had great financial significance. Designated international airports have received far greater support from the central government than the others. However, a number of Japanese airports that have not been “international” officially, such Tokyo/Haneda, do in fact cater to international passengers and cargo. Here again, the governmental label does not identify the functional systems.

The essential point of discussion is that governmental jurisdictions do not define airport systems. A single jurisdiction may include two or more reasonably distinct and competitive systems. Thus, California and Germany include systems centered, respectively, on Los Angeles and San Francisco, and on Frankfurt, Munich, and Berlin. Conversely, a single system may overlap several jurisdictions. The metropolitan multi-airport system around Boston includes airports in three states (Boston, Massachusetts; Providence, Rhode Island; and Manchester, New Hampshire). Similarly, the feeder system for Amsterdam airport in the Netherlands extends over a large part of Britain, Belgium, and Northern France.

The consequence of this observation is that governments rarely can plan airport systems effectively. If the government encompasses several airport systems, it will find it politically difficult to choose among the possibilities, to pick “winners” among the competitive systems. This explains why the U.S. NPIAS is a nonselective assembly of proposed developments of individual airports. This documents aggregates projects from the “bottom up,” as indicated in the previous section. On the other hand, if the government controls only part of the airport system, it may not be able to have a decisive impact on it.

Planning Airport Systems

In the late twentieth century, many national governments had a substantial effect on their airport systems. They were particularly able to develop regional airports. Typically, the national ministry in charge of transportation or aviation would use its resources to invest in provincial projects. Thus

- Australia built excellent facilities in the capital cities of each state and territory.

- Canada invested heavily throughout its provinces and territories, most notably constructing Montreal/Mirabel, then the airport with the largest area of property in the world.
- Japan endowed each prefecture with a series of remarkable airports, leveling mountains, filling in valleys, and creating airport islands at Hiroshima, Osaka, and Nagoya.
- Mexico built international airports at coastal resort areas throughout the country, strongly promoting the development of tourism in Cancun, Cabo San Lucas, and similar sites.
- The United States taxed airline tickets, placed the proceeds in the Airport and Airway Trust Fund, and used this money to improve airside facilities throughout the United States.

Directive national planning of airport systems is generally obsolete, however. Most nations have judged that they can no longer afford to subsidize such programs. Indeed, many regional airport projects sponsored by national planning were plainly not economically efficient, however desirable they might have been from a political perspective. Their traffic would never have justified the investments. The development of Montreal/Mirabel airport offers a prime example: built by Transport Canada as a second airport, it never served more than a few million annual passengers. When Aéroports de Montréal took over the city's airports, it concentrated all scheduled traffic at Montreal/Trudeau, effectively closing Mirabel. Such "white elephants" led to a drive for economic efficiency and reduction of airport subsidies. The result has been the breakup of government-owned national airport groups into local companies and authorities, as in Australia, Canada, and Mexico. Thus, the opportunities for national planning of airport systems are largely gone.

Many metropolitan authorities have also developed second airports as part of multi-airport systems.[1] Their general practice has been to use the revenues provided by major international airports to finance these projects. Typically, these secondary facilities took a long time to build up traffic. They were often financially premature and economically inefficient, as Chap. 5 discusses in detail. Thus

- The Aéroports de Paris built Paris/de Gaulle to be the premier facility for France, yet this platform took a generation to overtake Paris/Orly as the busiest airport for the region.
- The British Airport Authority (as the government agency later replaced by the privatized BAA) built London/Stansted airport, which remained largely underutilized compared to its design for 15 years.

- The Port Authority of New York and New Jersey built major new facilities at New York/Newark and had a major passenger building built that remained totally unused for over a decade.
- The FAA built Washington/Dulles and attempted unsuccessfully to build up significant traffic for almost 20 years.

Long-term subsidized investments in major facilities are likely to be rare in the privatized environment of the early twenty-first century. Airport authorities or companies that have to raise money in the private sector are replacing governmental bodies that acted as if they could afford to disregard interest payments. The British Airports Authority is now a company, BAA plc. The U.S. government transferred its responsibilities for the Washington airports to the Metropolitan Washington Airports Authority. Privatization limits the opportunities for planning and developing multi-airport systems.

Privatization is leading to the end of airport systems planning as it was practiced in the twentieth century outside the United States. The national governmental bodies that could direct airport development are disappearing, and the local and regional airport authorities are increasingly required to justify projects for secondary airports to demanding private investors. What will be the airport systems planning of the future?

The inevitable consequence of privatization is that private, local interests prevail. Airport planning in the future is likely to focus increasingly on the development of individual airports. Planning efforts will focus on increasing each airport's competitive advantage over other airports. To the extent that a competitive market economy maximizes the public welfare, this is desirable. However, airports do suffer from congestion and create externalities. Basic economics tells us that, under these circumstances, competition is not necessarily in the overall best interests of society, of a nation or a region specifically.

Airport planning in the twenty-first century is likely in practice to be narrowly defined around the development of the airport facilities under the control of a single authority or company. The focus will be on the configuration of the airfield; the set of passenger buildings; the supporting people movers, baggage, and communication systems; the complex of cargo and maintenance facilities, and the modes of access. The remainder of this chapter assumes this perspective.

Airport companies might eventually evolve into large international operators of major airports. They could in this case develop strategies for developing airports as part of a coherent global system competing with other chains of airport operators. So far, however, international airport companies such as Abertis or GMR manage independent airport operations (Table 1.12). Currently, the large integrated cargo shippers such as UPS and FedEx appear to be closer to planning for their systems of cargo hubs. How this will develop is an open question. Who knows what the future will bring?

4.3 The Forecast Is "Always Wrong"

Experience demonstrates that forecasts about airport traffic are "always wrong." Comparisons between what a forecast indicated for a given period and what actually occurred almost invariably show a significant discrepancy. This is especially true when one considers forecasts over 10 to 20 years, that is, over the normal periods for the planning of major airport facilities. The differences between forecast and reality are most apparent when they concern the total level of operations. However, they are equally significant for planning purposes when they concern the composition of the traffic. For example, 10 million passengers at an origin and destination airport require quite different facilities than the same number when half of them are transfers (Chap. 14). As this section illustrates, the accuracy of forecasts of all types is low.

The fact that forecasts are unreliable has crucial implications for airport planning. Responsible planners consequently must accept that they do not know what levels or types of traffic will use the facilities they design. They need to anticipate that these facilities will have to serve different loads than the ones they now think are most probable. They therefore need to make sure that any design they propose will function well in these different conditions. In practice, they need to check the performance of their designs under different loads and, when they find deficiencies, they need to alter these designs to avoid the potential for future problems. In general terms, the fact that forecasts are unreliable means that designers need to create flexible designs that can adapt easily to the range of future conditions.

The unreliability of forecasts has well documented for a long time. Ascher (1978) illustrated the phenomenon through case studies across a variety of issues. Makridakis and colleagues demonstrated the inaccuracy of all kinds of forecasts, even in the short run, through extensive analyses of all the major methods available (Makridakis and Hibon, 1979; Makridakis et al., 1984). de Neufville (1976) presents extensive evidence of the poor performance of forecasting for airport systems. The U.S. Office of Technology Assessment (1982) gave an official account of the unreliability of forecasts of airport activities in the United States. More recently, Friedman (2004) compared the divergence between forecasts and actual results for the FAA Terminal Area Forecasts. This section illustrates this evidence for the benefit of readers who cannot refer to these and other citations.

Forecasts are unreliable because it is basically impossible to get good forecasts. All forecasting is based on some extrapolation of past trends into the future. However, past trends are constantly changing for economic, technological, industrial, and political reasons. The financial crisis in Asia in the 1990s, terrorism in 2001, and the worldwide recession after 2008 caused aviation traffic to subside considerably. New aircraft, larger and quieter, enable new routes, lower fares, and more traffic. Airline mergers and alliances change the services and public consumption of air travel. Political changes, such as the collapse of the Soviet Union and the dismantling of the traffic barriers between Russia, China, and the West, vastly reconfigured traffic patterns. The list of reasons why trends do not continue over a reasonable planning period is practically endless.

Moreover, as Chap. 19 on forecasting indicates, even to the extent that trends do continue, the mathematical methods for determining them are too subjective to permit analysts to determine definitively what that trend might be. In short, better methods or better analysts will not make forecasts more reliable. In fact, in the increasingly deregulated world of air transport, forecasts are likely to become even more unreliable than they have been.

The presentation of the track record for aviation forecasts first covers two substantive contexts: the estimation of costs and the forecasts of overall levels of traffic. It then indicates how longer planning periods and deregulation of aviation further increase the lack of reliability of forecasts. The object is to provide a sense of the large range of uncertainty that should be attached to any aviation forecast—and thus to all planning scenarios.

Cost Estimation

Estimates of construction costs for major projects are notoriously inaccurate. Differences between estimated and actual costs of 30 percent are common on standard projects, because of surprises on site, changed orders from the architects or owner, and the whole litany of things that can go wrong. On innovative, high-technology projects, these differences can be much larger, and analysts of project costs have observed standard deviation equal to the estimate. Benz (1993) provides an account of what has been observed in all kinds of fields of construction and production, and Fig. 4.1 summarizes those findings. Notice that he reports an overall standard deviation of about 40 percent. In round numbers this implies that, in one case out of three, actual costs differ from estimated costs by more than plus or minus 40 percent.

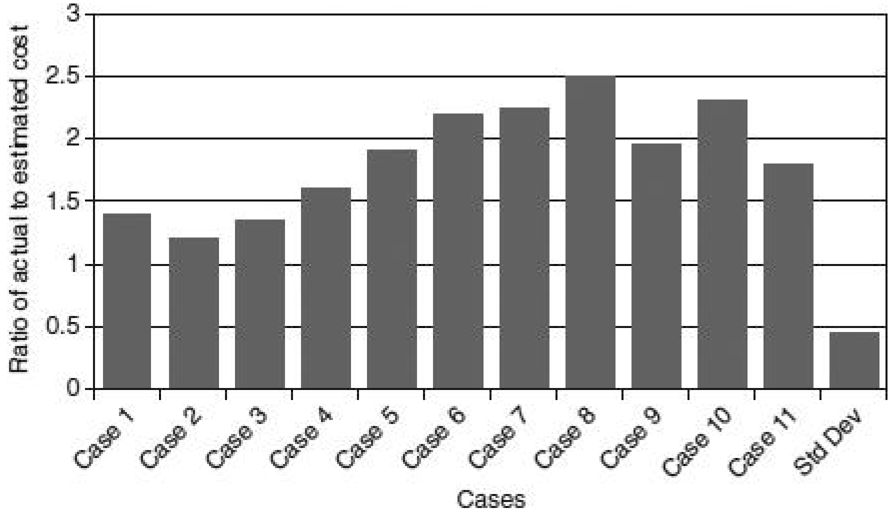


FIGURE 4.1 Average ratios of actual to estimated costs in various areas. (*Source: Benz, 1993*.)

An analysis of the cost of resurfacing airport runways illustrates the range of uncertainties in cost estimation. As this particular job is about the simplest to estimate, it provides a conservative indication of the uncertainties to be expected. The process of resurfacing runways uses primitive technology (asphalt is dumped off trucks and rolled to grade) on a clear surface with no hidden surprises. Figure 4.2 illustrates the distribution of the ratio of actual to estimated costs using data from the FAA Western Region of the United States (Knudsen, 1976). The analyst properly adjusted these data for inflation in the cost of construction over the time between the estimate and the execution of the job. Not surprisingly, the average and median actual costs are higher than estimated (about 25 percent in this case). What is remarkable is that the range of costs can be twice or half the average cost!

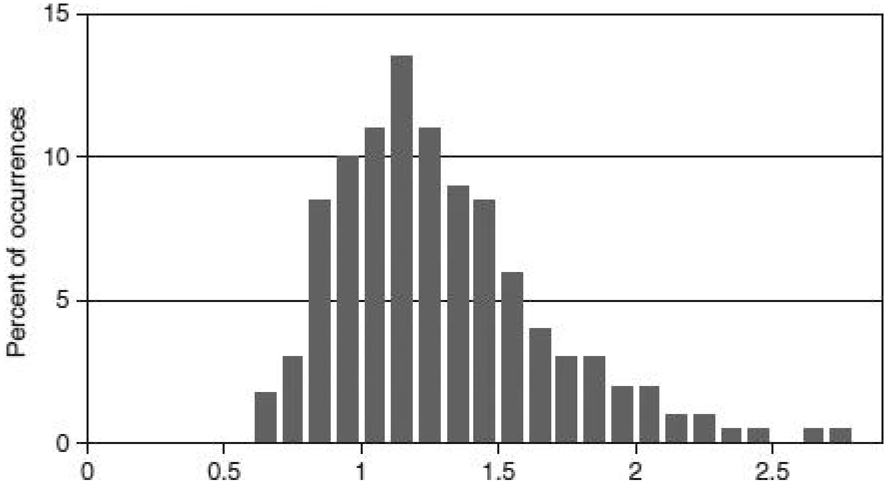


FIGURE 4.2 Distribution of ratio of actual to estimated costs for runway resurfacing projects. (*Source: Knudsen, 1976*.)

A general explanation for why actual costs vary from estimates has to do with the fluctuations in the real cost of materials and labor. Construction processes use commodities, such as steel and cement, whose production consumes a lot of energy. Thus their prices, and those of petroleum derivatives such as asphalt, rise and fall with the cost of petroleum, which is highly volatile: from 1999 through 2012 the price of a barrel of oil fluctuated between a low of about $10/barrel to over $130. Moreover, the variations in the price of oil, and the effective price of labor, depend largely on the state of the economy (de Neufville et al., 1977; de Neufville and King, 1991). During boom periods, the supplies of oil and labor are tight, so oil prices are high and employers have to pay overtime and premium wages. During recessions, however, supplies are plentiful, oil prices tend to drop, and workers are less demanding.

Aggregate Forecasts

The periodic swings in the overall economy naturally affect the overall level of aviation traffic. In boom periods, businesses need to travel and individuals have the money to do so. When there is a global or economic crisis, the growth in airport traffic correspondingly slows or decreases. The phenomenon makes medium-term forecasts, those covering 5 to 10 years, distinctly unreliable.

As a rule of thumb, half the medium-term forecasts differ from the forecast by more than 20 percent. This approximation is validated by repeated comparisons between forecasts and actual results over long periods and in different countries. For over 50 years for example, the U.S. FAA has each year been preparing both national forecasts of traffic and airport operations over the following 5 years—and reporting the actual levels observed for all the same categories (BTS, Annual). The comparison of such series, in the United States and elsewhere, demonstrates that 5- and 10-year forecasts are indeed easily wrong by more than plus or minus 20 percent.

Anyone can document the phenomenon by making similar comparisons. Figure 4.3 shows one such comparison. It shows how actual results easily deviate from forecasts in just a few years. Although many forecasts were off by over 30 percent, the data in Fig. 4.3 are especially conservative: the analysis done in 2004 deliberately omitted results for 2001 and 2002 when terrorism attacks and an economic downturn held back traffic growth. Those years were indeed extraordinary, but extraordinary events keep happening! If those years had been included, the deviations between forecasts and actual results would be even more striking.

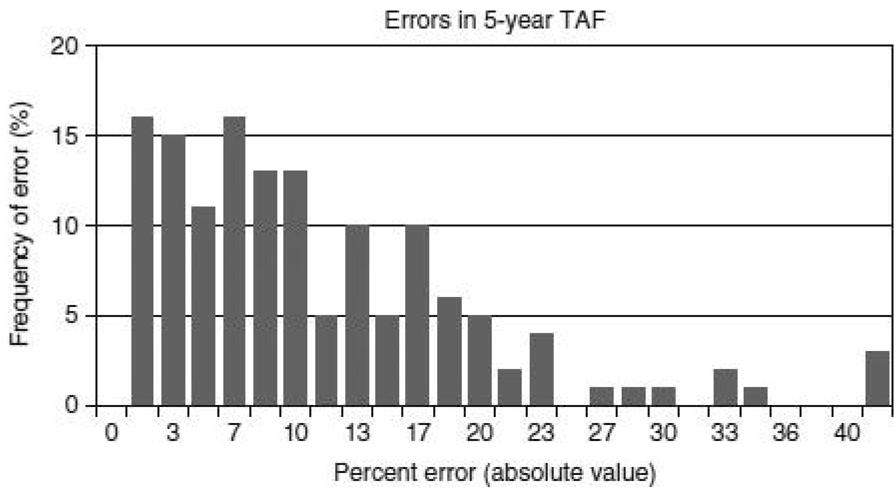


FIGURE 4.3 Distribution of ratio of actual to forecast passenger traffic based on FAA Terminal Area Forecasts done in 1992–1995 for the years 1997–2000. (*Source: Friedman, 2004*.)

Table 4.1 shows similar information from Japan, taken from their periodic forecasts of their national 5-year investment plans and subsequent statistics on passengers. For Japan, the average discrepancy between the forecast and the actual number of international passengers between 1980 and 1995 was 22 percent after 5 years and 40 percent after 10 years (Nishimura, 1999).

Forecast	Passengers (millions)			
For	Done In	Actual	Forecast	Percent Error over Actual
1980	1970	12.1	20	65
1985	1975	17.6	27	53
1990	1980	31	39.5	27
1995	1985	43.6	37.9	–13

Source: Nishimura, 1999, from Japanese Ministry of Justice Embarkation and Disembarkation Statistics, and Ministry of Transportation, Airport Investment 5-Year Plans.

TABLE 4.1 Comparison of 10-year Forecasts of International Passengers to Japan with Actual Results

Table 4.2 gives further insight into the inaccuracy of forecasts. It shows long-range forecasts for Sydney, Australia, prepared by three different authoritative groups. One came from a reputed international consultant, a second was the official forecast of the previous planning study, and the third was the forecast of the Australian Ministry of Aviation. Although these experts were all working with similar data, their estimates of the future differed widely. None of their forecasts was close to the actual level of traffic some 20 years later.

Forecast for Year	Consultant, 1974	Regional Study, 1978	National Ministry, 1983
1980	3.77	2.98–3.46	
1985	7.4	3.87–4.34	2.674–3.047
1990	9.8	4.71–5.51	2.762–3.751
2000 projected	12.0	6.27–8.66	2.938–5.159
2000 actual	10		

Sources: Australia Department of Aviation, 1985; Sydney Airport, 2001.

TABLE 4.2 Comparison of Actual and Forecast International Passengers through Sydney

Table 4.2 provides further lessons about forecasting in practice. Readers should notice the third- and fourth-place accuracy of the forecasts. This kind of precision is wholly unjustified. Most 20-year forecasts for aviation are lucky to get the first two decimal places right. Reporting more decimal places is pretentious. Two of the forecasts for Sydney have the great merit of providing ranges, reinforcing the notion that forecasts are not precise. However, these ranges are much too tight. They provide a range of only about plus or minus 20 percent over 20 years. By contrast, as the experience in the United States indicates, we can expect such deviations in a few as 5 or 10 years. The lessons are that aviation planners should do the following:

- Focus on the first two decimal points
- Use large ranges, on the order of plus or minus 30 percent or more over 20 years

Simply stated: large forecasting errors are normal.

Effect of Longer Planning Periods

The discrepancy between forecast and reality increases for longer forecasts. This is entirely to be expected. In the short run, inertia in the system keeps things moving as they were. In the longer run, trend-breaking events are more likely. Entirely new travel patterns may set in and make the actual results differ much more from forecasts.

A comprehensive analysis of airport master plans demonstrated the increase in forecast errors for longer-term predictions. Table 4.3 clearly shows how all measures of the error become larger with longer-term forecasts: the average discrepancy, the absolute range of the error, and the consequence standard deviation of the error.

	Error Characteristics (%)		
Forecast Years	Average	Range	Standard Deviation
5	23	36–96	23
10	41	22–140	34
15	78	34–210	76

Source: Maldonaldo, 1990.

TABLE 4.3 Discrepancies between the Forecast and the Actual Results Increase for Longer-Term Forecasts

The discrepancies between the forecasts and actual results apply equally to the content of master plans. By looking at old master plans and comparing them with what actually is built, it is easy to calculate statistics similar to those reported so far. In doing this exercise, the analyst has to consider both the projects in the plan that were not built and those that were built that were not in the original plan. Table 4.4 summarizes the results for one such study. It shows that the master plans accounted for less than half the projects constructed. As should be expected, the average discrepancy becomes larger for longer planning horizons.

	Error Characteristics (%)	
Forecast Years	**Average**	**Standard Deviation**
5	54	27
10	58	30
15	68	21

Source: Maldonaldo, 1990.

Table 4.4 Discrepancies between Projects Forecasted in Master Plans and Actually Built Increase for Longer-Term Forecasts

Effect of Economic Deregulation

Economic deregulation increases the volatility of traffic. This is because deregulation removes the barriers to changes in prices, frequency of service, and routes (de Neufville and Barber, 1991). Airlines can and do make sudden major changes in these circumstances, and may radically disturb the patterns and levels of traffic. These moves may have substantial effects on the largest airports. At smaller airports they may cause traffic to double or halve in just a few years. For example, Continental introduced a low-fare service to Greensboro, North Carolina, and doubled traffic from 2 to 4 million total passengers between 1993 and 1995. By 1997 Continental terminated this service, and traffic at Greensboro had fallen back to about 2 million passengers a year. Similarly, in the 1990s Delta created a connecting hub Cincinnati and built traffic up to a traffic peak of 22.8 million passengers in 2005—and then shut down these operations so that by 2010 traffic at Cincinnati had fallen below 8 million passengers.

Low-cost airlines such as Southwest or Ryanair can suddenly arrive on a market and generate huge increases in traffic. These may persist, or may fall if the airline fails. Southwest has continued to be successful over a generation, whereas PEOPLExpress rapidly doubled traffic at New York/Newark in the 1980s, from about 10 to about 20 million before it went bankrupt and deflated the traffic by half.

Major airlines can likewise make substantial moves. American Airlines moved a substantial block of its traffic from Chicago/O'Hare to Dallas/Fort Worth at the time deregulation became effective in the United States, thus dropping traffic through Chicago by about 15 percent in 1 year as Fig. 4.4 shows.

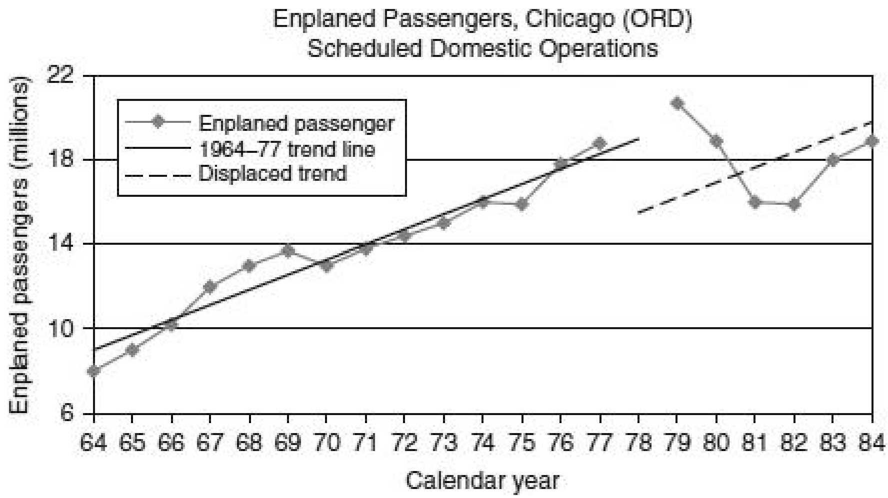


FIGURE 4.4 Volatility of air traffic at Chicago/O'Hare after deregulation in 1978.

Such radical changes in traffic can obviously affect the performance of an airport drastically. When US Airways moved the focus of its international operations from Baltimore to Philadelphia, the Baltimore airport was left with an underutilized international passenger building. Meanwhile, Southwest Airlines was expanding, so the total traffic at Baltimore stayed steady. However, Baltimore meanwhile had to build new facilities to accommodate this other form of traffic. Chapter 14 describes this case in detail.

The bottom line for airport planners and operators is that traffic can change rapidly in a deregulated environment. As of 2012, deregulation was the standard context in the busiest international markets: North America, Europe, Japan, India, Australia, and elsewhere. As this pattern spreads to other free-trade areas, and through "open-skies" policies allowing airlines to serve destinations in other countries freely, this volatility becomes increasingly important.

The fact that the forecast is "always wrong" means that master plans built around specific forecasts will also be "always wrong." This means that to get the planning right, it is necessary to move away from the notion of planning around a fixed forecast.

4.4 Concept of Dynamic Strategic Planning

Good planning needs to deal with reality. For airport systems, the fundamental reality is that future forecasts are highly unreliable. Forecasting errors of 20 percent or more after only 5 or 10 years is normal, and errors for longer-term forecasts are usually worse. Good planning therefore needs to deal with a broad range of possibilities.

The range of possibilities to be planned for includes both quantitative and qualitative factors. To demonstrate the inaccuracy of forecasts, the previous section stressed their measurable errors. These stem from changes in economic trends and policies, new technologies, new industrial organizations created by mergers and alliances, as well as from new political possibilities. The same factors clearly also influence the qualitative aspects of the loads on the airports. New technologies impose new requirements on the design of the airport. The introduction of the double-decked Airbus A380 meant that many airports had to modify their runways, taxiways, and passenger buildings. New political realities likewise have design implications. The creation of common market areas such as the European Community leads to different, generally easier, requirements for immigration and customs facilities. Meanwhile, a heightened concern for security imposes new requirements for examining and controlling baggage. Good planning needs to recognize the whole range of possibilities and anticipate solutions for the problems they pose.

This section develops the concept of dynamic strategic planning as an approach that incorporates the main elements needed for realistic planning: a good assessment of the issues, a flexible approach, and a proactive stance toward dealing with the future.

Assessment of the Issues

It is convenient to start with a SWOT analysis. This is a popular generic approach. Its name refers to a process whereby we systematically review the following:

- *Strengths* of the organization, in this case the airport, both internally and regarding its competition
- *Weaknesses*, again internally and regarding the competitors
- *Opportunities* for the airport, in terms of new markets, mergers, technologies, etc.
- *Threats* to the airport, in terms of the same kinds of events

A SWOT analysis guides us to an understanding of how the airport should develop its property activity, both physically and organizationally, so that it can shape and benefit from future developments. Physically, the airport might build new facilities. Organizationally, it might develop relationships with clients, post a favorable schedule of prices, and change its mix of products. Strategic planning in the business sense is a form of proactive, flexible planning.

Flexible Approach

Flexibility is essential. It is impractical to build now the facilities that will meet all eventualities. For example, facilities cannot both be large enough to satisfy the highest level of traffic anticipated, yet be small enough to avoid unnecessary expenses if traffic remains steady or drops to a low level. Planners need to establish some middle course, from which they can either grow the facilities as needed, or change them if some newer or lower level of traffic should arise. Consider the case of Baltimore-Washington Airport mentioned in Sec. 4.3. It faced a sudden drop in the level of international traffic, when their principal international carrier shifted the hub of those operations to Philadelphia. Although events like this are neither usual nor common, they are well within the range of possibilities and have happened elsewhere. Good planning in that case would have anticipated this possibility and would have designed the international passenger building with the flexibility to accommodate alternative traffic (Chap. 15). A flexible approach to planning and design would have avoided the difficulties associated with an underutilized building.

Dynamic strategic planning emphasizes flexibility. Its fundamental premise is that airport operators will inevitably have to adjust their plans and designs dynamically over time to accommodate the variety of futures that may occur. This emphasis distinguishes dynamic strategic planning from the traditional master or strategic planning, both of which build upon relatively fixed visions of the future.

Dynamic strategic planning represents a new vision of how airport systems planning should be done. It is particularly suitable for the current situation, in which privatized airlines compete in an increasingly deregulated environment, and increasingly privatized airports respond proactively to the opportunities and threats they perceive.

Although dynamic strategic planning is a new approach, it is entirely compatible with and builds upon the basic elements of traditional airport master planning and with strategic planning in management. It adds to the orderly process of the airport master plans by including the examination of several forecasts rather than one. It also assimilates the proactive approach of strategic planning, by encouraging planners to shape the future loads on the system, rather than reacting passively to whatever loads come to the airport. In short, this approach to planning represents a marriage of the best elements of both master and strategic planning, in a practical form suitable for routine use.

This new approach to planning is an extension of the master planning process outlined in Sec. 4.1 and detailed in standard guidelines. It differs in two ways. First, it substitutes a range of forecasts for the single forecast that the master planning process normally generates. In this regard, dynamic strategic planning simplifies the process, because it avoids the difficult and unsatisfactory process of trying to choose one forecast from among the many possible candidates. (See the discussion around Table 19.4 in Chap. 19 on forecasting.) In the subsequent phases of the process, dynamic strategic planning directs the planners to consider how each of their plans would

- Perform under the loads implied by the different forecasts
- Adapt to the new conditions these alternative scenarios represent

At this point, the dynamic plan is more complicated than the standard master plan. However, this additional effort can be managed by the appropriate use of computer-based tools such as decision analysis and simulation, as Sec. 4.5 indicates.

Proactive Stance

The approach is also strategic in that it is proactive. Dynamic strategic planning recognizes that planners can influence the nature of the airport traffic. They may preclude certain types or facilitate others. For example, as Chap. 14 describes, the construction of the passenger buildings at Kansas City made it impractical to service transfer traffic efficiently and impelled the locally based airline to establish its hub in another city. On the other hand, the planners went to great lengths to plan Denver/International to service transfer traffic efficiently, and thus maintained that airport as a leading transfer hub in the United States. Likewise, the developers of London/Luton airport consciously targeted the market of price-sensitive travelers and built their facilities to keep costs low. In a similar vein, Singapore has developed its facilities to offer premium services and thereby help establish and maintain that city as a favorite hub for business travelers. In each of these situations, the developments significantly influenced the traffic at the airport. Airport planners need to recognize the potential relationships between the possible airport designs and the airport loads. They should not apply a single range of loads to all possible plans, because the plans themselves may shape the loads. Good airport planners will incorporate this reality into the planning process.

Proactive planning is the alternative to the implicit attitude embedded in conventional master planning, which is that planners have to react to developments. Although a proactive approach has not been standard practice in airport planning, it is standard in business and totally possible in airport planning. The TBI Airport Management, Inc. demonstrated how this could be done in its development of Orlando/Sanford. Until around 1998, this airport had virtually no traffic and operated in the shadow of Orlando International, a magnificent first-class facility. A normal forecast would not have projected any significant traffic for the secondary airport in the near future. However, the private owners positioned Orlando/Sanford as an inexpensive base of operations, built appropriate facilities, and teamed up with holiday tours and charter carriers. By 2000, the airport operator had built up the traffic to about 1.2 million passengers, of whom nearly a million were international. In 2008, the airport traffic peaked at over 1.8 million passengers. The airport operator's planning and development shaped the future, rather than responded to it. As private airport companies become more significant in the industry, proactive planning is likely to replace conventional master planning where possible.

As dynamic strategic planning process and methods influence the type of traffic that may use the airport, analysts should correspondingly apply different loads may be applied to different sets of plans being considered. When the planning process examines airport configurations that favor transfer traffic, it should test them against forecasts with higher levels of transfers and of total traffic. Contrarily, when the process looks at plans that favor destination traffic, it should test these against forecasts that have little transfer traffic.

Most important, a dynamic strategic plan is phased. It focuses on finding the most appropriate initial developments. This first phase of development should permit the planners to respond appropriately to the future levels of traffic. For example, they might develop a passenger building that accommodates both international and domestic traffic in a first phase. In a later period, they could expand the capacity to serve either or both activities, or could substitute one capacity for the other, depending on the circumstances. See Example 4.1. The focus is not, as in the master plan, on describing a future long-range vision that in practice never is implemented. The focus of the dynamic strategic plan is on identifying the right initial position that permits effective responses to future opportunities and developments.

Example 4.1 The original master plan for the redevelopment of Mombasa airport in Kenya anticipated two distinct passenger buildings, one for domestic and the other for international traffic. Each was supposed to be large enough to meet its level of anticipated traffic.

The dynamic strategic plan recognized the major risks that the proportion of international traffic could shift radically, as passengers might come directly from Europe or transit through Nairobi. If this happened, one or the other of the new buildings might be crowded while the other was underused.

The strategy adopted was to build a single passenger building capable of serving about half the eventual growth. This facility was equipped to serve international traffic on one side, domestic traffic on the other side, and either traffic through shared use in the middle (Chap. 15 discusses shared-use facilities). This arrangement allowed the building to serve a range of mix of traffic immediately and to expand selectively in the future. It also enabled the airport to defer the decision about how much they should extend the building, for which kind of traffic. Postponing that decision until traffic patterns had matured allows them to choose an expansion appropriate for the actual traffic. This flexibility considerably improves the value of the design: deferring decisions until you know what you need leads to better choices; deferring construction saves interest—and even capital costs if the traffic does not develop as much as originally supposed.

The new elements require different analyses than those involved in master planning. To do a dynamic strategic plan, the analysts need to look at many scenarios, over several periods. In the twenty-first century this wider perspective can be obtained with a reasonable amount of effort. Planners can do these analyses, using computer models and computer-based analyses that simulate alternative outcomes.

Overall, dynamic strategic planning encourages planners to think like players of chess or other strategic board games. Planners should

- Think many moves ahead
- Choose an immediate development or move that positions them to respond well to whatever develops next
- Rethink the issues after they see what happens in the next phase
- Adjust their subsequent developments or moves correspondingly

Good planners for the uncertain environment of airport systems will, as good chess players do, emphasize good positions and flexibility.

4.5 Dynamic Strategic Planning Process and Methods

The process for executing a dynamic strategic plan is a modified form of the master planning procedure described in Sec. 4.1. In the following list of the essential elements of a dynamic strategic planning process, these additions appear in italics. The steps for preparing a dynamic strategic plan are thus the following:

- Inventorying existing conditions
- Forecasting range of future traffic, *along with possible scenarios for its major components (international, domestic, and transfer traffic, airline routes, etc.)*
- Determining facility requirements *suitable for the several possible levels and types of traffic*
- Developing several alternatives for comparative analysis
- Selecting the preferred *first-phase development, the one that enables subsequent and appropriate responses to the possible future conditions*

From an operational perspective, doing dynamic strategic planning requires the team to look at many more scenarios. However, computer models provide the basic tools for investigating the effects of different scenarios at a reasonable cost. They can analyze hundreds of situations easily. The real effort associated with using computer models is not in the calculations, which are virtually immediate. The tricky part is in finding models that are easy and cost-effective for each task. Many already exist, however, and consultants are developing many more. While better models will always become available and are to be hoped for, enough exist to make it possible to consider many design alternatives for many scenarios. (ACRP Report 76, 2012.)

Most computer models simulate the performance of facilities under loads. They show how a particular design or configuration—of a runway system, passenger building, or people mover, for example—performs under different conditions. Current practice routinely uses a wide range of models for different situations. These models can be deterministic or probabilistic. The former typically use a spreadsheet analysis to specify the result of particular configurations or patterns. Chapter 14, on the configuration of passenger buildings, describes a model that estimates the walking distances that result from various types of traffic on a passenger building. Probabilistic models are most appropriate when there are queues and delays (see Chap. 20). In short, sufficient models exist to enable good analyses of alternative developments under different conditions.

A different class of models facilitates the analysis of risk. Decision analysis and simulation enable planners to determine which initial developments lead to the best long-term development. These tools have the great merit of helping planners estimate the value of the flexibility that they might introduce into the design. While it often costs time and money to create to enable the development of alternative paths of development, flexible designs generally provide win-win solutions: they enable developers to save on immediate construction costs, subsequent interest payments, and add value by permitting designers to build the right kind of facility once the future arrives and the actual needs become clear (de Neufville and Scholtes, 2011).

4.6 Summary of Dynamic Strategic Planning

The value of dynamic strategy planning is that it helps planners understand the critical issues and prepare good responses for possible futures. Developing strategic thinking is the important ingredient. The mechanical parts of planning identified in the previous section, the analysis of the possible futures and their consequences, are necessary but not sufficient. This needs emphasis. It is easy for planners to focus on the computer models as they need to consider many issues under various circumstances, to calculate the performance and value of facilities under different loads. However, computer analyses are not sufficient to develop a good plan; they do not substitute for strategic thinking.

Strategic thinking has two elements. The first is a critical assessment of the competitive situation, the SWOT analysis. This is comparable to the inventory of current conditions part of the master planning process. The second is a creative process of identifying good responses to the prospective opportunities and threats.

In developing their strategic thinking, planners first need to go through a SWOT analysis. Specifically, they need to identify the following:

- *Strengths* of the existing site or airport, the characteristics that give it advantages over other sites or competitive airports—this may include a site central to an aviation network that favors transfer traffic, for instance
- *Weaknesses* of the facility, those that may limit its growth or opportunities—these might include the weakness of the local airline that might merge with another airline that focuses traffic elsewhere, or physical limitations of the site
- *Opportunities* for the region, which enhance its future prospects—this might be an expanding economic base that will lead to greater air travel and cargo
- *Threats* to the airport and region, from competitive airlines, airports, or other factors

Whereas it is easy to specify a way to inventory the situation, there is no checklist for creative thinking about strategic solutions. Good strategic thinking, like good chess playing, comes from observing examples and practice. Professional, experienced leaders in airport planning and development should be able to think through the possibilities and arrive at good ideas, provided they allow themselves the time and devote careful thought to this issue. By way of illustration, Example 4.2 describes the application of dynamic strategic planning to the development of the passenger buildings for Kuala Lumpur International Airport (KLIA).

Example 4.2 The original master plan for KLIA called for a main passenger building to serve domestic traffic, and satellite buildings for international traffic. It designed each terminal to fit the traffic forecast for its intended traffic.

The dynamic strategic plan revisited this master plan. It led to a strategy designed to strengthen the competitive position of the airport and its airlines. The SWOT analysis led to these observations:

- *Strengths:* The new airport would have enormous capacity and should be attractive for hub operations by a major airline.
- *Weaknesses:* The original design split the domestic and international operations. This is inefficient for any airline serving connecting traffic. Moreover, the split between international and domestic facilities lacked flexibility to deal easily with major shifts in the future level of domestic traffic, which was highly uncertain given the possibility that the Southeast Asian countries might form a common market.
- *Opportunities:* To become a leading aviation hub for Southeast Asia, given the capacity and low cost of operation due to inexpensive land.
- *Threats:* Competition from Singapore and Bangkok, both of which are strong aviation hubs served by excellent airlines.

Taken together, the SWOT analysis emphasized the opportunities for KLIA to become a major regional hub—if the design team configured the airport to provide integrated operations for the major carrier and achieve low costs.

In retrospect, the SWOT analysis was remarkably prescient. As of 2012, KLIA had become a major international hub for AirAsia, a rapidly expanding low-cost carrier.

Exercises

4.1. Look up previous forecasts for your local (or some other) airport. Compare these with the actual results. Calculate the deviations between forecast and reality, in terms of the percent of what actually occurred. To the extent possible, estimate how this percent error increases for longer-range forecasts.

4.2. Obtain previous master plans for an airport. Compare these with what actually has been constructed. To what extent has the airport invested in facilities that were not part of the original plan? Not invested in facilities that were in the plan?

4.3. Start by doing a SWOT analysis for an airport of interest. Then discuss: Which of the issues can the airport operators influence through their designs and developments? What kind of developments might position this airport to respond most effectively?

4.4. For the same airport as in Exercise 4.3, think about what elements of the future traffic might be most uncertain. What could the airport operators do to give themselves flexibility so that they could adjust their future developments to deal effectively and efficiency with these different scenarios?

References

ACRP, Airport Cooperative Research Program (2012) *Addressing Uncertainty about Future Airport Activity Levels in Airport Decision-Making*, ACRP Report 76, Transportation Research Board, Washington, DC. www.trb.org.

Ascher, W. (1978) *Forecasting: An Appraisal for Policy-Makers and Planners*, Johns Hopkins University Press, Baltimore, MD.

Australia Department of Aviation (1985) *Second Sydney Airport: Site Selection Programme*, Australia Department of Aviation, Canberra.

Benz, H. (1993) "Dynamic Strategic Plan for an Embedded Space Computer," S.M. thesis, Management of Technology Program, Massachusetts Institute of Technology, Cambridge, MA.

Bureau of Transportation Statistics (Annual) *Aviation Activity Statistics*, http://www.bts.gov/publications/airport_activity_statistics_of_certificated_air_carriers/.

de Neufville, R. (1976) *Airport Systems Planning: A Critical Look at the Methods and Experience*, MIT Press, Cambridge, MA, and Macmillan, London, U.K.

de Neufville, R. and Barber, J. (1991) "Deregulation Induced Volatility of Airport Traffic," *Transportation Planning and Technology, 16*(2), pp. 117–128.

de Neufville, R., Hani, H., and Lesage, Y. (1977) "Bidding Models: Effect of Bidder's Risk Aversion," *ASCE Journal of Construction Division, 103*, March, pp. 57–70.

de Neufville, R. and King, D. (1991) "Risk and Need-for-Work Premiums in Contractor Bidding," *ASCE Journal of Construction Engineering and Management, 117*, December, pp. 659–673.

de Neufville, R. and Scholtes, S. (2011) *Flexibility in Engineering Design*, MIT Press, Cambridge, MA.

FAA, Federal Aviation Administration (2005) *Airport Master Plans*, Advisory Circular 150/5070-7, U.S. Government Printing Office, Washington, DC.

FAA, Federal Aviation Administration (2004) *The Airport System Planning Process Master Plans*, Advisory Circular 150/5070-6B, U.S. Government Printing Office, Washington, DC.

Friedman, J. (2004) "Terminal Area Forecast (TAF) Accuracy Assessment Results," MITRE CAASD, unpublished report.

Hax, A., (ed.) (1997) *Planning Strategies That Work*, Oxford University Press, Oxford, U.K.

Hax, A. and Majluf, N. (1996) *The Strategy Concept and Process*, 2d ed., Prentice-Hall, Englewood Cliffs, NJ.

IATA, International Air Transport Association (2004) *Airport Development Reference Manual*, 9th ed., IATA, Montreal, Canada.

ICAO, International Civil Aviation Organization (1987) *Airport Planning Manual, Part 1, Master Planning*, 2d ed., Doc. 9184, ICAO, Montreal, Canada.

Knudsen, T. (1976) "Uncertainties in Airport Cost Analysis and Their Effect on Site Selections," Ph.D. thesis, Institute of Transportation and Traffic Engineering, University of California, Berkeley, CA.

Makridakis, S. and Hibon, M. (1979) "Accuracy of Forecasting: An Empirical Investigation," *Journal of the Royal Statistical Society*, Series A, *142*(2), pp. 97–145.

Makridakis, S., Andersen, A., Carbone, R., and Fildes, R. (1984) *The Forecasting Accuracy of Major Time Series Methods*, Wiley, New York, NY.

Maldonaldo, J. (1990) "Strategic Planning: An Approach to Improving Airport Planning under Uncertainty," S.M. thesis, Technology and Policy Program, Massachusetts Institute of Technology, Cambridge, MA.

Mintzberg, H. (1994) *The Rise and Fall of Strategic Planning—Reconceiving Roles for Planning, Plans, Planners*, The Free Press, New York.

Nishimura, T. (1999) "Dynamic Strategic Planning for Transportation Infrastructure Investment in Japan," S.M. thesis, Technology and Policy Program, Massachusetts Institute of Technology, Cambridge, MA.

Office of Technology Assessment (1982) *Airport and Air Traffic Control Systems*, U.S. Government Printing Office, Washington, DC.

Porter, M. (1985) *Competitive Advantage: Creating and Sustaining Superior Performance*, The Free Press, New York.

Porter, M. (1987) “Corporate Strategy: The State of Strategic Thinking,” *The Economist*, May 23, pp. 17–22.

U.S., Secretary of Transportation (2010) *The National Plan of Integrated Airport Systems (NPIAS) 2011-2015*, Report to the U.S. Congress, U.S. Department of Transportation, Federal Aviation Administration, Washington, DC.

---

[1] This discussion on the development of *second* airports as part of a system does not mention the many projects planners saw as *replacement* airports, such as those at Dallas/Fort Worth and Houston in the United States, or at Milan and Osaka. In those cases, the planners intended to close the older airports to commercial traffic yet the older facilities remained active. Although the airport developers in effect created multi-airport systems in those cases, that was not their intent.

CHAPTER 5

Multi-airport Systems

Multiple airport systems exist in all the metropolitan areas that generate sufficient originating traffic. They also exist around many other cities. They serve about 80 percent of world traffic each year (in 2009, over 2 billion of the estimated 2.5 billion total passengers). Multi-airport systems are an important aspect of the airport/airline industry.

Competition between airports in a multi-airport system inevitably leads to concentration of traffic at a primary airport. Except for the cities with the largest markets for air travel, traffic at secondary airports is generally much smaller and more volatile than at the primary airport. This is a crucial reality. Planners need to recognize and deal with it. Failure to do so has led to many politically and financially embarrassing planning failures, as Sec. 5.3 describes.

Effective planning of multi-airport systems requires an understanding of the dynamics of the competition between airports. Airport managers need to appreciate the factors that favor the growth and shape the opportunities of multi-airport systems. They should recognize that neither governments nor airport operators have much control over the market forces that shape the allocation of traffic between multiple airports. They can then properly assess the risks associated with these systems and invest accordingly.

Good planning will understand that the market for airport services has different components with distinct needs and uncertain futures. Hub-and-spoke operations, low-cost airlines, business commuters, and regional markets can and do shift opportunistically from one airport to another in a multi-airport system. Given these uncertainties, economically effective development of multi-airport systems therefore calls for investment in a range of flexible facilities, as Chap. 4 indicates. These will enhance the airport operators' opportunities to respond to the changing market patterns and enable them to cater to the range of services airport users desire.

5.1 Introduction

This chapter shows how to plan, develop, and manage multi-airport systems of airports. These are sets of two or more commercial airports in a greater urban area, as Sec. 5.2 defines in detail. They present unique difficulties for airport planners and operators, because the airports in the system compete with each other for traffic. Multi-airport systems demand special attention.

As of 2012, some 70 multi-airport systems already exist. This number is growing as traffic increases, as Bonnefoy and Bonnefoy et al. (2008) documented. They serve, for example, Chicago, New York, and San Francisco; Frankfurt, London, and Paris; Seoul, Shanghai, and Tokyo; and Mexico City and Buenos Aires. Figure 5.1 shows some examples.

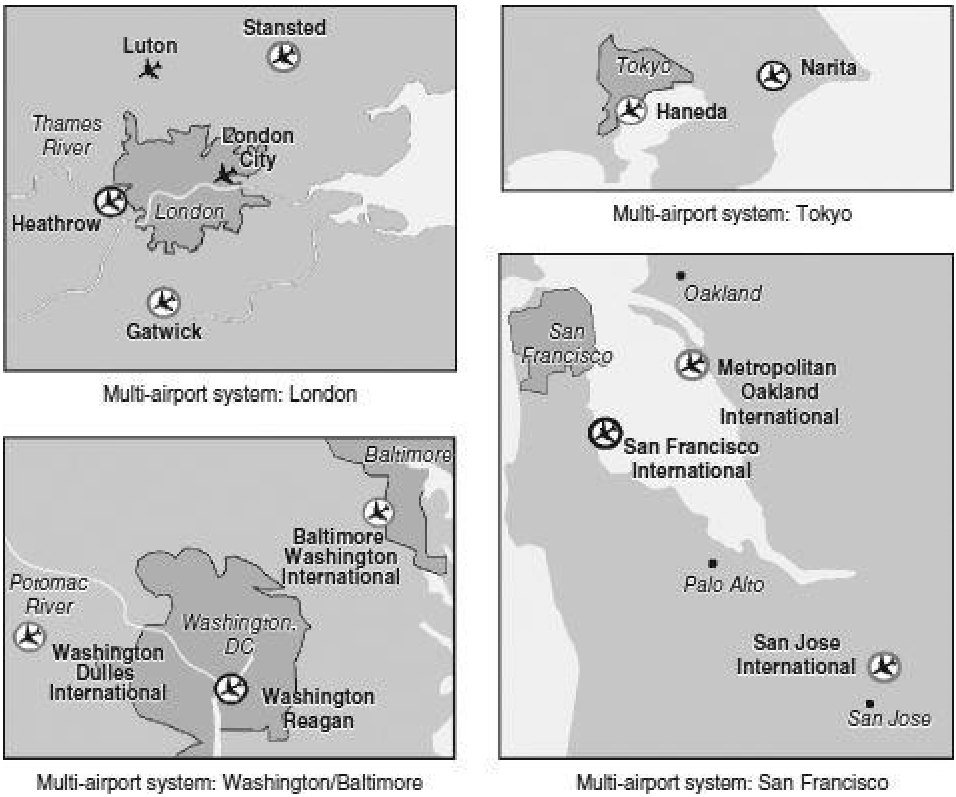


FIGURE 5.1 Airports in multi-airport systems of London, San Francisco, Tokyo, and Washington.

The issue with multi-airport systems is that airport planners, operators, and governmental sponsors frequently misjudge the development of individual, constituent airports. Lack of understanding of the way multi-airport systems perform has led to significant expensive and embarrassing failures. Section 5.3 describes some of these cases to motivate concern for the problem and to provide a foundation for the discussion of why and how multi-airport systems tend to evolve. This presentation particularly documents the systematic failure of planning policies that attempt to allocate traffic to the several airports in a multi-airport system. It also illustrates the volatility of customers for the services of secondary airports in a multi-airport system, a factor that has significant implications for how and when airport operators should invest in second or third airports around a city.

Market dynamics provide the basic explanation of the level and distribution of traffic among airports in a multi-airport system. Technical factors, political considerations, and chance do modulate the effects of the market. However, competitive market forces define the underlying structure of the outcomes. In brief, the competition of the providers of airport services for customers concentrates services for any market at specific airports. As Sec. 5.4 indicates, this concentration is a specific example of a wider phenomenon characteristic of the development of regional economies worldwide. Overall, the dynamics of the competition between the markets makes the longer-term outcomes uncertain.

An important practical issue in this context is the question of what services constitute a "market." Indeed, a market may focus on individual airlines, destinations, types of services, or fare levels. For example, Southwest Airlines is the prominent carrier at Dallas/Love Field and Miami/Fort Lauderdale. Paris/Orly specializes in North African and Caribbean destinations. Seoul/Gimpo serves internal Korean destinations. London/City is a focus for commuter business travel. Moreover, the situation may change, in some cases rapidly. Washington/Dulles, for instance, evolved from a minor airport serving only a few million annual passengers to a major international/domestic hub once United Airlines decided to use it as a base of operations. Markets can and do evolve in different ways and at different times.

Operators of airports within a multi-airport system face important strategic choices about how they will influence the development of these markets. Airlines also make similar choices. Jointly, the various airports and airlines are engaged in sequences of actions whose eventual outcome may be difficult to foresee.[1]

Airport planners and operators dealing with multi-airport systems have to deal with uncertain, unstable situations. They should actively seek to influence the situation for the benefit of their airports and the region. At the same time, they should be cautious in their investments, because the volatility in the level of traffic and the nature of their customers may make the facilities at secondary airports obsolete or unnecessary. As Sec. 5.5 concludes, airport planners should carefully assess the risks and invest in flexible facilities that give them appropriate options on future developments.

5.2 Basic Concepts and Issues

Definitions

Market dynamics are the central influence on the development of airports in a metropolitan area, as Sec. 5.4 demonstrates. It is crucial to use a concept of multi-airport systems based on how the customers and users of the system see it. A functional definition that reflects the realities of the market is appropriate.

Thus, for the purposes of airport planners and operators,

> a multi-airport system is the set of significant airports that serve commercial transport in a metropolitan region, without regard to ownership or political control of the individual airports.

This definition involves several important points:

1. It focuses on airports serving commercial transport. It leaves out military bases, such as Andrews Air Force Base near Washington, DC, or Yokota in the Tokyo area. It does not consider airports dedicated to aircraft manufacturing or shows, such as Boeing Field in Seattle and Le Bourget in Paris. It neglects general aviation airfields such as Van Nuys in the Los Angeles area. All these facilities are important from the perspective of air traffic control, but they are not factors in the market for the airport/airline services.

2. It refers to a metropolitan region rather than a city. As a practical reality, this region may include several distinct cities. The San Francisco region from this perspective includes the cities of San Francisco, Oakland, and San Jose. The airports of these cities all serve, at different times and in different ways, passengers and cargo associated with the San Francisco metropolitan area.[2] To reflect this reality, it is usual to refer to airports in a multi-airport system with a label indicating their metropolitan region: for example, San Francisco/International and San Francisco/Oakland.[3]

3. With its focus on the market, the definition does not pay attention to who owns the airport. The Paris multi-airport system consists of the two airports owned by the Aéroports de Paris (Paris/de Gaulle and Paris/Orly) plus the other one in the region that people use, Paris/Beauvais.[4] Similarly, the definition does not pay attention to administrative boundaries. The multi-airport system for Boston in the state of Massachusetts includes Boston/Providence in the state of Rhode Island and Boston/Manchester in the state of New Hampshire. These airports are within an hour or less of the Boston suburbs, often closer in terms of travel time than the primary Boston/Logan airport, and serve many customers in the greater Boston metropolitan region.

4. Finally, the definition focuses on significant airports, typically those that serve more than a million passengers a year. This focuses the discussion on facilities that contribute meaningfully to the air transport services of a metropolitan region.

Prevalence

Multi-airport systems constitute a sizable segment of the airport industry. As of 2009, they already catered to over 2 billion total passengers, well over 80 percent of worldwide traffic. These 70 multi-airport systems worldwide included 160 airports.

Multi-airport systems have been a feature of all metropolitan areas with the most originating and terminating traffic, without exception and over several decades. This is a remarkable fact. It stresses the strength of market forces to create and maintain multi-airport systems. Table 5.1 demonstrates this phenomenon. It presents the primary (or largest) and secondary airports at all cities that are the largest generators of traffic. Notice that, above a specific level of originating traffic, all the metropolitan areas feature a multi-airport system.

Metropolitan Region	Multi-airport System	Traffic, Millions of Passengers	
		Total	Originating
London	Yes	130	47
Tokyo	Yes	94	34
New York	Yes	101	33
Paris	Yes	85	30
Los Angeles	Yes	76	28
Beijing	Yes	66	24
Shanghai	Yes	57	24
Hong Kong	Yes	70	23
Washington	Yes	58	21
Chicago	Yes	80	20
San Francisco	Yes	54	19
Miami	Yes	52	19
São Paulo	Yes	48	18
Frankfurt	Yes	55	17
Seoul	Yes	44	17
Dallas/Fort Worth	Yes	60	17
Moscow	Yes	41	16
Istanbul	Yes	36	16
Houston	Yes	45	15
Rome	Yes	39	15
Milan	Yes	33	15

Source: de Neufville database.

TABLE 5.1 Metropolitan Regions Generating More Than 15 Million Originations in 2010

The level of originating traffic needed to justify and maintain a second airport has not been constant. It keeps rising, as Sec. 5.4 explains. As of 2010, the minimum level was about 15 million annual originating passengers for the entire metropolitan area, as Table 5.1 indicates. This threshold is likely to change over the coming generation.

The emphasis on originating traffic is vital to the understanding of multi-airport systems. The focus on locally generated traffic excludes transfers. The passengers beginning (and on return, ending) their trips in the metropolitan area create the pressure for multiple airports for their region. The transfers passing through the region want easy connections to their next flights and clearly prefer to be at a single airport.

Assuming that the number of passengers originating and returning to a region is equal and that either is half the total number of passengers less the transfers, we obtain that

$$\text{Originating passengers} = \frac{1}{2}\,(\text{total passengers} - \text{transfers}) \tag{5.1}$$

This number can be difficult to determine. Many airports and airlines do not release information about the number of their transfer passengers. Thus, the figures for originating traffic in Tables 5.1 to 5.3 are estimates. This inevitable lack of precision does not affect the overall association of multi-airport systems with the biggest traffic generators.

Many cities with lower levels of originating traffic also have multi-airport systems. Some of these are developing multi-airport systems (Table 5.2). Other metropolitan areas feature several airports primarily for technical or political reasons (Table 5.3). Technical reasons, for example, led Taiwan to develop a major international airport for its capital, Taipei/Taoyuan, equipped with 3600-m runways capable of handling large transoceanic aircraft. The downtown airport, Taipei/Sung Shan, is popular with local traffic but simply cannot handle long-distance aircraft with its 2550-m runways. Political reasons led the U.S. Department of Defense to develop Montreal/Plattsburg originally as a base for long-distance bombers.

Metropolitan Region	Multi-airport System	Traffic, Millions of Passengers	
		Total	Originating
Osaka and Kobe	Yes, in different cities	32	15
Boston	Yes, in different cities	32	14
Bangkok	Old airport struggling	40	14

Source: de Neufville database.

TABLE 5.2 Metropolitan Regions almost Generating 15 Million Originations in 2010

Metropolitan Region	Multi-airport System	Traffic, Millions of Passengers	
		Total	Originating
Düsseldorf/Bonn	Political, former capital	31	13
Taipei	Technical, runway length	25	11
Berlin	Political, divided city	21	10
Rio de Janeiro	Technical, runway length	15	6
Buenos Aires	Technical, runway length	14	6
Belfast	Technical, runway length	7	3

Source: de Neufville database.

TABLE 5.3 Multi-airport Systems due to Political or Technical Reasons

Unequal Size

Airports within a metropolitan multi-airport system characteristically have significantly different levels of traffic. The typical pattern is that a city has a primary airport (with the most traffic) and one or more secondary airports with between 10 and 50 percent of the traffic of the primary airport. Table 5.4 indicates relative levels of traffic of the secondary airports, compared to the primary airport in each of the multi-airport systems associated with the cities with the most originating traffic. The secondary airport rarely has as much traffic as the primary airport serving the most passengers.

Metropolitan Region	2010 Traffic at Secondary Airports as % of Primary		
	Second	Third	Fourth
London	44	30	14
Tokyo	52		
New York	73	47	4
Paris	43	3	
Los Angeles	15	8	8
Beijing	2		
Shanghai	78		
Hong Kong	54		
Washington	95	78	
Chicago	26		
San Francisco	24	22	
Miami	64		
São Paulo	58	20	
Frankfurt	7		
Seoul	46		
Dallas/Fort Worth	14		
Moscow	79	42	
Istanbul	22		
Houston	21		
Rome	14		
Milan	47	41	

Source: de Neufville database.

TABLE 5.4 Traffic at Secondary Airports Is a Fraction of That at Primary Airports

The primary airports with the most traffic are not necessarily the largest by size. Washington/Dulles has more runways and land than downtown Washington/Reagan, but for the first 20 years of its existence it served only about 20 percent of the traffic of Washington/Reagan (about 3 million passengers annually, compared to about 14 million). In fact, airport operators have often built major new airports far away from prospective clients, and these users have preferred to stay with the older primary airport until a major market shift occurs. For example, traffic at Washington/Dulles only grew significantly once United Airlines located a transfer hub there. This action completely changed the market for local customers.

Exceptionally, a secondary airport may have about the same level of traffic as the primary airport. This happens when a secondary airport grows and overtakes a primary airport as occurred when Paris/de Gaulle grew past Paris/Orly and Washington/Dulles overtook the other regional airports.

5.3 Difficulties in Developing Multi-airport Systems

The development of airports in a multi-airport system has always been problematic. The symptoms of this phenomenon come in several forms. Sometimes these overlap, as the examples indicate. As an overview, the following issues arise:

1. Not enough traffic comes to the new airport, resulting in an expensive and embarrassing "white elephant." Many cities have built major new airports and then had problems attracting customers. This has happened at London, Montreal, New York, São Paulo, Tehran, and Washington.
2. It is politically and economically difficult to close an old airport. The traffic then divides between the airports, resulting in poor service and insufficient traffic at the new airport until sufficient traffic builds up. This occurred at Buenos Aires, Edmonton, Milan, and Osaka.
3. There is not enough traffic to support a multi-airport system. This situation also leads to poor service, low traffic at each airport, and financial losses. Montreal provided the prime example.
4. It is impractical to allocate traffic away from a congested primary airport to alternatives in the region. This frustrates planners, who would like to reduce noise and congestion in one part of the region and provide service to another. Los Angeles, London, Milan, and Osaka have each experienced this difficulty.
5. Traffic at the secondary airport is volatile, both in level and in type. The result is that the operators of these facilities can alternately face underutilization and congestion and often have inappropriate facilities for their clients. Operators at London/Stansted, New York/Newark, and San Francisco/Oakland have faced these issues.

Difficulties in developing multi-airport systems keep occurring. For example, the operator of the Milan airports tried (and failed) to force airlines to leave the convenient Milan/

Linate airport and move to major new facilities at Milan/Malpensa. Thailand faced similar issues when it opened Bangkok/Suvarnabhumi and tried to close Bangkok/Don Muang. To avoid such problems, it is useful to learn from the examples of the past, discussed as follows.

Insufficient Traffic at New Airport

Planners have often mistakenly assumed that new airports would collect traffic from their "catchment areas." A catchment area for a facility includes all the places that have easier access to it, in terms time, distance, and expense, than to its competition. This concept applies to the siting of facilities for industries selling undifferentiated products where cost considerations dominate (e.g., the shipment of bauxite ore to aluminum smelters) (Weber, 1929). This approach is not suitable for airport planning, however, as the following examples demonstrate and the theory in Sec. 5.4 explains.

Thus, the Port of New York Authority in the early 1970s created three large passenger terminals at New York/Newark in the state of New Jersey.[5] An important part of the planners' reasoning was that, because this facility would be much more convenient to New Jersey residents and businesses, it would therefore attract about a third of the metropolitan traffic. Indeed, by going to New York/Newark the travelers from New Jersey avoid the congested crossings of the Hudson River, Manhattan, and the East River to get to the other two regional airports (New York/LaGuardia and New York/Kennedy). However, for a long time many passengers—and airlines—avoided New York/Newark airport. Travelers would drive by and continue on to New York/LaGuardia, for example. Airlines did not increase their flights or service to make use of the new capacity at New York/Newark, since the traffic was not there. The result was that for over a decade the Port Authority had to board up and close one of their three new passenger buildings. Lack of understanding caused a very large and embarrassing mistake.

Similarly, the British Airports Authority[6] built London/Stansted airport to the northeast of London, with the idea that it would serve that "catchment area" and relieve the pressure on London/Heathrow and London/Gatwick to the west and south of London. They created capacity for between 10 and 15 million annual passengers. However, London/Stansted served fewer than 5 million annual passengers for most of its first decade. Half its capacity, one of its two midfield concourses, was unneeded during this time. Nearby travelers would systematically bypass London/Stansted to catch flights from the other London airports because airlines provided service at those airports rather than at London/Stansted. The passenger and airline decisions to avoid London/Stansted reinforced each other.

These examples show that the development of an airport in a multi-airport system requires airlines willing to serve that facility. Passengers do not simply follow the easiest path to an airport. They will not flow like drops of water, from a "catchment area" to the most accessible exit. Passengers go to an airport to catch flights to specific destinations at an acceptable price. If these services are not available at an airport, they will not go there. The problem for New York/Newark and London/Stansted was that for many years airlines did not want to provide much service to these airports.

Airline strategy can and does evolve, as Sec. 5.4 explains. As of 2012, a generation after the openings that left an entire passenger building vacant for decades, New York/Newark was close to the busiest airport in its region and London/Stansted was at its original capacity. In both cases, this evolution was due to airline strategies unanticipated by the original planners.

Difficulty in Closing Old Airport

Developers of major new airports often assume that they will be able to close the older convenient airport and avoid having two airports active simultaneously. This is possible—with both luck and careful planning. Denver closed the convenient Stapleton airport when it opened Denver/International, and Greece turned its close-in Athens airport into an Olympic site. However, it often happens that economic and political pressures intervene to keep the convenient older airport open, despite assurances made by authorities in the original planning process, some 5 to 10 years earlier. Airport planners need to anticipate this possibility.

The case of Osaka illustrates this point. The government built Osaka/Kansai on a man-made island far from the center of Osaka and affordable housing. Its original plan was to close the older Osaka/Itami airport that was congested, in the middle of an urban area, and distributed considerable noise and dirt over its neighbors. However, the old airport is also much more convenient, for both its passengers and workers. Their political pressures kept Osaka/Itami open and it is still one of the busiest airports in the world. This has been costly for the new Osaka/Kansai airport, keeping its regular traffic and the landing charges high (it has been about $7000 per operation!). Also, because the runways at the two airports are almost at right angles to each other, their simultaneous operation complicates flight paths. Appropriate contingency planning might have avoided these difficulties.

This kind of difficulty is not unique. At Edmonton for example, the government built a new airport for long-distance service associated with the city. However, passengers continued to use the downtown airport for long-distance trips for years after the new airport opened. Until this closed to airline service, travelers used the frequent shuttle to Calgary (a city about 1 hour away by air) to connect to excellent long-distance service there.

Likewise, when the operator of the Milan airports opened their major new facilities at Milan/Malpensa, it attempted to close the older Milan/Linate airport to all non-Italian carriers. This policy would have given Alitalia a virtual monopoly on that airport, which is more convenient to downtown Milan. The competitive airlines, backed by their passengers, protested vigorously and managed to stay at Milan/Linate.

The lesson from these examples is that markets attempt to maintain operations at older, more accessible facilities. Often, they succeed despite governmental and other commitments. The fact is that the authorities in charge when the new airport opens are generally not those who made commitments during its planning, about a decade earlier. Moreover, they inevitably confront new realities—for example, that the access to the new airport is inadequate because a highway or railroad is incomplete or not yet built. Good planning will recognize the likelihood that older facilities often do not close as planned. Good planners will recognize their limited ability to make passengers go where they do not want to go.

Insufficient Traffic Overall

All metropolitan areas generating more than a threshold of traffic feature a multi-airport system, as Table 5.1 indicates. As Sec. 5.2 explains, these regions have enough traffic to sustain two significant airports at the same time. Conversely, regions with less than the threshold amount of traffic may have difficulty sustaining two airports.

Metropolitan regions with less than the threshold amount of traffic will be able to maintain two airports when there are technical or political reasons that compel these airports to exist. For example, both Taipei and Buenos Aires have multi-airport systems although their current originating traffic is far below the prevailing threshold. This is because their convenient downtown airports (Taipei/Sung Shan and Buenos Aires/Aeroparque) simply cannot handle transoceanic aircraft, so that traffic must go to the alternative airport (Taipei/Taoyuan and Buenos Aires/Ezeiza). Theoretically, the airport operators in those cities could close the older, convenient airports. However, such moves would certainly be unpopular with the airport users and difficult to sustain absent compelling reasons. Meanwhile, the split between the airports disrupts international airlines and trade. For example, the most convenient connections between the rest of the world and Mendoza and other Argentine provinces often pass through Santiago, Chile, rather than through Buenos Aires.

Metropolitan regions with less than the threshold amount of traffic and no compelling reasons to have two facilities have great difficulty sustaining both airports. Montreal is the prime example. Its convenient older airport, Montreal/Trudeau, is fully capable of handling transoceanic aircraft. The region never had enough traffic to sustain two major airports; in 2011 it only generated about 13.6 million total passengers. When the authorities developed the huge Montreal/Mirabel airport in the 1970s, they thought they could force international traffic to use it. However, many passengers avoided the inconveniently distant Montreal/Mirabel by taking flights to Toronto and then proceeding on to Montreal/Trudeau. Airlines scheduled flights to serve this alternative pattern, offering fewer flights to Montreal/Mirabel and further weakening its position. This diversion of traffic was bad for Montreal and the operation of Montreal/Mirabel was highly uneconomical. Correspondingly, once the authorities established the Aéroports de Montreal as the commercially oriented airport operator for the region, it effectively closed Montreal/Mirabel and moved international operations back to Montreal/Trudeau.

This experience is a warning to airport operators seeking to establish a major second airport before the traffic is sufficiently high, when they otherwise do not need to do so for technical reasons. For example, planners for Lisbon and Chennai should be cautious, so long as their traffic is below the current threshold of about 15 million originating passengers. For example, they might focus on securing a new site for an airport, without actually committing to its construction. This was the strategy Bangkok, Sydney, and Toronto adopted, along the lines suggested in Chap. 4. In effect, they took out a "real option" to protect their future (de Neufville, 1990, 1991). Bangkok eventually exercised its option and opened Bangkok/Suvarnabhumi on land reserved some 40 years earlier. Planning for second airports should recognize the risks and deal with them by securing sites while deferring commitments until sufficient demand exists.

Impractical to Allocate Traffic

Many airport operators have tried to force passengers and traffic to move from a busy primary airport to a secondary airport with under-used capacity. Their motivation is straightforward: moving traffic from a crowded to an uncongested airport should reduce congestion and delays, make better use of the existing facilities, and perhaps avoid further capital investments. With few exceptions, these attempts have been futile.

In the United States, the national government tried unsuccessfully for years to move traffic from Washington/Reagan airport to their new Washington/Dulles airport. To that end, in 1981 the Federal Aviation Administration (FAA) designated Washington/Dulles as the international airport for the capital and limited direct flights from Washington/Reagan to airports within 1000 miles (1600 km).[7] These restrictions did not succeed in forcing either passengers or airlines to move substantially to the distant Washington/Dulles. Airlines scheduled departures from Washington/Reagan to London and Tokyo by the simple device of changing aircraft at intermediate points such as Boston or Chicago. In the early 1990s, almost 20 years after Washington/Dulles had opened, the Official Airline Guide showed more international departures to London and Tokyo from Washington/Reagan than from the supposed international airport. Additional flights went overseas from the Baltimore/Washington International airport. As for domestic flights, the airlines and passengers evaded the spirit of the restrictions by scheduling flights from Washington/Reagan to San Francisco, say, via intermediate stops. Moreover, politically influential cities such as Chicago, Los Angeles, and New Orleans obtained exemptions. In the end, the governmental restrictions did not force traffic to grow at Washington/Dulles. Only when United Airlines decided to make that airport one of its hubs in the 1990s did traffic at Washington/Dulles grow rapidly.

The British government continuously tried to move traffic out of London/Heathrow and over to London/Gatwick. It unsuccessfully attempted to persuade travelers within Britain to shift their travel patterns, by offering $30 discounts (in terms of year 2012 dollars) in the regulated fares. It pressured foreign countries to have their national airlines fly into the secondary airport but most successfully resisted second-class assignments to the less popular airport.

Only the strongest government pressures can compel the allocation of airlines and traffic between airports. Thus, the Japanese government closed Tokyo/Haneda and Osaka/Itami to international traffic, forcing service beyond Japan to go to Tokyo/Narita or Osaka/Kansai. (This situation was highly inconvenient and the downtown Tokyo/Haneda airport is now again open to international traffic.) The French government largely developed Paris/de Gaulle by compelling Air France (which it owned) to move all its operations from Paris/Orly. This move imposed enormous costs on the airline. It had to build and operate duplicate facilities at both airports. For most of the next 20 years, it lost substantial traffic to foreign competitors who continued to operate at Paris/Orly, which remained the primary airport and had the best connections throughout France. Only an airline with generous government financial support could persist in the face of such long-term economic adversity. Today's investor owned airlines will not be able to comply with directives that go against the market forces.

The general rule is that market dynamics ultimately prevail. Government efforts to force traffic shifts between airports are impractical, except in limited circumstances. The emphasis is on the dynamics of the market. The outcomes are often unexpected.

Volatility of Traffic at Secondary Airport

Traffic at secondary airports is typically much more volatile than at the primary airports with the most traffic. One explanation for this phenomenon is that secondary airports relieve the congestion at the primary airports. Their traffic grows in boom periods but falls back when traffic returns to the primary airfield during recessions. Another is that a secondary airport often is a base for a startup airline. As these ventures often grow rapidly and then collapse, so does the traffic at the secondary airport. For these and other reasons, traffic at secondary airports often grows and falls rapidly. This feature makes planning difficult and investments risky and potentially unprofitable.

The example of Chicago/Midway suggests the point. As of 1987, it was also the hub for Midway Airlines, a startup that chose to operate out of this secondary airport rather than compete with United Airlines at its major hub of Chicago/O'Hare. When Midway Airlines failed in 1992, traffic at Chicago/Midway dropped over 40 percent, from its high of about 3.2 million enplanements. Within the following 7 years, however, the traffic tripled to 6.2 million enplanements. These kinds of rapid changes make it difficult to create the infrastructure needed when the traffic does occur—and to pay for facilities that have become relatively empty when traffic collapses.

The experience of Chicago/Midway is not unique. Sudden spurts or drops are the common experience of small secondary airports. The traffic at Boston/Manchester, for example, grew gradually for a decade and then at over 50 percent a year for the next 2 years as a Southwest Airlines moved in. Between 1995 and 1997 Orlando/Sanford grew from 50,000 to over 1 million total passengers. On the other hand, its traffic fell by a one-third between 2009 and 2010, from 1.7 million to under 1.2 million.

Statistically, the traffic at the individual airports in a multi-airport system is much more volatile than it is for the region. Moreover, airline traffic worldwide has become more changeable due to deregulation of air transport industry, as Chap. 4 explains. These facts mean that it is more difficult to plan and manage a multi-airport system than a single airport for a city. The traffic varies more; planning has to respond more quickly to these rapid changes, and investments are more risky and difficult to justify.

Overall Perspective

Premature ambition to create a major new airport is the cause of many of the difficulties with the development of a second airport. Airport operators worldwide have built facilities at second airports that proved to be too large for many years. They counted on being able to move traffic from the crowded primary airports to the new facilities and were not able to do so. The volatility of the traffic at second airports worsens their difficulties by complicating planning and hurting investments. The next section explains why the difficulties are inevitable. As Sec. 5.5 indicates, airport planners and operators should develop second airports flexibly and incrementally, so that they can avoid these premature, unwise investments.

5.4 Market Dynamics

Concentration due to Sales Opportunities

Market dynamics in the airline/airport industry lead to concentrations of traffic at specific airports. This concentration is a specific manifestation of a widespread phenomenon. It needs emphasis because airport planners might not appreciate this important effect. As a leading researcher of international competitive markets wrote

> …what is less understood is how prevalent [geographic concentration] is. British auctioneers are all within a few blocks in London. Basel is the home base for all three Swiss pharmaceutical giants. In America, many leading advertising agencies are concentrated on Madison Avenue in New York City… General aviation aircraft producers are concentrated in Wichita, Kansas… (Porter, 1998)

This concentration effect contrasts sharply with the concept that passenger traffic flows to airports from their catchment areas, defined as those areas that are most accessible to the airport. The model of catchment areas derives from the earliest explorations of location theory (Weber, 1929). It is widely applicable to many situations in which transport costs are most important. However, it does not apply to companies such as airlines, for whom sales and profits depend on their locations. As a leading later researcher pointed out, “Weber’s solution for the problem of location proves to be incorrect as soon as not only cost, but sales possibilities are considered” (Lösch, 1973). The important point is that the “catchment area” notion does not apply to airports. The focus must be on concentration effects.

Airlines concentrate their activities to avoid giving their competitors a decisive advantage in the marketplace. Airlines make investments and deploy aircraft as strategically as they can. Their goal is to get the largest and most profitable shares of the market. Meanwhile, their competitors do their best to counter these moves. The dynamics of this competition leads to concentration. The parable of the ice cream sellers on the beach illustrates this kind of behavior between economic entities; see Example 5.1. The situation for the airlines is similar in concept, although the motivation for the concentration is different.

Example 5.1 The parable of two ice cream sellers on the beach illustrates how sales possibilities can override cost considerations and impel the concentration of economic activities. Suppose that sellers A and B serve a 1000-m-long beach with potential customers spread equally along the shore. If we consider only the cost of travel, the optimal location for the sellers is 250 m and 750 m from the end of the beach. This placement minimizes both the maximum and average distance customers might have to go (250 and 125 m). Thus

|---------------A---------------'---------------B---------------|

In this situation, seller A can increase his market by moving to the center. He will then be closer not only to all the people on the left-hand side but also to some who would otherwise be closer to B. If seller B does not move, A will be more convenient for 625 m or five-eighths of the market. Thus

|---------------'---------------A---------------B---------------|

Seller B’s logical response is to move to the center. She can recapture the potential customers lost to A. When both sellers are at the center, neither can gain any competitive advantage by moving and both will have equal access to the market. Thus

|---------------'---------------A,B---------------'---------------|

Concentration at the center is a stable solution, in contrast to the original location that allows either seller to gain a competitive advantage by moving toward the center. From the customers’ perspective, however, this is an inferior solution—their maximum and average distances (500 and 250 m) are twice those associated with the original solution.

The moral of this story is that geographic concentration occurs in a market because participants recognize the importance of sales possibilities.

Airlines Concentrate on Routes

The primary factor that impels the concentration of traffic for airlines is the S-shaped relationship between an airline’s share of the market on a route and the frequency of service it offers. It applies to the extent that all other factors—such as fares—are equal. Figure 5.2

sketches the curve for two airlines operating in a market. Three points anchor the function. First, if the airlines offer identical frequency and service, they will each have half the market, as indicated by the mark in the middle of the sketch. Second, if one of the two airlines withdraws, it has no frequency and no market share. Complementarily, the competing airline will be offering all the frequency of service and have all the market. The latter two points are marked at the end of the dashed line diagonally across the sketch. The crucial factor is the S-shape between these extreme situations.

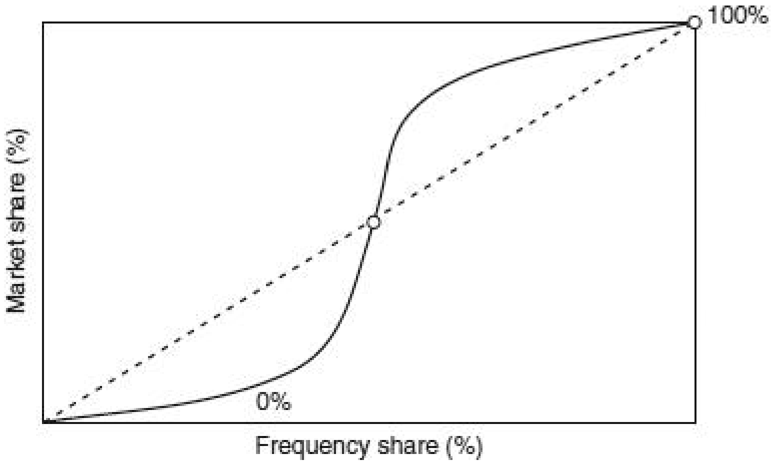


FIGURE 5.2 S-shaped relationship between frequency share and market share for two airlines operating in a market.

Empirical studies have shown that when two airlines compete on a route, the airline with the greater frequency of service gets more than its share of the market—all else such as fares and size of aircraft being equal (Fruhan, 1972). This means, for example, that if one of two airlines offers 60 percent of the flights along a route, it may get 65 or 70 percent of the traffic. Correspondingly, the airline with the 40 percent frequency share might only get 30 to 35 percent of the traffic. The reason for this is simply that passengers will go to the airline that has the most departures for a destination, is therefore more likely to provide service when desired, and has more backup in case delays or other setbacks occur. This simple fact has tremendous implications for the profitability of the airlines, and thus for their behavior, as Example 5.2 shows. (See discussion in Chap. 2.)

The S-shaped nonlinear relationship between frequency share and market share motivates airlines to match frequencies in a market unless they have some particular competitive advantage. If they cannot match frequency, they may only serve a route occasionally for special reasons—for example, as part of a large route or an extension of a continuing flight—or they may exit a route altogether. In 2001 for example, Delta Air Lines abandoned the shuttle service it had been operating from Washington to Boston. Its competitor, US Airways, had many more flights and was the airline travelers flocked to when they wanted convenient departures.

Airlines Concentrate at Primary Airports

The matching behavior on routes has important implications for airline and passenger traffic at airports in a multi-airport system. In this case, there are two S-shaped relationships at work. One concerns the airlines. The other concerns the airports. Just as the airline with the greater service in a market attracts passengers who appreciate the convenience of more departures and return flights, so the airport with the greater frequency attracts more passengers in a market—all else being equal.

Example 5.2 Consider a route that has two airlines operating 100-passenger aircraft. Suppose that there is enough traffic to fill 70 percent of the seats on 20 daily flights, that is, 1400 passengers. Suppose further that the breakeven load factor for the shuttle service is 65 percent. If either airline has a lower load factor, it loses money. If it has a greater load factor, it makes a profit.

If both airlines offer the same frequency of service and split the market evenly, they each offer 10 daily flights and carry 700 passengers. Each then has a profitable load factor of 70 percent and makes money.

Now assume that one airline manages to offer 60 percent of the frequency, that is, 12 flights out of the 20. Assume further that, according to Fig. 5.2, it then gets 65 percent of the market or (0.65)(1400) = 910 passengers. Its load factor is then:

$$\text{Load factor for more frequent airline} = \frac{910}{12} = 75.8 \text{ percent}$$

Meanwhile, the situation for the competitor is disastrous. With 40 percent of the frequency, it offers 800 seats a day yet carries only 35 percent of the traffic, that is, 490 passengers. Its load factor is ruinous:

$$\text{Load factor for less frequent airline} = \frac{490}{8} = 61.2 \text{ percent}$$

Airlines cannot afford to fall behind on the S-curve. They avoid this situation by matching frequency of service on a route. This effort is conceptually identical to the behavior of the hypothetical ice cream sellers on the beach (see Example 5.1).

The dynamics of the competition between the airlines serving a multi-airport system leads them not only to match their flights but also to place them preferentially into the airports with the greater traffic. Any extra flight that they can allocate to the airport with the most traffic helps their sales. It will either match a flight of their competitors and protect their share of the larger market, or give them an advantage in this larger market (de Neufville and Gelerman, 1973, demonstrated how this works in detail). Although the airlines might provide more convenient service overall at less trouble to themselves if they split their flights proportionally between the airports in a multi-airport system, they do not do this in a competitive economy.

The competitors' attempts to gain an edge lead them to a competitively stable position. They tend to concentrate flights at the primary airport in any multi-airport system. This explains the observed pattern of concentration of traffic at primary airports (see Table 5.4).

Factors Favoring Multi-airport Systems

The analysis based on frequency share has limits. These define the principal conditions that enable secondary airports to develop. They are that

- The assumption that the airlines and airports operating in the same "market" does not always hold—they may serve distinct markets defined by quality or fare differences
- It is often not true that "all else is equal,"—the airports in the multi-airport system may offer different prices, destinations, or quality of service
- There are limits to the value of increased frequency of service in terms of attracting passengers, which appears to define the threshold for the meaningful operation of secondary airports
- Airports in the system do have geographic advantages
- Technical and other necessities

Secondary airports typically develop around specialized airlines that operate in markets different from those of the airlines at the primary airports (Table 5.5). Most frequently, low-cost airlines appeal to a different range of passengers (and thus a different market) than the legacy airlines that operate at the primary airports. In their case, the assumption that "all else is equal" does not hold, and the arguments concerning frequency are not decisive. As of 2012, Ryanair in Europe is a prime example of a low-cost airline with the explicit strategy of implanting itself at small or secondary airports that are neither congested nor expensive, such as London/Stansted, Brussels/Charleroi, Frankfurt/Hahn, Paris/Beauvais, and Rome/Ciampino. Southwest in its early years followed a similar strategy in developing service at Dallas/Love Field, Miami/Fort Lauderdale, San Francisco/Oakland, and Boston/Providence.

Metropolitan Area	Secondary Airport	Airline Served
Dallas/Ft. Worth	Love	Southwest
Düsseldorf	Köln/Bonn	FedEx, UPS
Frankfurt	Hahn	Ryanair
Houston	Hobby	Southwest
London	Stansted	Ryanair
London	Luton	easyJet
Los Angeles	Ontario	UPS
Paris	Beauvais	Ryanair

TABLE 5.5 Examples of Secondary Airports Developed around Specialized Markets

Low-cost airlines have often dominated and been responsible for the success of secondary airports in large metropolitan areas. Southwest's development of Boston/Providence in the late 1990s illustrates this phenomenon. Its cheap fares attracted passengers, and traffic tripled to around 6.5 million in just 3 years. Thanks to Southwest, this regional airport, of little consequence for decades, grew to be a major second airport for the Boston metropolitan region.[8]

Integrated cargo airlines such as FedEx and UPS have also been responsible for the development of secondary airports. These carriers offer door-to-door handling of individual shipments and do not compete in the same market as the passenger airlines that carry belly cargo. Some secondary airports serve special destinations or regions. For example, Tokyo/Narita is almost exclusively is international, Osaka/Itami is a domestic airport, and Paris/Orly has traditionally served Africa and the Caribbean. None of these services competes directly with those at the primary airport. These airports serve their own markets and market dynamics do not concentrate their traffic at the primary airport.

The analysis of the dynamics of competition between the airlines based on frequency presumes that greater frequency is more attractive to their potential customers. At some point, however, additional frequency is no longer valuable. Hourly flights on a shuttle service may be enough, for example. This implies that when the traffic is high enough, airlines will lose interest in further concentration and will be willing to place additional flights in the secondary airports. Indeed, this seems to be what happens.

Thus, the examination of metropolitan regions with the largest number of originating passengers has consistently shown that, beyond a threshold of traffic, all these areas had a viable multi-airport system. As of 2012, the traffic threshold that seems to justify an effective multi-airport system is around 15 million annual originating passengers for the metropolitan region, as Sec. 5.2 indicates. This threshold has been steadily increasing. In the early 1970s, it was at around 8 million annual originating passengers. In the intervening years, aircraft became larger and airlines could handle more passengers with the same frequency. The interpretation of this evolution is that frequency becomes less important above some level, at which point a second airport can develop more easily. This level translates into a rising number of passengers, as the size of the aircraft increases.

Geographic considerations become more important when the importance of frequency diminishes. At some point, secondary airports receive substantial traffic because they are in fact more convenient. This effect is particularly significant when travel throughout the metropolitan region is inherently difficult. Hong Kong is a prime example of this situation. Although both Hong Kong/Shenzhen and Hong Kong/Macao are geographically close to the primary airport at Hong Kong/Chek Lap Kok, they are actually quite distant in time because of inadequate roads for one and a long water crossing for the other.

Finally, runway limitations may impel the development of a multi-airport system. For example, the short runways at Dallas/Love Field (< 9000 ft or 2700 m) led to the development of Dallas/Fort Worth. Likewise, Buenos Aires/Ezeiza has significant traffic because the more convenient downtown airport, Buenos Aires/Aeroparque, simply does not have runways adequate to serve intercontinental aircraft. Similar situations apply to Taipei/Taoyuan, Rio de Janeiro/Galeão, and São Paulo/Guarulhos.

5.5 Planning and Developing Multi-airport Systems

As the previous sections discuss, multi-airport systems are both

- A necessary or inevitable feature of many major metropolitan regions, because of either the high level of locally originating traffic or technical limitations of an existing airport
- A source of problems—for the airport operators due to absence and volatility of the traffic at these airports and the difficulty in paying for them, and for the airlines and the region because of the fragmentation of the traffic and the inefficiency operations

Responsible planning agencies and airport operators need to anticipate the development of new airports for many metropolitan areas. At the same time, they should proceed carefully. Specifically, they should

- Secure the possibilities of future developments as necessary, for example, by land banking sites for new or expanded airports
- Develop new facilities incrementally, in line with demonstrated traffic, rather than speculatively on the hope that traffic will move voluntarily to a second airport

- Build flexible facilities that can serve the several different types of traffic that may develop at the second airports, in light of the experience that the airlines using these facilities come and go and each has different requirements
- Work closely with airlines that target markets distinct from those served by the primary airport and that are consequently more likely to implant themselves at the second airport (de Neufville, 1995)

These recommendations should also be useful to planners and developers of major new airports designed to replace older airports. This is because the developers of major replacement airports often have been unable to close down the existing airports completely. Many new large airports intended to replace the older facilities end up being second airports, at least for a while. This was the case for Osaka/Kansai, Paris/de Gaulle, and Washington/Dulles. Airport developers should anticipate this possibility and plan accordingly.

Land Banking

This is the practice of securing land for possible future development. Properly executed, it represents a major way of implementing long-term plans for the development of new airports at a reasonable cost.

Land banking is a form of insurance. It protects the region from the risk of not being able to find a site for a future airport when needed. This risk comes from the fact that significant new airports require at least several square miles (or in the range of 1000 hectares) of vacant, reasonably flat land. However, this is precisely the kind of land that is most attractive for the expansion of the city. If the region waits to acquire land until the city has grown to the point where it needs a new airport, it may well find that no convenient sites are then available. Land banking gives the region the option of building some kind of airport when needed, without requiring the region to do so.

Land banking is relatively inexpensive. It is obviously much less expensive than buying land and then also building a major airport prematurely—it avoids construction costs that are easily 10 times the price of the land. Moreover, although the initial cost of the land may be expensive in absolute terms, it can be a good long-term investment. As the metropolitan area grows, the value of the land should appreciate. If the airport never uses the land, it will be available for other purposes such as housing and industry. For example, the Australian federal government paid about US $100 million to acquire 1700 hectares for a possible second Sydney Airport in the late 1980s. This cost only a few percent of the estimated cost for the government's alternative, the construction of a major new international airport. A generation later, this large block of property is worth many times its original price. From an investment perspective, this land banking was both inexpensive and profitable.

To be effective, land banking needs to maintain the option intends it to serve. Planners setting aside land for a future airport need to ensure that they will be able to develop a new airport at the site should they need to do so. They need to control local zoning and development to inhibit obstructions and ensure access. Importantly, they can develop or maintain a small airport at the existing site. This facility and its operations will be useful in maintaining both the principle of the airport at that site and the necessary clear zones for aircraft operations (as Chap. 9 describes). In this vein, regional planners and airport operators should try to maintain existing airports in a metropolitan region, as insurance against future needs. For that reason the Boston airport authority, Massport, has been subsidizing the continued operation of Boston/Worcester, which otherwise might have closed due to lack of traffic.

Conversely, land banking will fail if planners do not maintain the option to develop the site as an airport. For example, the Toronto region acquired a 7500-hectare site at Pickering for a possible second airport. For nearly 40 years, this area has effectively been a vast nature preserve. The Pickering site may thus no longer available for a major airport. In practice, the second Toronto airport is more likely to develop at Toronto/Hamilton, an airport that has been in continuous operation and that, as of 2012, is a major center for integrated air cargo carriers.

Effective land banking often involves the maintenance and eventual recycling of military airports. Examples include Austin (Texas) and London/Stansted. Future opportunities in this line exist for Washington (Andrews Air Force Base), Tokyo (Yokota), and other cities.

Incremental Development

In developing new airports, airport operators should stage development incrementally, along with the actual traffic at the new facility. They will save money and be able to build the right facilities for the traffic that eventually occurs. Financial and operational problems arise when the airport operator constructs a first stage of development that is far too big for the traffic that actually occurs.

Airport planners should, and regularly do, plan new airports on major sites, typically much larger than the older airports. A large area gives the airport operator room to expand easily when traffic makes this desirable. A large site is a form of land banking that provides inexpensive insurance for future capacity expansion. Having the option to build large in the future is, however, very different from building large at the beginning.

Traffic generally builds up slowly at new airports, unless the airlines are compelled to move. This is primarily because it is advantageous for them to keep flights at the busy airport. So airlines will tend to move away from the established airport slowly. The experience at London/Stansted, Montreal/Mirabel, New York/Newark, and Washington/Dulles document this phenomenon. Both London/Stansted and New York/Newark, for example, had major airport passenger buildings standing empty for a decade or more. Each of these second airports experienced far less traffic than planners expected over a long period, because of the reluctance of the airlines to move to the new facilities.

Strong financial reasons reinforce the reluctance to move based on market forces. Airlines implanted at the old airport may have major investments in hangars, maintenance facilities, and other properties. They naturally resist abandoning these facilities and may find it difficult to raise the money to replace them.

Experience indicates that traffic is likely to develop slowly at a second airport. Airport operators should therefore stage development accordingly. The way the Aéroports de Paris (AdP) developed Paris/de Gaulle offers a good example of how planners can do this. It built passenger buildings in increments, each capable of handling about 10 million annual passengers. This incremental approach has meant that the airport has been under nearly continuous development for all this time. However, AdP anticipated this and designed the airport so that it could easily accommodate this construction without unduly disrupting ongoing operations.

This incremental approach to the development of a second airport has several advantages. Most obviously, it defers construction of capacity until needed. Postponing the construction and maintenance costs for several years may effectively halve the present value of the investments. Perhaps even more important, incremental development permits the airport operator to design each addition according to the changing needs of the airlines. Each of the increments of passenger buildings at Paris/de Gaulle has a different configuration, representing opportunities and requirements at the time of construction. The second stage (Terminals 2A and 2B) solved a difficulty with baggage handling associated with the first stage (Terminal 1), Terminals 2E and 2F enabled connections to high-speed rail service. Terminal 2G serves domestic trips and so on. Each stage represents an important addition to earlier facilities. Overall, the incremental approach has allowed the AdP both to save money and to keep up to date.

Flexible Facilities

Because forecasts are uncertain, as Chap. 3 emphasizes, airport operators should build flexible facilities that can accommodate a range of loads and types of traffic. This recommendation applies especially to the development of second airports, because their traffic is particularly volatile.

The traffic at second airports is variable as to both level and type of traffic. Because second airports are often bases for startup airlines, they go through boom and bust periods. Chicago/Midway went through this as Midway Airlines grew and failed around 1990.NewYork/Newark had a similar experience when PEOPLExpress grew rapidly and then collapsed in the 1980s. In both cases, passenger traffic rebounded, although with different airlines having different objectives and requirements. In other situations, the type of traffic might change significantly. At Washington/Baltimore, for instance, US Airways pulled out most of its international flights in the 1990s. Although the growth of Southwest maintained the overall level of passengers, the empty international gates were not flexible enough to serve Southwest. The airport thus had to construct a new passenger building, an expense it could have been spared if it had built flexible facilities in the first place (see Chap. 15 and ACRP, 2012).

Airport operators should therefore configure their facilities so that they can both accommodate different types of traffic and change easily to meet different needs. San Francisco/Oakland offers an example of how airports can do this. Their facilities have been inexpensive and designed to cater to domestic, international, and cargo facilities. These developments are perhaps not architecturally impressive, but they have met the varying requirements of the airlines that have come and gone from this airport over the last generation.

Careful Marketing

Airport operators should develop a careful strategy for marketing second airports to likely users. Airlines that operate at the primary airport are unlikely candidates as they normally will be reluctant to withdraw flights and weaken their position at the more important source of traffic.

Airlines or operators serving different markets are the most likely candidates for second airports. These may

- Aim at special market segments (such as Ryanair's emphasis on cheap fares)
- Cater to particular clients (Florida bound family vacationers)
- Orient toward particular destinations (such as business service to Rome out of Milan/Linate) or serve a special business center (as Houston/Hobby does for the NASA Space Center and the refineries, and London/City does for the financial center)
- Provide specialized services, such as integrated cargo (as Los Angeles/Ontario and Toronto/Hamilton do for integrated cargo airlines)

To entice airline clients to secondary airports, operators should develop facilities that particularly serve their needs. Orlando/Sanford provides a good example. In the late 1990s, the private operators of this airport aimed to attract low-cost airlines catering to tourists. They therefore made a special effort to reduce the cost of operating at that facility, particularly when compared to the primary airport Orlando/International. They built inexpensively, managing for example to construct their parking garages for about half the cost paid by Orlando/International (see Chap. 17). They pioneered the shared use of gates between international and domestic services (see Chap. 15). In short, they had a specific marketing strategy to develop traffic at this secondary airport and built their facilities for this market. They were then successful, too. The development strategy for Orlando/Sanford built the traffic at this secondary airport from about 50,000 to well over 1 million annual passengers in little more than a decade.

5.6 Take-aways

Multi-airport systems are truly significant elements of the airport industry. They thus deserve careful attention, especially because so many efforts to develop complementary new airports have been embarrassing financial failures.

The most fundamental lesson is that traffic in any air transport market tends to concentrate at specific airports. Correspondingly, it is very difficult for planners in a market economy to force traffic to use new facilities. To complicate the task of developing complementary airports, experience further shows us that traffic at secondary airports is volatile, both in terms of the level of traffic and the nature of the users.

These facts imply that airport planners and designers need to be flexible in their efforts to develop secondary airports. They should be cautious about committing too much too soon and about the type of facilities they create.

Exercises

5.1. Select a multi-airport system for which you may be able to obtain data about their operations. Describe how the traffic has developed at each airport over the past 10 years or so. How would you describe the relative size of the primary and the secondary airport(s)? What might account for any changes in this ratio?

5.2. For some multi-airport system, consider several important destinations for air travelers. What is the distribution of flights from each of the airports in the multi-airport system to these destinations? What do you observe about the concentration of flights at particular airports? To what extent do airports provide service to the same destinations? What factors do you think account for this: Special markets? Geographic advantage? Frequency saturation at the primary airport? Or some other factor?

5.3. Consider the capacity of each of the airports in the same system. To what extent does the actual traffic at the airports use this capacity? Would you say that some airports are underutilized? If so, what impact do you think this has had on the financial performance of the investments at these airports?

References

ACRP, Airport Cooperative Research Program (2012) "Addressing Uncertainty about Future Airport Activity Levels in Airport Decision-Making," ACRP Report 76, Transportation Research Board, Washington, DC. www.trb.org.

Bonnefoy, P. (2008) "Scalability of the Air Transportation System and Development of Multi-airport Systems: A Worldwide Perspective," Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Bonnefoy, P., de Neufville, R., and Hansman R. J. (2008) "Evolution and Development of Multi-airport System: A Worldwide Perspective," *ASCE Journal of Transportation Engineering, 136*(11), pp. 2010–2019, 2010.

de Neufville, R. (1990) "Successful Siting of Airports: The Sydney Example," *ASCE Journal of Transportation Engineering, 116*, February, pp. 37–48.

de Neufville, R. (1991) "Strategic Planning for Airport Capacity: An Appreciation of Australia's Process for Sydney," *Australian Planner, 29*(4), December, pp. 174–180.

de Neufville, R. (1995) "Management of Multi-airport Systems: A Development Strategy," *Journal of Air Transport Management, 2*(2), June, pp. 99–110.

de Neufville, R. and Gelerman, W. (1973) "Planning for Satellite Airports." *ASCE Transportation Engineering Journal, 99*(TE3), August, pp. 537–552.

Fruhan, W. (1972) "The Fight for Competitive Advantage: A Study of the United States Domestic Trunk Air Carriers," Graduate School of Business Administration, Harvard University, Boston, MA.

Lösch, A. (1973) *The Economics of Location* [translated from German], Yale University Press, New Haven, CT.

Porter, M. (1998) Chap. 4, "The Dynamics of National Advantage," in *The Comparative Advantage of Nations*, Free Press, New York, and Macmillan, Basingstoke, U.K.

Weber, A. (1929) *Theory of the Location of Industries* [translated from the German], University of Chicago Press, Chicago, IL.

---

[1] Analysts call these situations "games" and use "game theory" to investigate their properties. These names should not fool us. These competitive "games" represent serious economic struggles.

[2] This practical definition may not correspond to administrative norms. For example, the U.S. Census Bureau divides the San Francisco Bay region into several "primary metropolitan areas."

[3] Understandably, secondary airports do not refer to themselves in this way that suggests a subordinate role.

[4] The Airports Council International uses a different definition. It reports data on members who own multiple airports. For them, the Paris multi-airport system includes only the airports operated by the Aéroports de Paris.

[5] In recognition of the state of New Jersey, the formal names eventually changed to the Port Authority of New York and New Jersey, and Newark Liberty airport.

[6] The British Airports Authority was the government organization that preceded the privatized BAA.

[7]The Congress abolished this rule in 2012.

[8]Boston/Providence also highlights the volatility of traffic at secondary airports cited in Sec. 5.3: by 2009 its traffic had dropped one-third to 4.4 million passengers.

CHAPTER 6

Aviation Environmental Impacts and Airport-Level Mitigations

6.1 Introduction

Reducing environmental impacts while also meeting the needs of growing demand is a key challenge for the air transportation system in the twenty-first century. Remarkable technical progress over the last several decades has made aircraft significantly quieter, cleaner, and more fuel-efficient. However, continued growth in demand threatens to outpace future technical progress, while political and public awareness of environmental concerns continues to increase. Many major airport developments have been significantly delayed (or even cancelled) at great expense because of environmental issues (GAO, 2000) (see Example 6.1). It is therefore critical that airport planners, designers, operators and managers understand and mitigate environmental impacts from aviation. To address these needs, this chapter provides an overview of the key issues, mitigation opportunities, and the important environmental review processes that should be complied with during planning, construction, operation, and modification of airports.

Historically, noise has been the dominant environmental concern (partly because it is directly perceived), but other issues are becoming increasingly important, such as those shown in Fig. 6.1. Airport operations can have direct impacts on noise, air quality, water quality, and wildlife. Climate change impacts primarily arise from operations at high altitude (due to the majority of fuel burn occurring during flight), but airports have an important role to play in promoting mitigations in this area too.

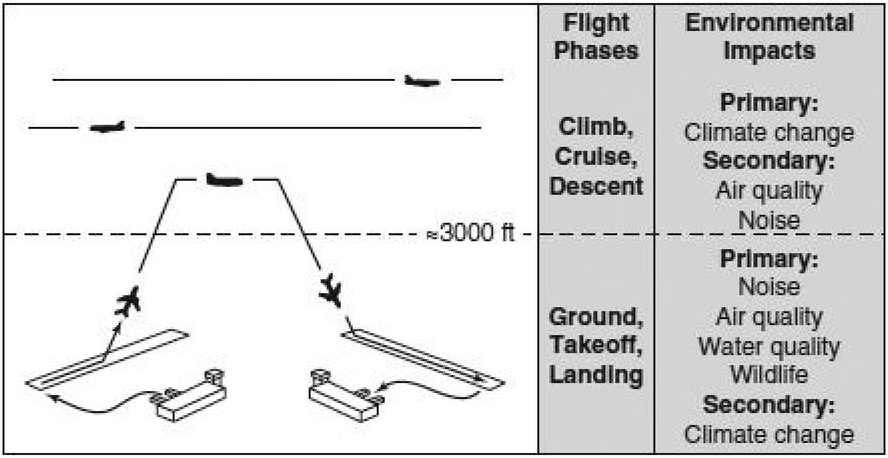


FIGURE 6.1 Aviation environmental impact areas.

Example 6.1 Airport Development/Environment Interactions At Boston/Logan, a heated controversy surrounded the construction of Runway 14/32. Some communities around the airport opposed it because of its perceived environmental impacts, whereas proponents claimed that it would provide environmental benefits by distributing noise more equitably among the affected communities, facilitating more overwater approaches to the airport and reducing congestion. The runway opened in 2006, some 30 years after Massport, the airport's operator, had first proposed it.

The possibility of a fifth passenger terminal at London/Heathrow, was first discussed in the early 1980s. Additional terminal capacity was badly needed but generated fierce opposition. Environmental impact studies, public hearings, and lawsuits delayed the opening of Terminal 5 until 2008. A third runway to increase the capacity of London/Heathrow has similarly been proposed for decades, but the support of successive British governments has wavered due to environmental, societal, and other concerns. To maintain U.K. airport capacity (and hence not lose traffic and associated economic benefits to other European countries), planners are also exploring other options, such as a new airport in the Thames estuary that has its own major environmental challenges.

In Asia, concern about noise impacts, coupled with often difficult terrain and high population densities, has led to the construction (at enormous cost) of offshore airports on artificial islands at Osaka/Kansai, Nagoya/Chubu, Kitakyushu, and Hong Kong/Chek Lap Kok (Fig. 6.2). Moreover, much of Tokyo/Haneda is built on reclaimed land in Tokyo Bay. Even though these offshore airports mitigate some environmental concerns, they create others, such as interfering with the marine environment.

FIGURE 6.2 Hong Kong/Chek Lap Kok artificial island airport. (*Source:* Wylkie Chan, Wikipedia.)

Given their general relevance to most airports, this chapter focuses on noise, air quality, climate change, water quality, and wildlife impacts, as well as mitigation opportunities relevant to airports. Depending on specific circumstances, airport planners and operators may also need to consider other areas during formal environmental review processes, such as impacts on wetlands, coastal resources, or farmland, but these are not explored in detail in this chapter.

Managing environmental impacts from anthropogenic (i.e., generated from human activity) emissions is complex because issues are experienced at local, regional, national, and international geographic scales and across timescales ranging from seconds to centuries. In the case of aviation activities, takeoff noise from a single aircraft is experienced for only a short time in a relatively small area immediately around an airport, while carbon dioxide emissions resulting from fuel burn remain in the atmosphere for centuries and potentially cause impacts on a global scale. In addition, interactions between aviation environmental impacts and other system performance metrics (e.g., environmental mitigations may have adverse consequences on throughput or vice versa) and between environmental impacts themselves (e.g., mitigations that reduce noise impacts but increase emissions) compound the challenges. The following discussions of each environmental impact area highlight some of these tradeoffs.

6.2 Aircraft Noise

Background

Noise is any undesirable or unwanted sound. During the early decades of aviation, there were few aircraft movements and hence limited aviation noise concerns. The first-generation jet aircraft in the 1950s led to a rapid expansion in commercial aviation and their engines created significant noise. The resulting severe disruption of living patterns in nearby communities prompted the establishment of formal and informal groups opposing airport expansion, drawing considerable media attention and, ultimately, government intervention. To allay public concerns in the 1960s, authorities put in place airport-specific noise limits as traffic grew at major airports such as London/Heathrow and New York/Kennedy. In the 1970s, the U.S. Federal Aviation Administration (FAA) introduced the first noise certification standards and the International Civil Aviation Organization (ICAO) promoted similar standards globally (Smith, 1989). Chapter 2 of ICAO's "Environmental Protection/Annex 16 to the Convention on International Civil Aviation" (ICAO, 2008a) defined noise standards for aircraft certified before October 6, 1977 (with some exemptions); Chapter 3 for aircraft certified between then and December 31, 2005; and Chapter 4 for aircraft certified thereafter. ICAO member states adopt these standards into national legislation, for example U.S. Federal Aviation Regulation (FAR) Part 36 Stages 2/3/4 correspond to ICAO Chapters 2/3/4. The standards outline noise limits at approach, sideline and flyover certification points (Fig. 6.3) and cumulative across all three points. All new aircraft must meet these certification standards in order to gain approval to operate. Sound level is measured in decibels (dB), and each new ICAO chapter imposes increasingly stringent noise limits, resulting in a 10-to 20-dB cumulative reduction in allowable noise. These standards have significantly driven down the noise impacts of individual aircraft of a given size over time. For example, the first-generation Boeing 747-100/200 was introduced in 1970 under Chapter 2 rules, the Boeing 747-400 in 1989 under Chapter 3 rules, and the Boeing 747-8 under Chapter 4 rules. Each successive generation of the aircraft has been required to be significantly quieter than its predecessors.

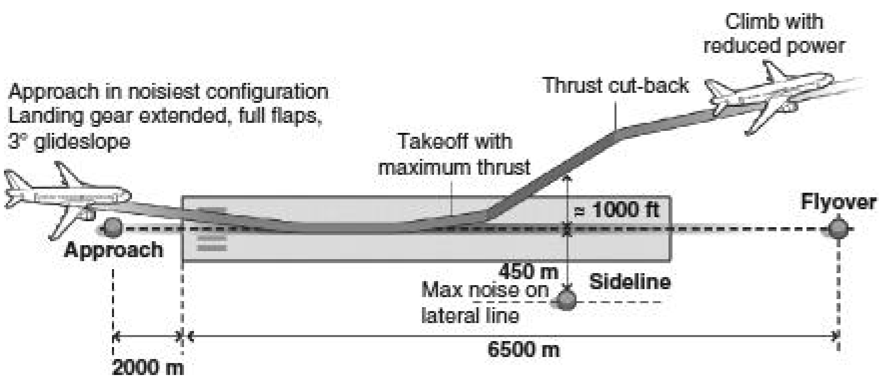


FIGURE 6.3 Aircraft noise certification points (Chapters 3 and 4).

These increasingly stringent certification standards (coupled with the other mitigations to be discussed) have dramatically decreased the number of people exposed to significant noise levels from airport operations in the last several decades. For example, from 1975 to 2005 there was a 95 percent reduction in the number of people in the United States living inside 55-dB DNL contours around airports (a noise metric discussed in detail later) (NRC, 2002). However, as technology enhancements experience diminishing returns and demand for aviation continues to grow, ICAO projects that the number of people exposed to 55-dB DNL noise will increase globally, from approximately 20 million in 2005 to 25 to 35 million in 2035 (ICAO, 2010a). As a result, airports will need to continue to address noise impact concerns.

Aircraft Noise Sources

There are two general sources of noise from aircraft: the engines and the airframe, as shown in Fig. 6.4.

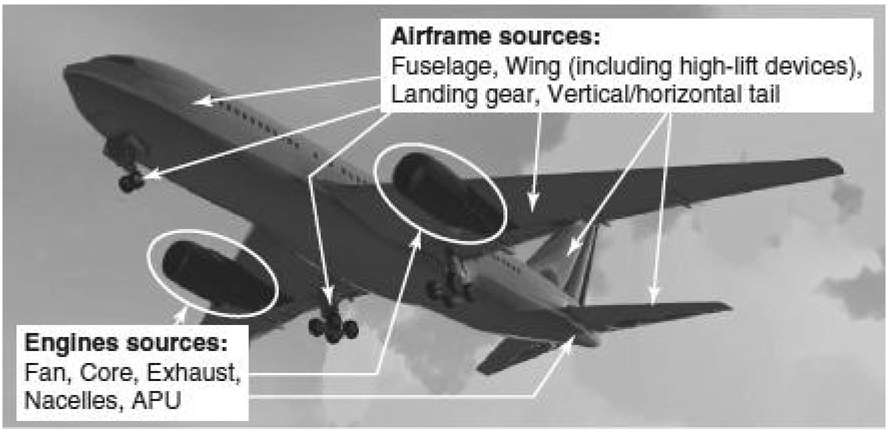


FIGURE 6.4 Primary aircraft noise sources.

Aircraft generate noise whenever there is high-speed or turbulent airflow and/or high-speed mechanical movement and rotation. Turbofan engine noise [and noise from auxiliary power units (APUs) used to provide power when aircraft are on the ground] comes from the flow of air through and rapid rotations of the various components of the engine fan and core elements, as well as the high-speed gases in the engine exhaust being expelled into the outside air. Turboprop (propeller) engine noise also includes the turbulent air shed from each blade and the interactions between the blades. Airframe noise is caused by the flow of air over the surfaces of the aircraft and the turbulent flows created by the structure and cavities introduced by the deployment of high-lift devices and landing gear. See Smith (1989) for more detailed discussion of aircraft noise sources.

Engine noise tends to dominate on the ground, especially during takeoff when the engines are at very high thrust level, on landing when using thrust reversers and when taxiing at low speed. By contrast, airframe and engine noises are about equally important during approach and landing operations when aircraft are at low altitudes in "dirty" aerodynamic configuration with high-lift devices and landing gear extended and engines at lower thrust levels than at takeoff. Another source of aircraft noise is the sonic boom created by aircraft flying at supersonic speeds which can be very disruptive to activities on the ground. This issue severely limited the market for supersonic commercial aircraft introduced in the 1970s. Only the Aérospatiale-BAC Concorde found a niche market serving transatlantic routes (overland flights were banned due to the sonic boom concerns) until its retirement in 2003 on economic grounds.

The increasingly stringent noise certification standards have spurred the development of low-noise technologies for new aircraft. These have significantly reduced noise impacts, as illustrated in Fig. 6.5. Most reductions in aircraft noise have been achieved through improvements in engine technology, especially the transition from turbojets to high bypass ratio turbofan engines. The bypass ratio is the ratio between the amount of air drawn in by the fan that bypasses the engine core relative to that passing through the core. Large modern turbofan engines have a bypass ratio of around 10:1; that is, ten times more of the air that is ingested by the fan goes around the engine core than goes through it. This configuration achieves a given thrust level with minimum size of core and the slower moving bypass air mixes with the high-speed core air, resulting in a significantly lower exhaust velocity that in turn reduces exhaust noise. Although bypass ratios have generally increased over time for modern turbofan engines, a limit is being reached which manifests as the plateauing in the noise reduction curve in Fig. 6.5. Higher bypass ratios require larger fan diameters that increase the weight and drag of the engine and thus increase fuel burn. This implies a tradeoff between environmental impacts of noise and climate change from fuel burn emissions discussed later in this chapter.

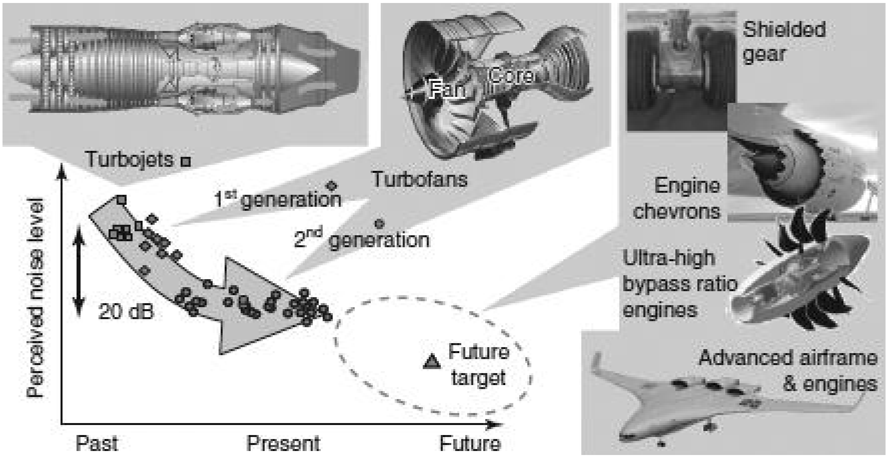


FIGURE 6.5 Aircraft source noise reduction.

Meeting future noise targets [such as the European Commission's goal for a 65 percent reduction in perceived aircraft noise level relative to 2000 levels by 2050 (EC, 2011) and NASA's long-term goal for a cumulative 62 dB reduction below Chapter 4 standards (NSTC, 2010)] will require new noise reduction technologies. Some candidate technologies are illustrated on the right side of Fig. 6.5. Near-term incremental technology enhancements include engine core and nacelle chevrons (which increase the mixing of the core and bypass air, reducing engine exhaust noise), and streamlined landing gear fairings (but these also increase weight and hence have fuel burn impacts). In the medium-term (possibly by 2020), geared turbofans and ultrahigh bypass ratio (UHBR, also called unducted fan) engines are being promoted for significant fuel savings, but their impact on noise needs to be carefully monitored. Longer-term (unlikely to be available commercially until at least 2025), more integrated airframe/engine designs afforded by blended-wing body configurations are being explored. These absorb or heavily shield engine noise, leading to significantly lower noise impacts on the ground. Their operational usability is an area of ongoing research, for example regarding airport infrastructure. It is unlikely that new-generation supersonic aircraft will reenter the commercial airline fleet in the foreseeable future, although new airframe technologies (such as low boom shaping) are being developed, which may enable smaller supersonic business jets to become a reality in the near future.

Measuring Aircraft Noise and Its Impacts

Aircraft noise propagates in the form of sound waves that travel through the atmosphere. When these waves reach the human ear, they create pressure fluctuations that are processed mentally. The wide range of pressures to which the human ear responds and the nonlinear response to pressure levels have led to the use of a logarithmic scale for quantifying sound levels. As previously introduced, the unit of measurement used internationally is the decibel (dB): a tenth of a bel, a unit named after the Scottish innovator Alexander Graham Bell. A sound level of intensity, I measured in dB (L_{dB}) is defined as

$$L_{dB} = 10 \log_{10}\left(\frac{I}{I_{ref}}\right) \tag{6.1}$$

where I_{ref} is the sound intensity at the threshold of hearing for the healthy human ear, which by convention equates to 0 dB. The range of sound levels perceptible to the human lies in a range of roughly 0 to 120 dB: those just above 0 dB are barely perceptible by the most sensitive ears in a perfectly quiet environment, whereas those above 120 dB lie at the threshold of causing pain and physical injury to the ear.

Common everyday events mapped to the decibel scale are illustrated in Table 6.1, along with their intensity ratios and approximate perceived loudness ratios. The formulas for determining perceived loudness are complex and vary significantly with sound characteristics such as frequency (Smith, 1989). As a rough rule of thumb, the human ear perceives an increase of 10 dB in sound level as approximately twice as loud. Noise events within 2 miles of a major airport when under the flight paths from aircraft taking off and landing typically fall within the 70-to 110-dB range, depending on aircraft type, exact location, and atmospheric conditions.

Sound Level (dB)	Relative Intensity (I/I_{ref})	Approx. Perceived Loudness Relative to 60 dB	Typical Event
0	1	1/64	Threshold of hearing
10	10	1/32	Rustle of leaves
20	100	1/16	Background in recording studio
30	1,000	1/8	Quiet rural nighttime
40	10,000	1/4	Quiet suburban nighttime
50	100,000	1/2	Quiet urban nighttime
60	1 million	1	Normal speech 3 ft away
70	10 million	2	Busy office
80	100 million	4	Urban daytime
90	1,000 million	8	Truck at 100 ft
100	10,000 million	16	Power mower at 3 ft
110	100,000 million	32	Rock band
120	1 million million	64	Threshold of pain

TABLE 6.1 Sound Levels and Typical Noise Events

Although the logarithmic scale for measuring the loudness of sound is technically convenient, it causes immense confusion in informing the public about aircraft noise. When told that measurements at some location show that the noise generated on takeoff by the average aircraft has been reduced from a typical value of 100 to 90 dB, most people will (not surprisingly) interpret this statement to mean that aircraft noise has been reduced by 10 percent, when in fact the intensity of the sound has dropped by 90 percent (i.e., a factor of 10) whereas the human perceives the relative loudness has dropped by approximately 50 percent.

In addition, the decibel measurement of the sound generated by an aircraft movement does not fully characterize its impact on humans: the frequency or pitch is also important. People may perceive the loudness of two sounds with equal decibel level but different frequencies as significantly different. Although the healthy human ear can hear sounds in the general frequency range of 16 to 16,000 hertz (Hz), it is most sensitive to sounds in the range of 2000 to 4000 Hz. Measurements of the loudness of sound thus typically undergo a further calibration, resulting in an “A-weighted adjustment”, to better reflect the human response to noise in the different frequencies. In practical terms, this adjustment adds approximately 2 to 3 dB to sounds in the high-sensitivity frequency range of 2000 to 4000 Hz and subtracts a few decibels from sounds outside this range. Noise measurement devices installed around airports are designed to report A-weighted sound levels automatically. To indicate explicitly that the decibel scale has been adjusted to account for the sensitivity characteristics of the human ear, the A-weighted decibel units are denoted as dB(A) or dBA.

The most commonly used measures of airport noise can be subdivided into single-event metrics (associated with a single aircraft movement) and cumulative metrics (measuring noise from many movements over a specified time period). Audible noise generated by a single aircraft movement lasts for an amount of time T that varies from about 10 seconds to a few minutes, depending on the location of the listener relative to the aircraft and on the type of movement (approach, departure, overflight, surface movement, etc.). Analysts and regulators typically use three measures to describe single event noise:

- L_{max} (maximum sound level) measures the peak sound level reached during T. It is simply the highest reading, in dBA, recorded by a noise sensor during T.
- SEL (sound exposure level) is a measure of the total noise impact of an event by integrating the noise impacts over time T which is then normalized to a 1-second duration.
- EPNL (effective perceived noise level) is similar to SEL but accounts for the duration and tone of an event (e.g., by assigning additional weight to certain discrete frequency tones that are particularly irritating to the ear). It is the measurement used for certification purposes, and its units in this case are termed EPNdB. Because of the complexity of its definition, the generation of EPNL estimates requires sophisticated computation. As a result, airport environmental studies typically utilize SEL to measure single-event noise, not EPNL.

Cumulative measures of noise estimate the total noise effect over multiple aircraft movements over a specified time near a particular location. They are thus more appropriate for representing the general noise environment around an airport. Their definitions attempt to capture the combined impact of the A-weighted loudness of the total individual noise events. Two cumulative measures are particularly important:

- L_{eq} (equivalent sound level) is a time-averaged cumulative equivalent sound level whose specific parameters can be adapted to a given situation. It measures noise exposure by computing the average dBA of noise per unit of time during the specified period. For example, to compute L_{eq} for a 2-hour period, the SEL of all the aircraft-generated noise events occurring during that period would be added on a logarithmic scale and the resulting total would be averaged (i.e., spread equally) over 7200 seconds.
- L_{dn} or DNL (day–night average sound level) is a special case of L_{eq} for an entire day (86,400 seconds) with a 10-dB increase for nighttime (10:00 pm–7:00 am) noise to account for its greater impact at these times. Importantly, it is the standard metric of the FAA for determining the noise impacts of aggregate operations around airports in the United States.

Because cumulative measures represent average noise exposure over time, they may not be able to distinguish between quite different situations. For example, one noise event generating a painfully loud noise for a short period of time might have the same average noise over a longer period as many events each generating moderate noise. The L_{eq} value may be similar for both cases, but most people would distinguish between them. Public hearings on airport noise often bring up this deficiency of cumulative measures of noise.

In most cases, the main product of noise analyses is a set of noise contours. These are lines on a map defining the areas around an airport that are estimated to be subjected to specific levels of noise after completion of the proposed project: an example is shown in Fig. 6.6. It can be seen how noise exposure areas are impacted by arrival and departure flight patterns. Airports often publish aggregate contours annually to show their noise performance over time. Airports often use noise monitoring systems with sensors at strategic locations to assess their actual noise performance. Airports may also employ web applications to allow the public to have timely access to aircraft flight track and noise impact information, which can facilitate communication between airports and community stakeholders.

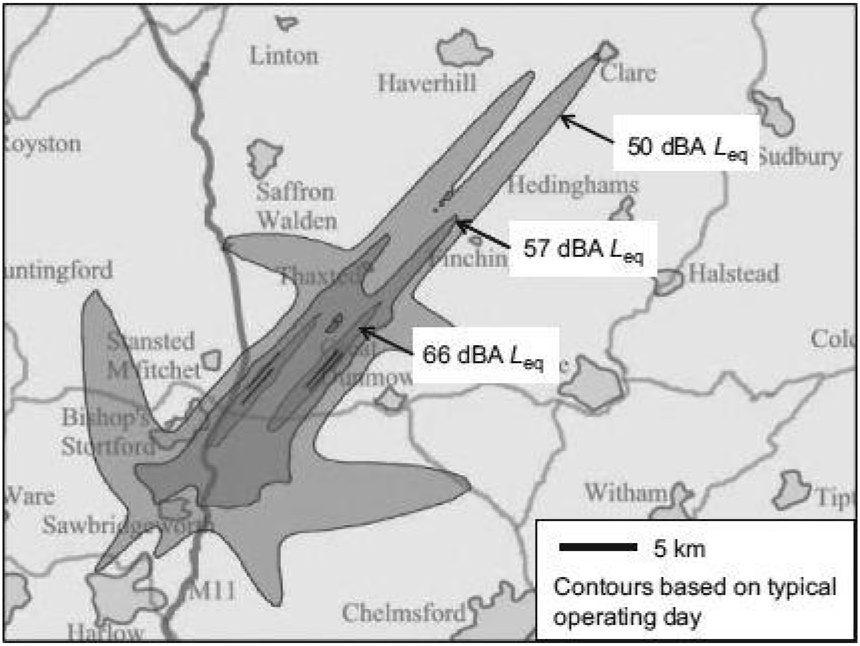


FIGURE 6.6 Sample noise contours. (*Source: NATS.*)

Noise contours are typically drawn for DNL or L_{eq} values in the range 50 to 80 dBA in appropriate increments. They can be generated by computer models, such as the U.S. FAA's Integrated Noise Model (INM)[1] and Noise Integrated Routing System (NIRS), the U.K. Department for Transport's Aircraft Noise Contour version 2 (ANCON-2) model, and EUROCONTROL's SysTem for AirPort noise Exposure Studies (STAPES). For INM, trajectory information is required in the form of aircraft type, ground track, and altitude profiles over a given time period, as well as airport characteristics such as length of runways and proportion of time in each runway configuration. The model then uses noise-power-distance (NPD) characteristics for different aircraft to calculate the noise contours on the ground. Users need to spend considerable effort preparing good-quality, location-specific inputs for INM and carefully calibrate the results with appropriate field measurements. Population distributions (e.g., from census data) can then be overlaid on the noise contours to determine how many people and properties are subject to noise of different levels. This can then be used to determine which properties qualify for noise mitigation funds for sound-proofing or relocation, as well as help land planners determine appropriate usage of certain areas regarding new development. For example, approval of noise-sensitive activities (e.g., schools, hospitals, religious institutions, and residences) would generally not be recommended in high-noise areas, but they might be acceptable for industrial and commercial purposes (“employment zones”) instead.

Past airport environmental assessments in the U.S. have concentrated on the number of people living in areas that experience noise above a certain level (e.g., DNL values of 65 dBA or higher), based on the premise that this group suffers the most and reacts most strongly to noise. This was consistent with research conducted in 1970s relating transportation noise exposure to annoyance level (Schultz, 1978) which became the U.S. government’s preferred noise impact metric based on the recommendations of the U.S. Federal Interagency Committee on Noise (FICON). However, more recent research (e.g., Fidell and Silvati, 2004) suggests that the annoyance curve is shifting such that people are becoming highly annoyed at lower DNL levels (see Fig. 6.7). For example, at the 65 dBA DNL level, the fraction of people expected to be highly annoyed has moved from 15 percent using the older data to around 25 percent using more recent studies. Although most of these studies have been conducted in the United States and Europe, and scatter in the data is relatively high, the general trends are likely to be similar in other world regions.

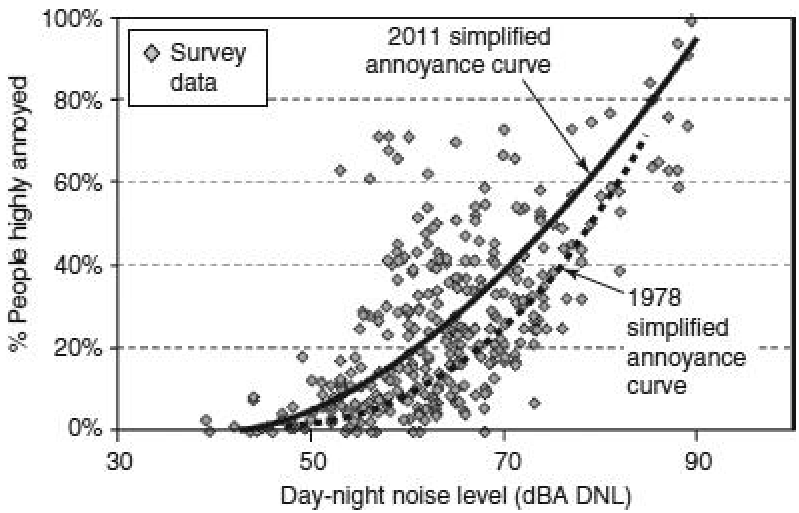


FIGURE 6.7 Annoyance level as a function of noise level. [*Source: based on data from (Schultz, 1978; Fidell and Silvati, 2004; Mahashabde et al., 2011*).]

There are extensive studies into the behavioral and physiological impacts from short- and long-term exposure to aircraft noise. Potential impacts include sleep disturbance; stress-related health effects such as hypertension, hormone changes, and mental health effects (Mahashabde et al., 2011); deteriorations in work performance; and child learning disruption. Attributing impacts to specific aircraft operational and performance parameters is challenging due to the many confounding variables such as income and dietary habits, but research is identifying some well-defined exposure-response relationships at much lower levels than 65 dBA DNL. The World Health Organization has recommended that a limit of an L_{eq} value of approximately 50 dBA (16-hour time base) in exterior sound levels is necessary to avoid serious annoyance (WHO, 1999). In addition to suspected human health effects, aircraft noise also leads to monetary impacts in terms of reducing property values. This effect is commonly captured using noise depreciation indices (NDIs) that relate the percent loss in housing stock value for each dB of aircraft noise. Typical NDI values of 0.5 to 1.0 percent per dB of noise have been reported (Nelson, 2004), that is, a 0.5 to 1.0 percent loss of housing value for each dB of noise. However, significant variations exist between regions and countries and careful consideration is required in any specific analysis.

Given this information, noise levels of 55 dBA DNL and above are now becoming important for aviation impact analyses. This is consistent with maintaining annoyance levels at no more than around 15 percent in Fig. 6.7 to account for the increasing sensitivity to noise in recent years. This is significant because, not only are many more people then included, but also the communities in the 55-to 65-dBA DNL exposure zone are often wealthier (compared to the higher-noise 65 dBA DNL and above regions) and are more effective politically in objecting to airport activities.

Airport-Level Noise Mitigations

ICAO recommends a “balanced approach” to aircraft noise management (ICAO, 2007a, 2010b). This comprises the following:

1. Reductions at source
2. Land use planning and management
3. Noise abatement operational procedures
4. Operating restrictions

Noise charges are a complementary mitigation mechanism and each of these elements is examined in turn.

Reductions at Source

Reductions at source decrease the amount of noise being generated by the aircraft. They comprise the engine and airframe modifications and technology improvements previously described. Airframe and engine manufacturers are developing and implementing these improvements in response to the certification environment and airline customer requirements. The main impact for airport planners and operators is to ensure compatibility of airport infrastructure to any new airframe and engine configurations introduced in response to source noise reduction efforts. For example, future alternatives such as blended wing bodies would have to overcome significant barriers from an airport operating perspective.

Land Use Planning and Management

Land use planning and management policies should minimize the impact of any noise that is generated. They include appropriate zoning, building codes, and mandated noise disclosures in real estate transactions set by the local authorities of residential, municipal, and commercial areas around airports given the noise environment. The airport authority does not directly control these policies but needs to interact with local authorities to ensure effective implementation. In addition, the airport is sometimes required to provide sound insulation upgrades to certain properties within the highest noise contours (e.g., as described under FAR Part 150 Airport Noise Compatibility Planning Program in the United States).

Noise Abatement Operational Procedures

Airports can promote runway, taxiway, and airspace designs and associated operational procedures that minimize the noise generated and the number of people impacted by aircraft movements. Some of the best practice operational procedures (regarding noise, air quality, and climate change mitigations) are identified in Fig. 6.8. Note that many of the mitigations can help against multiple environmental impact areas, but, given the low altitudes involved, surface, departure, and approach phases are seen to be most important from an airport noise perspective.

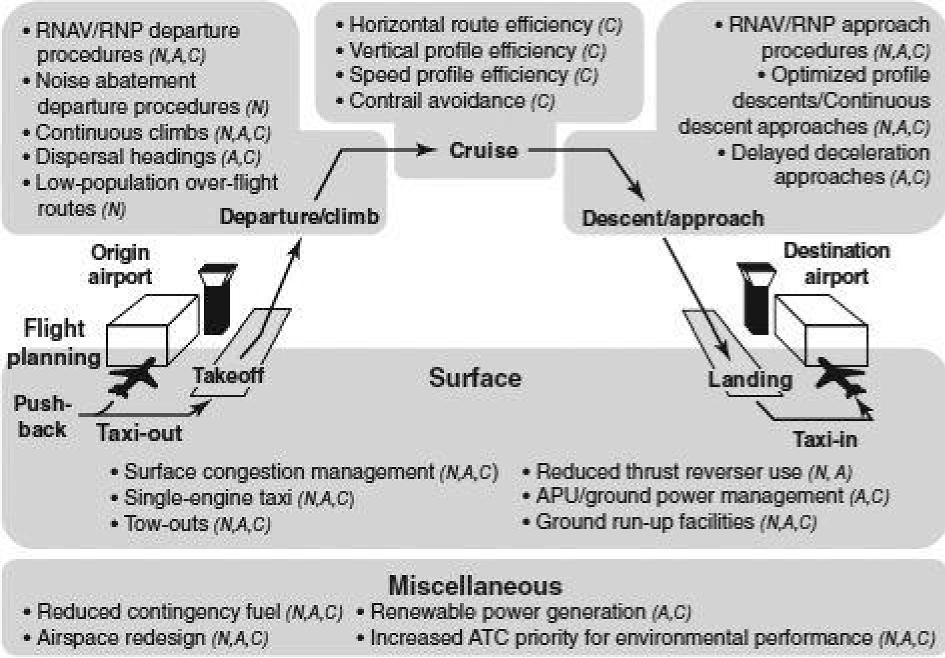


FIGURE 6.8 Sample operational mitigation alternatives to address noise (N), air quality (A), and climate change (C) impacts.

Ground operations that minimize noise include taxiing with one engine turned off (single-engine taxiing for two-engine aircraft); extended towing of taxiing aircraft by ground vehicles; minimizing APU usage (using ground-supplied power instead); limits on the use of thrust reversers on landing; and provision of ground run-up facilities with appropriate noise barriers for engine testing. The amount of time that an aircraft is taxiing and generating noise on the ground can be reduced by appropriate runway and taxi route design and assignment given the terminal and gate that each aircraft is coming from/going to, as well as surface congestion management techniques (to be discussed in detail in Sec. 6.4).

Noise abatement departure procedures (NADPs) minimize noise impacts on the ground by modifying the takeoff and initial climb phases of flight. Current ICAO guidance defines two types of NADPs that modify the thrust and speed profiles (ICAO, 2005): “NADP 1” reduces noise for areas close to the departure end of the runway by delaying the acceleration to climb speed until after 3000 ft altitude is reached; “NADP 2” reduces noise to areas more distant from the runway end by accelerating aircraft earlier to get them higher more quickly. Departures off some runways at an airport can have much lower noise impacts than off other runways (e.g., if one runway sends departures over a body of water initially, compared to over residences) and in these cases noise-preferred runway assignments may be an important aspect in the decision making of ground air traffic controllers. Planners should consider such issues in the very early stages of airport design when they configure runways, taxiways, and terminals. Many airports also have "noise containment corridors" within which departing and arriving flight tracks should be maintained (e.g., consistent with the "high-noise" corridors used by the land use planning authorities). Defining departure and approach procedures to remain within these regions is another important part of noise management.

Noise impacts for arriving flights have received greater attention than for departures because arriving aircraft spend more time at lower altitudes. Continuous descent approaches (CDAs) are the most common noise abatement approach technique. They eliminate level segments present in conventional "step-down" approaches, keeping aircraft at higher altitude and lower thrust for longer prior to intercepting the final approach glide slope, thereby reducing noise impacts (as well as fuel burn and emissions as discussed later). Figure 6.9 illustrates the basic CDA concept and shows that most noise benefits are achieved within 10 to 40 nautical miles (nm) to touchdown.

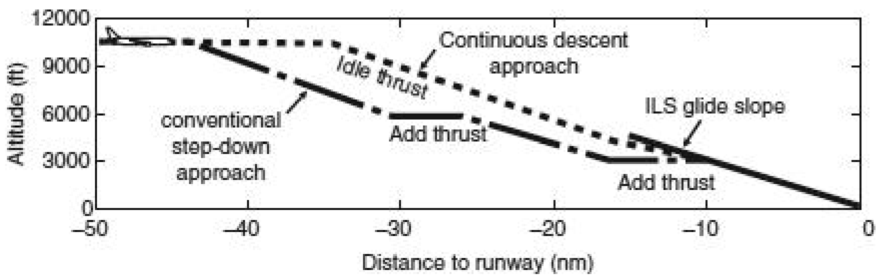


FIGURE 6.9 Continuous descent approach concept.

Proper design of CDA procedures can reduce noise impacts by several dBA outside the airport perimeter and significantly reduce the number of people within certain noise contours. CDAs can also be built into the next generation of approach procedures that take advantage of area navigation/required navigation performance (RNAV/RNP) technologies. These enable aircraft to fly more precise approach (and departure) paths, so that they can be directed over lower-impact corridors such as rivers and less populated regions (see Fig. 6.10). Noise is therefore more concentrated in these regions compared to the more dispersed noise impacts with conventional procedures. As an example, approaches to Washington/Reagan airport follow the curves of the Potomac River rather than flying directly over sensitive parts of the city. Careful consideration is required to determine whether predefined flight paths are appropriate for any given airport, for example regarding "concentrated" versus "dispersed" noise impacts and when the noise procedures result in longer flight tracks with associated higher fuel burn and emissions impacts.

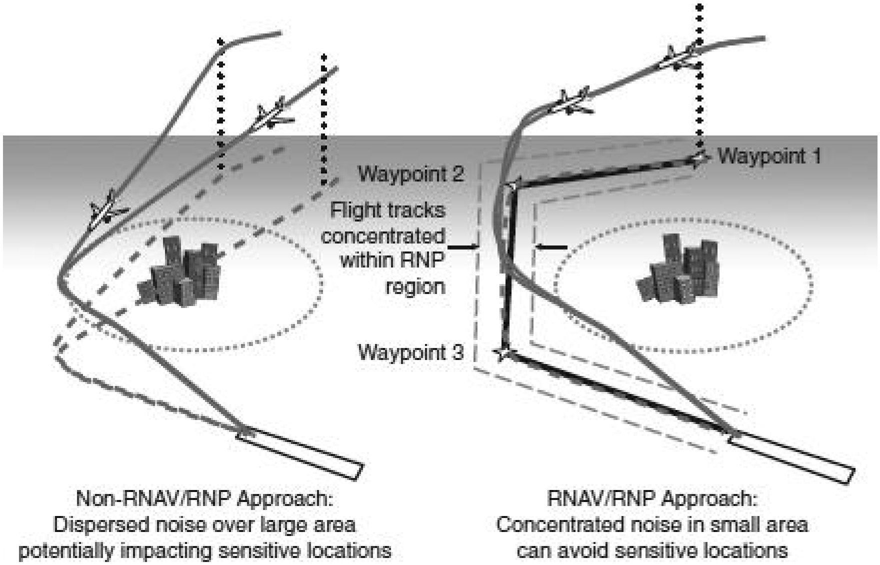


FIGURE 6.10 Dispersed versus concentrated approach paths.

Another technique to reduce noise impacts from arriving aircraft involves displacing the landing threshold further down the runway. This keeps aircraft higher outside of the airport perimeter. However, this has significant infrastructure and procedural implications and can only be contemplated on runways long enough to accommodate displaced landings without adversely impacting safety standards. Frankfurt/International is one airport where a displaced landing threshold has been studied for certain aircraft. Longer-term, steeper approach paths (above the approximate 3° flight path typically used today) are being studied. These would also achieve greater aircraft altitudes outside the airport boundary and hence reduce noise impacts, but again extensive safety studies are needed and the technology and infrastructure changes required at the airport can be significant.

Operating Restrictions

Operating restrictions reduce or limit access to given airports. They potentially provide relatively rapid and sizeable noise reductions. However, airports should carefully consider other consequences to the airport, its operators, and local regions (e.g., reduced economic benefits) before implementing such restrictions. ICAO encourages member states only to utilize operating restrictions after first applying the other elements of the balanced approach and to account for special circumstances of operators from developing countries so as to not unduly disadvantage them or where other modes of transport are unavailable.

Restrictions can take the form of outright bans, curfew limits to certain aircraft at certain times and noise quotas. Chapter 2-certified aircraft were largely banned throughout the world between 2000 and 2010. Curfew limits to certain aircraft at certain times can restrict the noisiest Chapter 3 aircraft during the most sensitive nighttime hours. Finally, noise quotas are becoming increasingly common as a way of "capping" aircraft noise impacts at a given airport to an agreed level during a certain time interval. Under these schemes, every movement by an aircraft of a given type during nighttime hours is assigned a noise value (typically based on its noise certification values) which is then subtracted from the total quota. Once the quota is reached for a given period, no further operations are permitted. In this way, airports encourage operators to fly quieter aircraft as much as possible, given that more operations of quieter types would be allowed while still complying with the quota. The total noise quota also typically decreases over time to further encourage operators to transition to quieter types just to maintain a given number of operations: see Example 6.2.

Example 6.2 Noise Quotas, Noise and Emissions Charges Noise quotas are used at several airports in Europe, including London/Heathrow, Gatwick, and Stansted, as well as Madrid and Brussels. Each aircraft type is given a quota count (QC) rating for departure and arrival operations based on their certified noise levels. For example, Airbus A380 aircraft are QC2 on departure and QC0.5 on arrival. Limits are then set at each airport on the total QC and number of movements allowed in the 11:30 pm–6:00 am period, totals which decrease year on year to encourage adoption of quieter aircraft (Table 6.2).

	2006/2007 Total Quota		2008/2009 Total Quota		2010/2011 Total Quota		Total Movements	
Airport	Winter	Summer	Winter	Summer	Winter	Summer	Winter	Summer
Heathrow	4140	5610	4110	5460	4080	5100	2550	3250
Gatwick	2300	6700	2180	6500	2000	6200	3250	11200
Stansted	3510	4900	3430	4800	3310	4650	5000	7000

TABLE 6.2 London Airport Quota Count Details

At London/Heathrow, landing charges vary by noise certification category and time of day, as shown in Table 6.3 [as of August 2011, see IATA (2011)].

Noise Certification Category		Landing Charge, £	
		Day	Night
Chapter 2	If allowed	4,912	12,280
Chapter 3	High: less than 5 EPNdB below limit	4,912	12,280
	Base	1,637	4,093
Chapter 4	High: less than 5 EPNdB below limit	982	2,456
	Base	819	2,047
	Low: QC1.0 or less	491	1,228

TABLE 6.3 London/Heathrow Landing Charges

In August 2011, emissions charges at London/Heathrow were £6.09/kg of NO_X according to the certification values (see discussions on air quality in Sec. 6.3). The noise and emissions charges are seen to be significantly less for quieter and cleaner aircraft types (landing fees are 50 percent lower for a base Chapter 4 compared to Chapter 3) and a factor of 2.5 times higher for nighttime operations compared to daytime for a given category.

Noise Charges

Noise charges are fees airports impose on operators in proportion to the noise level of the aircraft being used. ICAO recommends (ICAO, 2010c) that airports only levy charges if they have noise problems and then at a level that reasonably recovers costs incurred by noise mitigation programs (such as acoustic insulation programs) around the airport. It provides advice on determining appropriate levels for noise charges and their collection in its *Airport Economics Manual* (ICAO, 2006). In practice, some noise-sensitive airports levy charges by aircraft type against published formulas (e.g., as a function of their noise certification level [see Example 6.2]), whereas others use noise monitoring systems to measure the actual noise impacts of every movement at specific locations and charge operators as a function of these measurements or if they exceed some specified thresholds. Figure 6.11 shows the growth in number of airports worldwide imposing noise charges, together with the other types of operating restrictions. It is apparent that all types of restrictions and charges are increasing with time, showing the growing sensitivity to aircraft noise impacts around the world. Noise abatement procedures (NAPs) and curfews are the measures that have grown the most, but noise charges are also becoming increasingly common.

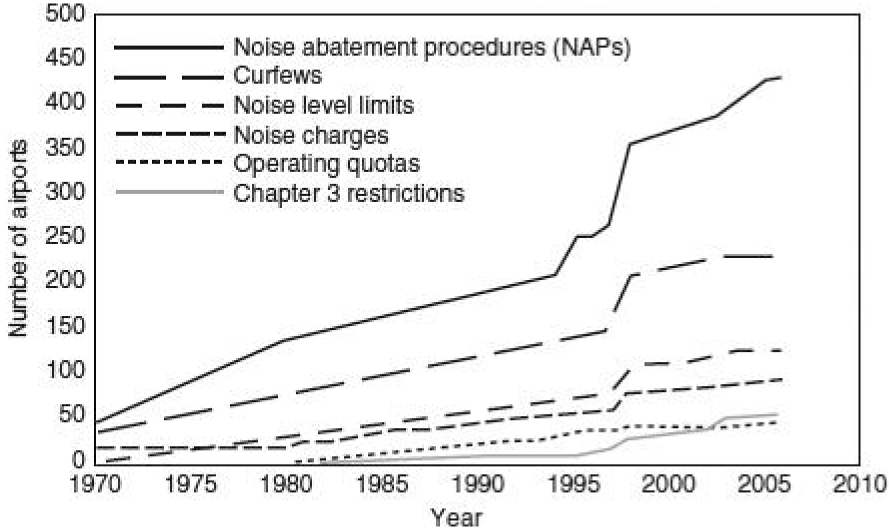


FIGURE 6.11 Growth of noise restrictions worldwide. [*Source: adapted from (ICAO, 2007a)*.]

6.3 Air Quality

Background

This section provides an overview of the potential impacts from the emissions produced by airport activities (including aircraft engine and ground support equipment fuel burn) that are of concern from a local and regional air quality perspective. The link between poor air quality and health impacts is well known, and air quality issues are becoming as important as noise impacts to local communities in some locations.

Similar to noise, ICAO sets certification limits for specific air quality pollutants that a new aircraft engine must comply with to get approval to operate (ICAO, 2008b). The standards cover hydrocarbons, nitrogen oxides, carbon monoxide, and smoke emissions. The certification test takes place on a test bed where a new engine is run at four different fractions of maximum thrust settings for specified times to simulate the various phases of a standardized landing and takeoff (LTO) cycle, as illustrated in Fig. 6.12. The LTO cycle covers the typical taxi, takeoff, and approach operations of aircraft below 3000 ft because emissions below this altitude are thought to be the primary contributors to surface air quality impacts. However, research suggests that aircraft emissions from flight phases above 3000 ft (e.g., the significant fraction of emissions during cruise flight) may constitute a substantial portion of the total air quality health impacts of aviation and this may influence how air quality certification standards are defined in the future (Barrett et al., 2010).

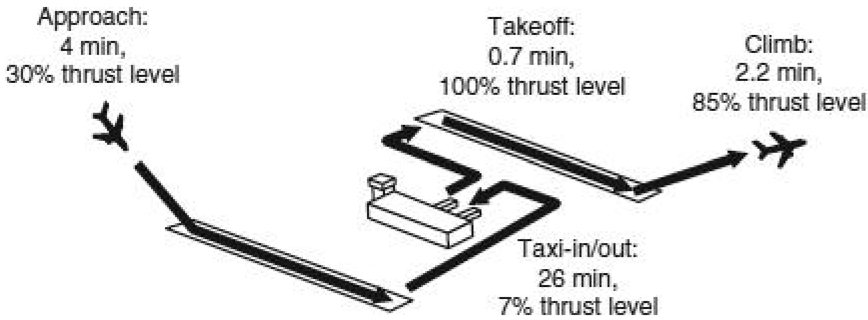


FIGURE 6.12 ICAO landing and takeoff (LTO) cycle.

Since air quality certification criteria first came into effect, the stringency of nitrogen oxide standards has reduced those emissions from new engines by around 40 percent and further reductions are expected in the coming years. Reductions in carbon monoxide and unburned hydrocarbon emissions have been equally impressive. This success has primarily been achieved through advanced engine technologies (such as combustor design), reduced fuel consumption, and modified fuel composition.

Air Quality Emissions Sources

Air quality impacts can come from many different sources and it is often difficult to determine their contributions, for example from on-airport activities compared to road traffic on a nearby major highway. Many countries have ambient air quality standards designed to protect human health and do not distinguish between pollutants from different sources. In the United States, the Environmental Protection Agency (EPA) sets the National Ambient Air Quality Standards (NAAQS) that specify acceptable levels of different "criteria pollutants" that are considered harmful to public health and the environment, including carbon monoxide (CO), lead (Pb), nitrogen dioxide (NO_2), ozone (O_3), particulate matter (PM), and sulfur dioxide (SO_2). A district meeting-acceptable levels of a given pollutant is known as an "attainment area" for that standard, while those that do not are termed "nonattainment areas" (EPA, 2011) and require special efforts to become compliant.

Aircraft and ground support equipment emissions contain many different chemical species. Carbon dioxide and water vapor are the largest components of engine emissions by mass (approximately 70 and 29 percent, respectively), but they are not a concern from an air quality perspective. The primary species that are of interest include the following criteria and noncriteria pollutants:

- *Particulate matter (PM)/smoke*: includes primary nonvolatile soot emitted from the engine as a by-product of jet fuel combustion and secondary aerosols (e.g., sulfates and nitrates) which form later in the exhaust plume through physical and chemical processes in the atmosphere. $PM_{2.5}$ and PM_{10} (where the subscript denotes the typical particle size in micrometer [μm]) are of primary interest (see the health impacts discussion in Table 6.4).

Pollutant	Health Effect
Particulate matter (PM)	• Premature mortality • Aggravated respiratory and cardiovascular disease • Lung function impairment
Nitrogen oxides (NO_x)	• Lung irritation • Lower resistance to respiratory infections
Unburned hydrocarbons (UHCs)	• Eye and respiratory tract infections • Headaches/dizziness/memory impairment
Ozone (O_3)	• Lung function impairment • Lower resistance to respiratory infections
Carbon monoxide (CO)	• Aggravation of cardiovascular disease

TABLE 6.4 Air Quality Pollutant Health Effects [*Source: adapted from (ICAO, 2010a)*.]

- *Nitrogen oxides (NO_x)*: nitric oxide (NO) and nitrogen dioxide (NO_2) form from jet fuel combustion. Levels and proportions vary significantly with engine settings and ambient conditions.

- *Unburned hydrocarbons (UHC)/volatile organic compounds (VOCs)*: C_mH_n compounds arising from incomplete combustion of fuel.
- *Ozone (O_3)*: secondary formation from NO_x and UHCs.
- *Sulfur oxides (SO_x)*: from sulfur in fuel.
- *Carbon monoxide (CO)*: from incomplete combustion of fuel.

These species are of interest because of their various impacts on human health, as shown in Table 6.4. Air quality impacts can be monetized (i.e., converted into estimated monetary terms) using value of a statistical life (VSL), willingness-to-pay (WTP) and cost-of-illness (COI) parameters that are often recommended by appropriate agencies. Premature mortality associated with the small-scale $PM_{2.5}$ is the largest monetized impact because of the very much higher value associated with mortality (death)[2] compared to the morbidity (sickness) effects from other species. That said, the health impacts of the other species are of significant community interest during airport environmental assessment studies.

Measuring Air Quality and Its Impacts

Air quality health impacts are typically estimated through concentration-response functions (CRFs). These relate the concentration of a pollutant that a human is exposed to over a certain time period to the observed health response (e.g., from epidemiological studies). Concentrations are typically measured using air quality sensors located at strategic locations around airports and supplemented with computer models of the generation and dispersion of the species of interest. Airports and local councils increasingly provide real-time (and archived) data from their air quality sensors to the public. Common sensor types include automatic monitors (which typically measure hourly pollutant concentrations from a continuous stream of air pumped through them) and diffusion tube monitors. These measure less frequently but include more species and are generally more reliable because they collect samples by chemical reaction on a filter or substrate within the tube that is then sent off to a laboratory for analysis. Guidance on collecting and interpreting aircraft gaseous and particulate matter emissions data can be found in ACRP (2008).

Computer models approved by ICAO's Committee for Aviation Environmental Protection (CAEP) to estimate air quality concentrations include the U.S. FAA's Emissions and Dispersion Modeling System (EDMS), U.K. Department for Transport's Atmospheric Dispersion Modelling System (ADMS), EUROCONTROL's Airport Local Air Quality Studies (ALAQS) model, and the Swiss/German Lagrangian Dispersion Model for Airports (LASPORT). The U.S. EPA's Community Multiscale Air Quality (CMAQ) simulation system is also being used increasingly for aviation air quality studies. Figure 6.13 shows examples of computer model and field sensor measurements of air quality impacts around Greater London. Clearly there are many contributors to the modeled air quality concentration levels, including road traffic, industry, and airport activities that are evident in the figure.

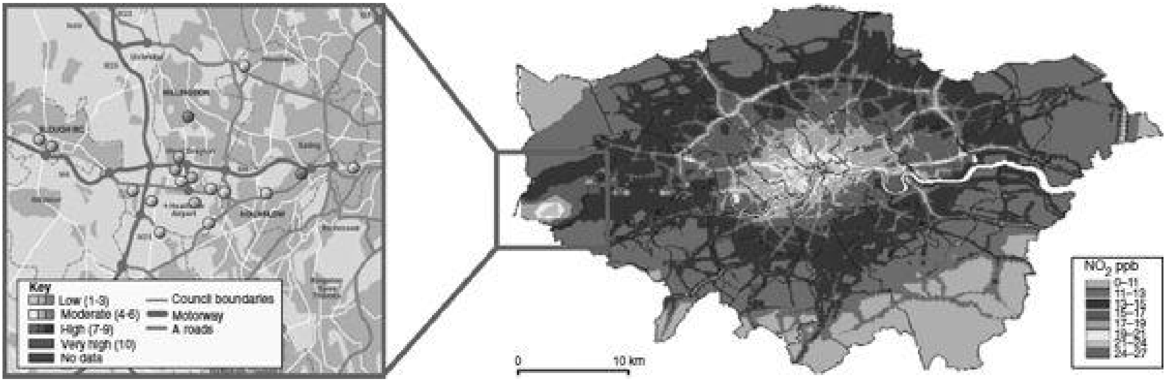


FIGURE 6.13 Air quality modeling around Greater London and London/Heathrow airport monitor status (insert). (*Source: www.heathrowairwatch.org.uk and CERC.*)

Airport-Level Air Quality Mitigations

There are ongoing activities to reduce emissions in all flight phases through more efficient aircraft operations, such as the U.S./European Atlantic Interoperability Initiative to Reduce Emissions (AIRE) and the U.S./Australasian Asia & South Pacific Initiative to Reduce Emissions (ASPIRE). As previously noted, operations on the surface and in the LTO flight phases are of particular interest from a local air quality perspective, but attention on other flight phases is increasing as our understanding matures regarding pollutant transport in the atmosphere. Airport air quality impacts can be mitigated by a number of means, including the operational procedures, emissions charges and airport authority policies discussed below. Further guidance can be found in ICAO (2007b).

Operational Procedures

The operational mitigations identified with an “A” in Fig. 6.8 are effective at reducing air quality impacts of aviation. Many of the policies previously identified for reducing noise impacts on the ground result in less engine-on time and therefore lower fuel burn and emissions on the ground. These include surface congestion management (see Sec. 6.4); single-engine taxi; extended towing of aircraft using efficient (and even electric) tugs; restrictions on when engines can be run-up for test; preferential runway assignments; airfield designs that reduce taxiing distances and time; and limited use of APUs. To reduce APU usage, many airports provide aircraft electric power and cooling capabilities at the gate that are more efficient and cleaner than APUs powering aircraft generators and air conditioning packs (Fig. 6.14). Detailed guidance on how to evaluate these strategies can be found in ACRP (2012).

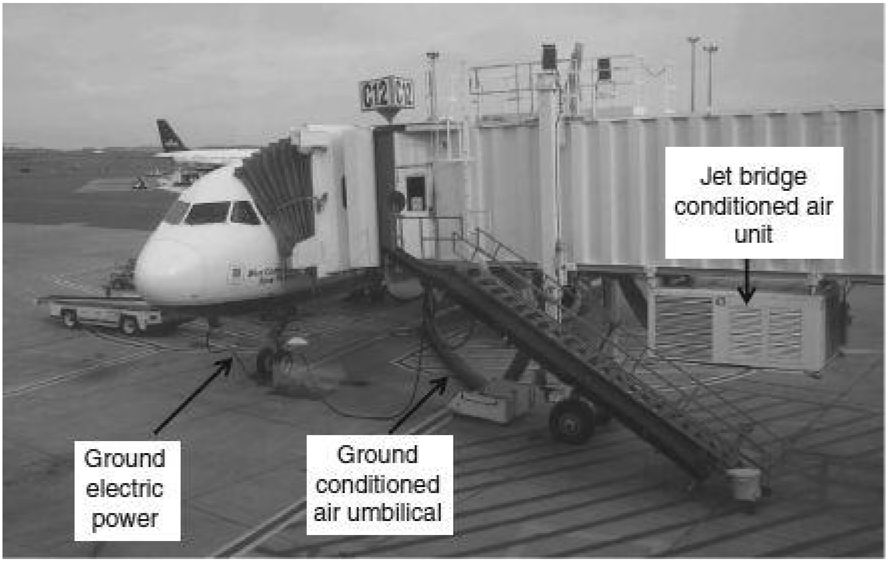


FIGURE 6.14 Ground supplied power and conditioned air.

Emissions Charges

Airports can use emissions charges to encourage operators to fly cleaner aircraft. ICAO guidance on implementing such schemes recommends that airports should only levy air quality emissions charges when they have a local air quality problem (e.g., they are in a nonattainment area), and should design charges to recover no more than the cost of measures to mitigate or prevent the damage caused by the aircraft emissions, while accounting for special needs of developing countries (ICAO, 2007c). Emissions charges are becoming increasingly common, especially in Europe. Similar to some noise charge mechanisms, they assign engine types to emissions categories and charge higher rates for more polluting categories (see Example 6.2).

Airport Authority Policies

Airports can also improve their air quality by implementing various policies to reduce emissions from their ground transportation. For example, they can encourage use of public transportation or high-occupancy, hybrid, and electric vehicles through market-based mechanisms such as reduced tolls and parking rates for suitable vehicles. They can make their own vehicles (e.g., airfield vehicles, shuttle buses, etc.) cleaner from an air quality perspective by using electric or lower impact fuels such as compressed natural gas (CNG). They can mitigate the effects of fuel vapor loss on air quality by using state-of-the-art fuel storage and distribution systems.

6.4 Climate Change

Background

This section provides an overview of the potential impacts of aviation on climate change, as well as the reverse impacts of climate change on aviation (e.g., changing weather and tourism patterns). It also indicates what mitigations are relevant from an airport perspective.

Concerns about the impacts of anthropogenic emissions on climate change have increased significantly since the 1990s, especially in terms of global warming. A wealth of literature on the subject promotes a bewildering range of opinions. In 1988, the United Nations Environment Programme and the World Meteorological Organization established the Intergovernmental Panel on Climate Change (IPCC) to independently assess the science, impacts, economics, and mitigation alternatives of climate change.[3] The IPCC has produced a series of documents that have become important works of reference, especially their regularly updated assessment reports. Their 2007 Assessment Report (IPCC, 2007) concluded that “warming of the climate system is unequivocal” and that “most of the observed increase in global average temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic greenhouse gas (GHG) concentrations.”

Transportation across all modes is a significant source of GHG emissions (Schäfer et al., 2009). Although aviation accounts for less than 3 percent of anthropogenic carbon dioxide emissions (one of the key GHGs discussed later), special attention has often been focused on the climate impacts of aviation because of its high rates of growth, the high visibility of the industry, and the unique characteristics of some of its impacts (e.g., emissions being deposited at high altitudes). The IPCC published a special report on the specific impacts of aviation on the global atmosphere (IPCC, 1999). The NO_x certification standards described under Sec. 6.3 do help reduce climate change impacts (due to the role NO_x plays in GHG concentration levels discussed later), but there are no fuel burn or carbon dioxide certification standards specifically designed to reduce GHG emissions similar to the noise and air quality areas (although they are under investigation). Despite the fact that fuel burn and GHG emissions from aircraft operations at airports are a small fraction of the total systemwide GHG emissions, airports have a highly visible role in the industry’s response to climate change. It is therefore important for airport planners and operators to understand the key issues.

Climate Change Emissions Sources

The burning of aviation fuel produces chemical species that potentially cause global temperature impacts by producing GHGs and/or aerosols. Their net effect is to trap thermal infrared radiation and thus may change the energy balance of the atmosphere. The key species of interest for climate change are the following:

- *Carbon dioxide (CO_2)*: CO_2 makes up 70 percent of exhaust emissions by mass. It is a long-lived GHG having an atmospheric residence time on the order of centuries. It is thus of particular concern with respect to climate impact potential. As a result of its long lifetime, aviation CO_2 emissions get mixed in the atmosphere around the globe and become indistinguishable from CO_2 emissions from other anthropogenic sources.
- *Water vapor (H_2O)*: H_2O makes up 29 percent of the exhaust emissions by mass and has a warming impact. However, in the troposphere (the portion of the atmosphere up to approximately 50,000 ft) where all subsonic commercial aircraft fly, water vapor only has a lifetime on the order of days and therefore has a negligible climate impact. Water emissions in the stratosphere (higher altitudes where supersonic aircraft can fly) can remain for much longer, but there are very few emissions in this region currently.
- *Nitrogen oxides (NO_x)*: NO_x gas species are not themselves GHGs, but through atmospheric processes they lead to ozone (O_3) production and methane (CH_4) destruction, both of which are strong GHGs, so these processes have warming and cooling effects. The interactions are complex with differing temporal and spatial characteristics: short-lived ozone production causes warming effects that last on the order of months and primarily influence the northern latitudes where most aircraft fly, whereas long-lived methane destruction has cooling impacts over decades and occurs on a global scale. Therefore, although the globally averaged warming and cooling impacts largely offset, there can be significant regional variations.
- *Sulfate (SO_x) and soot aerosols*: these are solid or liquid aerosols suspended in the atmosphere which can reflect sunlight (a cooling effect) or trap infrared radiation (a warming effect) depending on their characteristics (e.g., size, composition, concentration) and time of day. These aerosols have an atmospheric residence time of days to weeks, so impacts are short-lived. However, they can also act as nuclei for cloud condensation and trigger changes to naturally occurring cloud properties that may have much larger climate impacts.
- *Condensation trails (contrails)*: these are the line-shaped trails visible from the ground that sometimes form behind aircraft (typically at high altitude) under certain atmospheric and engine conditions. They often only last a few minutes and have negligible environmental impact. However, under certain conditions, the contrails can persist for an hour or more and may also produce induced cirrus clouds that can

last for days. The climate impacts of these effects is not well understood; current estimates of their impact range from negligible to being more important than the impact from carbon dioxide while they exist.

Measuring Climate Change and Its Impacts

A variety of metrics are used to measure climate change impacts. The two most common ones are radiative forcing (RF) and global warming potential (GWP). RF is a measure of the influence that a factor has in altering the balance of incoming and outgoing energy in the Earth–atmosphere system and GWP is the cumulative RF effects of an emission over a specified time horizon. RF is often used because the estimated surface temperature change due to a given pollutant is directly proportional to its RF value. However, RF does not readily account for the vastly different timescales of the impacts of the pollutants described above and GWP can be more meaningful in that regard.

Figure 6.15 shows the currently estimated RF impacts for key climate pollutants. Positive bars represent warming effects, negative bars represent cooling effects, and the black whiskers represent uncertainty bounds. This figure also lists the spatial and temporal impact scales and the level of scientific understanding as of 2005. It is apparent that CO_2 accounts for about half of the total estimated RF impacts from aviation and the level of scientific confidence is high for this GHG. NO_x and linear contrails are the next most important contributors, but for them the science is very much less certain. As the science progresses, climate impact priorities will continue to be refined.

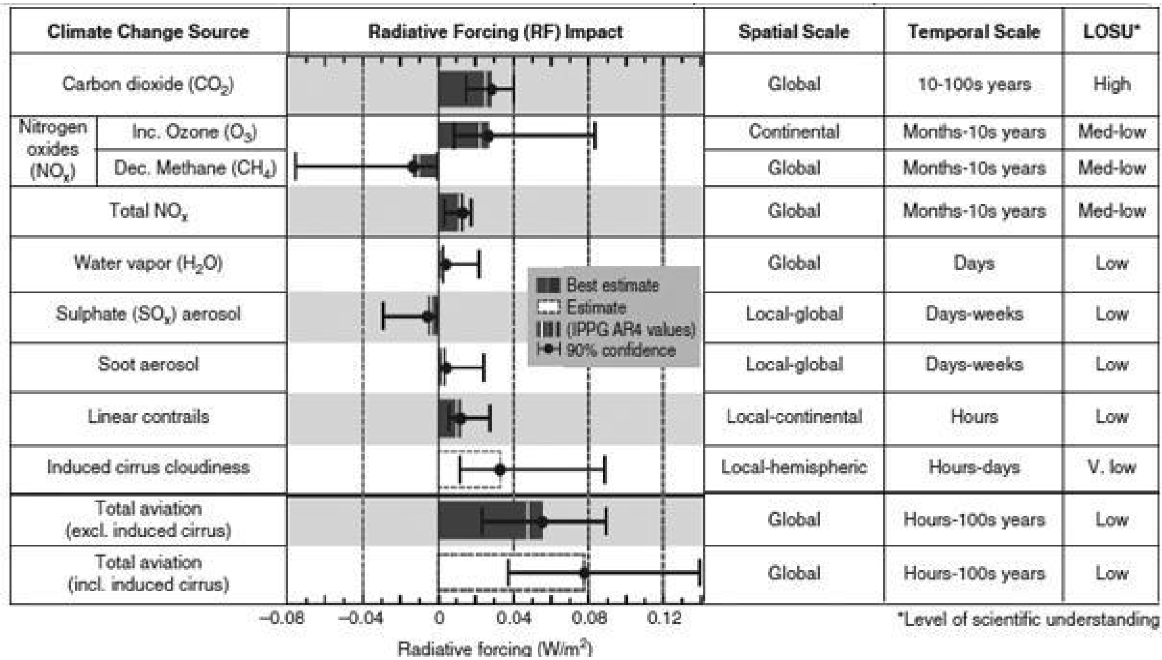


FIGURE 6.15 Radiative forcing impact estimates of aviation emissions. [*Source: adapted from (ICAO, 2010a)*.]

Figure 6.16 shows the pathway from fuel combustion to climate change impacts. The emissions species and RF impacts are evident in the middle steps of the pathway. The climate change effects include changes in surface temperature, sea level, ice/snow cover, and precipitation patterns. These have agriculture, ecosystem, energy, human health, and social consequences that can be monetized through appropriate damage functions (Mahashabde et al., 2011). As this figure illustrates, the latter stages of the pathway are most relevant from a policy-making perspective but also coincide with the greatest levels of scientific uncertainty.

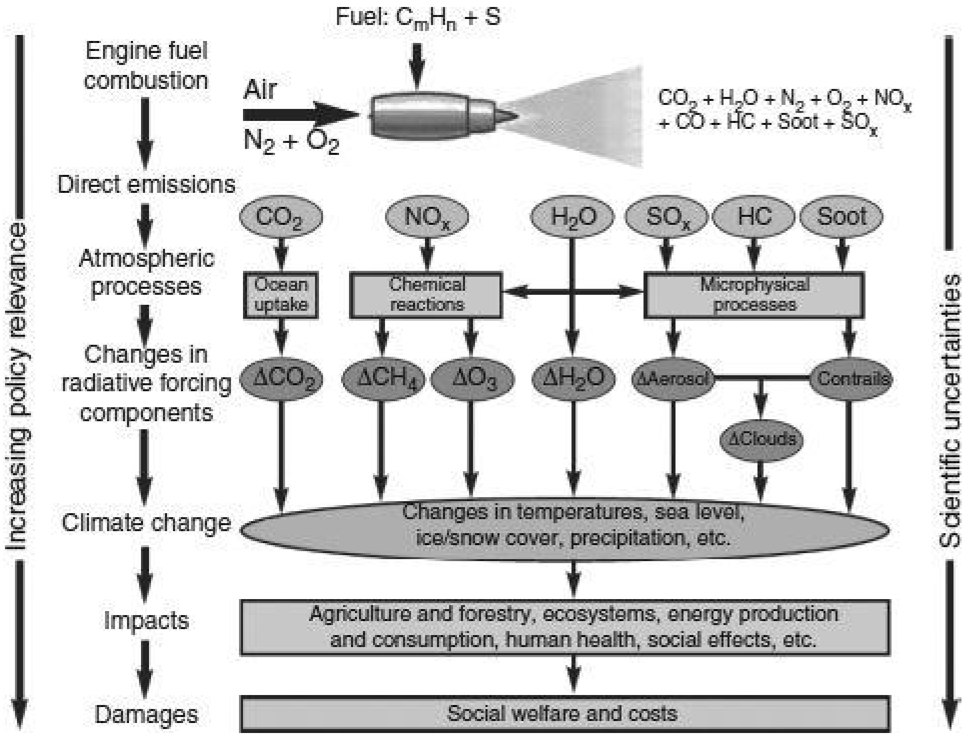
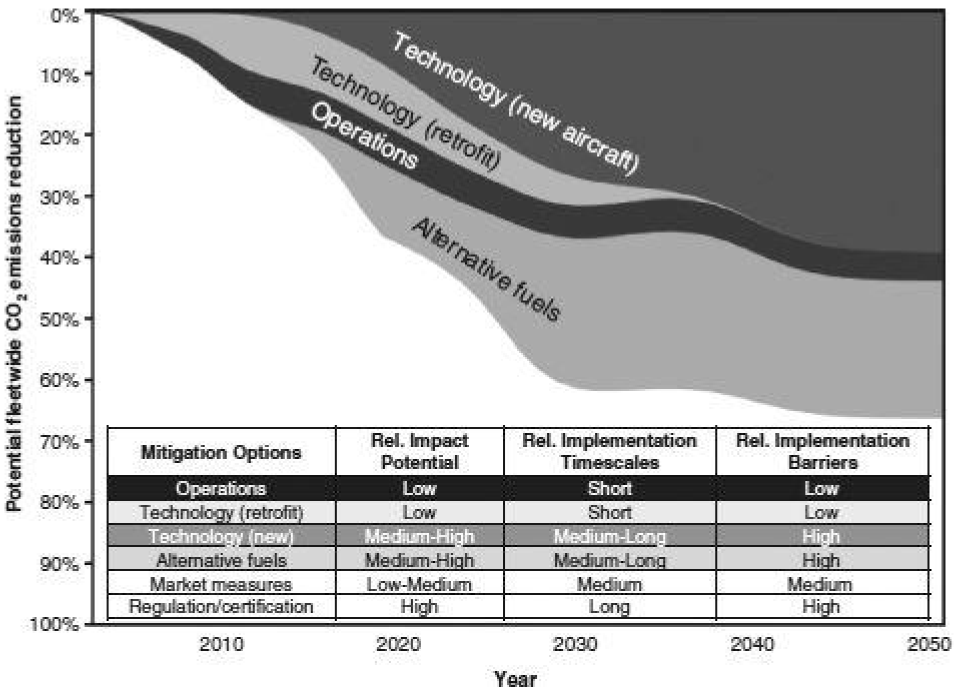


FIGURE 6.16 Aviation emissions impact and damage pathways [*Source: adapted from (ICAO, 2010a*).]

From an airport perspective, climate change potentially has several important impacts. First, it could alter the gross domestic product (GDP) of nations and hence the wealth of the traveling public in different countries. This would impact the evolution of demand for air transportation. Second, long-term changes to climate could change tourism patterns around the world (e.g., snow cover at ski resorts, peak temperatures of beach destinations, etc.). The network appropriate to service this new distribution of demand would thus evolve. Third, sea-level and weather changes (e.g., wind patterns, frequency of adverse weather conditions) would impact airport needs for drainage, snow/ice clearing equipment, runway orientations, etc. (EUROCONTROL, 2008).

Airport-Level Climate Change Mitigations

There is a growing library of guidance to airports, airlines, and other stakeholders to assist in GHG emissions reductions. For example, ACRP (2011a) lists strategies for reductions in a wide range of areas, including airfield design and operations; business planning; construction; carbon sequestration; energy management; ground service equipment; ground transportation; operations and maintenance; performance measurement; and renewable energy. Airlines can mitigate climate-related emissions by reducing fuel burn and by taking other actions, such as contrail avoidance, that involve changes to aircraft flight paths (ACI, 2009; EC, 2005; IATA, 2009). Because fuel accounts for a high proportion of airline operating expenses, there is alignment between the economic and climate impact pressures in this case. Figure 6.17 shows how some of the key opportunities for reducing climate emissions (in this case CO_2) may be implemented over the next several decades. The relative potential impacts on emissions, implementation timescales, and barriers vary significantly between the alternatives. For example, changes to operations have relatively small impact reduction potential, but they can be implemented relatively quickly because their implementation barriers are lower compared to other actions. By contrast, new certification standards take longer to establish (because of the lengthy international negotiations required to reach agreement), but then they may greatly reduce environmental impacts once implemented. Key mitigations are discussed next.



Mitigation Options	Rel. Impact Potential	Rel. Implementation Timescales	Rel. Implementation Barriers
Operations	Low	Short	Low
Technology (retrofit)	Low	Short	Low
Technology (new)	Medium-High	Medium-Long	High
Alternative fuels	Medium-High	Medium-Long	High
Market measures	Low-Medium	Medium	Medium
Regulation/certification	High	Long	High

FIGURE 6.17 Carbon dioxide emissions reduction opportunities. [*Source: adapted from (Kar et al., 2010)*.]

Operational Procedures

The operational mitigations identified with a "C" in Fig. 6.8 are effective at reducing climate impacts of aviation. The proportion of total fuel burn in each phase of flight depends on the type of aircraft and mission being flown (e.g., short-haul vs. long-haul), but a general approximation is that 5 to 10 percent is burnt on the ground, 10 to 30 percent in climb and descent phases (including the terminal areas around the origin and destination airports), and the balance during cruise. There are ongoing efforts to reduce fuel burn in all phases of flight (e.g., the AIRE and ASPIRE initiatives previously discussed), but the cruise portion of flight obviously offers the biggest potential for GHG emissions reduction because it accounts for most of the fuel burn. However, the cruise phase also presents the biggest challenge to implementing changes given its large geographic scope. Therefore, although the ground and airspace regions around airports account for relatively little of the fuel burn and climate-impacting emissions, airport stakeholders can play a relatively large role in their mitigation by promoting improved airport operations, airspace design, and procedures.

Many of the operational techniques for reducing noise and air quality impacts on the ground result in less engine-on time and therefore lower fuel burn and GHG emissions. These include single-engine taxi; extended towing of aircraft; limiting the use of APUs; and preferential runway assignment and airfield design aimed at reducing taxiing distances and time. Airports can also promote other ground, departure, and approach/landing flight phase operational improvements. Surface congestion management approaches are especially effective at reducing taxi fuel burn and associated air quality and GHG emissions. Every airport has a limit to the number of aircraft it can efficiently handle as a function of characteristics such as runway configuration, weather conditions, and demand. During periods of high demand, surface congestion management aims to allow just enough aircraft to taxi out to keep the airport operating at this limit. Excess flights are held at gates or other appropriate location with engines off until they can be released to the departure runway efficiently, as Fig. 6.18 shows. By restricting the number of aircraft moving on the surface, it reduces engines-on taxi-out time, fuel burn, and emissions. Such concepts have been widely studied in operational trials in the United States (including Boston/Logan, New York/Kennedy and Memphis) and significant benefits are observed: for example, see Simaiakis et al. (2011).

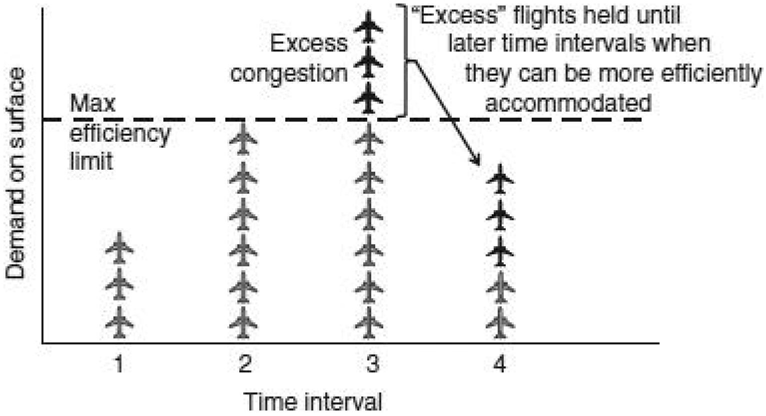


FIGURE 6.18 Surface congestion management concept.

Continuous climb departures (CCDs) minimize level-offs during the climb phase to get aircraft to their more efficient cruise altitudes as soon as possible. They are similar to the CDA procedure previously described for approach. These CDA procedures, primarily designed to reduce noise impacts, also save approach fuel and emissions given that the engines are generally at lower thrust levels for longer during the approach. Complementary to CDAs, delayed deceleration approaches (DDAs) keep the aircraft speed higher for longer during the initial stages of the approach, and hence in a cleaner aerodynamic configuration resulting in lower engine thrust and fuel burn. Various optimized profile descents (OPDs) are being explored which can incorporate CDA vertical and DDA speed profiles all the way from the top of descent down to final approach to achieve maximum fuel and emissions reductions. Increased use of RNAV/RNP capabilities with well-designed departure and arrival procedures that are sensitive to operational and environmental needs are also critical to delivery of fuel reduction benefits into the future.

Technology

Aircraft technology alternatives available to operators include retrofits and all-new equipment. Many older, long-range aircraft types can be modified with new winglets (as available for B757 and B767 aircraft) and fuselage elements (e.g., tail cone retrofit for MD80 aircraft) that can reduce drag and decrease fuel burn and emissions by as much as 5 percent. Although these retrofits are expensive, high fuel costs often reduce the payback period to a few years or less. In the longer term, new aircraft offer greater reduction potential but require larger capital investments on the part of the operators. Geared turbofans and unducted fan engines (see discussion of noise impacts in Sec. 6.2) can be integrated with conventional tube and wing aircraft and hold promise of 15 to 25 percent reductions in fuel burn and emissions compared to existing conventional turbofan engines. New airframe and engine designs, such as blended wing-body configurations, could potentially reduce fuel burn and emissions by up to 70 percent but, as previously described, they may have airport integration issues because of their very different configuration from current types.

Alternative Fuels

One of the big problems with burning aviation fuel derived from conventional fossil fuel is that it emits carbon that had been locked away from the atmosphere for millions of years. Alternative fuels based on renewable sources such as cellulosic biomass have the advantage that the growth of the feedstock absorbs nearly as much carbon dioxide from the atmosphere as the burning of the resulting fuel creates. The net lifecycle carbon dioxide emissions from alternative fuels (the so-called "well-to-wake" emissions of aviation fuel) are thus significantly lower. The aviation industry is actively exploring drop-in alternative fuels (i.e., ones that do not require significant modifications to aircraft/engine technologies or airport fuel distribution infrastructures). These would be aviation fuels chemically almost identical to those derived from conventional fossil fuels. Alternative fuels based on the hydroprocessing of plant oils derived from soybean, palm, Jatropha, and algae show promise. However, many obstacles need to be overcome before alternative aviation fuels are available in enough quantities to deliver large environmental benefits (RAND, 2009).

Climate Emissions Charges

Governments are exploring emissions trading schemes that include aviation to incentivize the reduction of climate emissions. A central, typically governmental, authority implements these market-based arrangements. It sets a limit or cap on the total amount of a pollutant that can be emitted. It then allocates or sells credits within the cap to trading entities such as airlines. These credits are emissions permits that represent the right to emit or discharge a specific volume of a pollutant. Entities that need to increase their emissions permits (e.g., airlines for whom it is expensive to implement mitigations) must buy them from entities that require fewer permits (e.g., industries for which it is cheaper to implement mitigations than the price at which they can sell their permits). Emissions trading schemes are thus designed to reduce emissions in an economically efficient way. Such schemes operate effectively in several parts of the world, for different pollutants and industries.

The European Union Emissions Trading Scheme (EU ETS) is the first that explicitly includes the aviation industry. Guidance on ETS can be found in ICAO (2008c) and ACRP (2011b), and on preparing GHG emission inventories in ACRP (2009a).

Some airports may have the option of buying "offsets" to mitigate their climate impacts, especially if they have set targets for carbon neutral growth. Instead of achieving aviation emissions reductions, offset credits fund projects in other areas to achieve certified GHG reductions. The ACI-Europe Airport Carbon Accreditation (ACA) scheme offers a framework for airport operators to gain formal recognition of their efforts to achieve carbon neutral status.

6.5 Water Quality

Water Quality Impacts

Airport operators need to pay attention to their wide range of fluid discharges. These particularly include those from de-icing of aircraft, handling fuels, and stormwater runoffs.

De-icing Fluids

Proper prevention and removal of ice on aircraft is critical for safe operation in cold climates. Accumulated snow and ice reduces the lift produced by the wings and makes aircraft unstable. This has been the cause of multiple accidents, such as the 1982 Air Florida flight that crashed shortly after takeoff from Washington/Reagan airport. Operators in snowy climates thus devote substantial resources to removing ice and snow from aircraft and preventing it from re-forming before the aircraft has taken off. With climate change effects, the frequency of such conditions in different locations may change and airports need to adapt accordingly, for example in terms of how much investment to make in de-icing and snow removal equipment.

A central element in the de-icing process is the spraying of aircraft with anti-icing/deicing fluids (ADF) that melt existing and inhibit the further formation of snow and ice. Currently, the most effective methods involve the application through high-pressure hoses of heated glycol-based fluids with additives such as corrosion inhibitors, flame retardants, dispersion agents, and thickeners (Fig. 6.19). These chemicals are toxic to mammals and implicated in neurological, cardiovascular, and gastrointestinal health problems (EPA, 2002).

FIGURE 6.19 Aircraft de-icing. (*Source: Munich International Airport*.)

Airports in cold climates use millions of gallons of de-icing fluid that can pollute groundwater and thus pose environmental issues unless properly managed. Airports may also use other chemicals on airfield surfaces to loosen snow and ice before being removed by mechanical means such as snowplows.

Fuel and Other Chemical Leaks and Spills

Airports normally store substantial amounts of aviation and other fuels. The storage and distribution systems do not ordinarily pollute the environment because distributors take care to prevent losses of this expensive resource. However, when leaks and spills do occur, they can severely damage the environment, particularly groundwater and wildlife (discussed in the next section). Other chemicals such as aircraft servicing and maintenance fluids, fire-fighting fluids, pesticides, and herbicides used on the airport property can also have major adverse environmental impacts if leaks and spills occur.

Stormwater Runoff

Because airports consist of large areas of paved surface, rainwater runs off quickly and flushes away pollutants that may have accumulated on these surfaces. Moreover, the large quantities of water can create flash floods that can cause runway or taxiway closures if the airport drainage system is not designed properly.

Airport-Level Water Quality Mitigations

De-icing Fluids

Guidance for the use and management of de-icing fluids can be found in ACRP (2009b, 2011c). No effective substitutes have been found so far for glycol-based ADFs (acetate-based fluids used at some airports are less toxic but ineffective at lower temperatures). Therefore, the main strategies for managing the environmental impacts are reduction of the amounts used, collection and disposal of the fluids, and recycling. In general, airports and airlines use a combination of these approaches (EPA, 2000).

A number of airports, such as Denver/International, Montreal/Trudeau, and Toronto/Pearson, use centralized de-icing facilities. These are located as conveniently as possible to the ends of the departure runways to reduce the time between fluid application and takeoff. They include a network of drains intended to capture and channel the ADF to special retaining areas. In some cases, these centralized de-icing facilities make it possible to recycle the runoff from de-icing and sell it for other purposes.

Where central facilities are not available or oversubscribed, specially equipped trucks de-ice aircraft at their gates or other airport locations (see Fig. 6.19). To prevent ADF from mixing with groundwater, airports may install special drainage systems in ramp areas, use valves and sewer plugs, or have special vehicles vacuum up the fluid. Other alternatives include using settling pools where microorganisms can break down the toxic chemicals into less harmful by-products and sludge that can be removed and disposed of more easily than the original toxic chemicals.

None of these procedures collect all the ADF. Some ADF inevitably drips from the aircraft as they taxi and during takeoff. Recently, some carriers and airports (including New York/Newark and Kennedy) have been developing a system of infrared heaters in hangars to melt ice and snow. These often need to be supplemented with conventional de-icing to prevent further ice from forming. While such heaters can significantly reduce ADF use, their power demands are very high, which lead to other environmental considerations given this power still needs to be generated by power plants.

Fuel and Other Chemical Leaks and Spills

Airport operators mitigate leaks and spills through reliable storage and distribution, secondary containment, and effective cleanup procedures. They need to maintain fuel storage and distribution systems carefully to prevent leakage and contamination of the groundwater. Some airports use zoned leak detection systems for underground fuel lines to identify volumetric changes of product and thus leaks. They can also use a distributed fuel system that delivers fuel directly to gate areas via pipes, reducing the need for fuel trucks and dispensing from centralized fuel farms.

Secondary containment protects other parts of the airport against accidental spills or deliberate sabotage. Berms are often built around tanks to contain possible massive spills. Strict management protocols also need to be in place and enforced for the other chemicals used around the airport to ensure that leaks and spills are prevented as much as possible. When fuel or other chemical leaks do occur, well-defined cleanup procedures are essential.

Stormwater Runoff

As for any urban area that has large areas of paved and hard surfaces, airports need properly sized and located drainage systems to mitigate stormwater runoff issues. These involve drainage inlets around the airport, coupled to drainage lines and culverts connected to appropriate outflow sites and/or settling ponds that will retain water to prevent flooding and to settle material and chemicals swept away from the airfield. In the United States, airports have to secure discharge permits under the National Pollution Discharge Elimination System (NPDES: see EPA, 2001) if their discharges go directly to surface waters. NPDES permits require monthly grab samples (single samples collected at a particular time and place) from the appropriate discharge points. These must be tested against allowable EPA limits for pH, oil and grease, total suspended solids (TSS), benzene, surfactants, and bacteria levels during both wet and dry weather—and for de-icing compounds during the winter months.

6.6 Wildlife

Wildlife Impacts

Airports are home to many kinds of wildlife, particularly flocks of birds and other small animals, although large mammals may be an issue in some locations. An airport may also have impacts on nearby coastal resources, rivers, wetlands, floodplains, and farmland. Airports need to understand both the impacts that wildlife could have on their safe operation and development, and the potential impacts of the airport on the wildlife around them (especially if there are any endangered species).

One of the biggest operational concerns is collision with wildlife on or near the airport. Bird strikes are major safety hazards that can cause significant damage to aircraft and engines. They have been responsible for several high-profile incidents, including the 2009 forced ditching in the Hudson River of an Airbus A320 that struck a flock of Canada Geese during initial climb-out from New York/LaGuardia. Less dramatic bird and other animal strikes are common occurrences at major airports (especially those in coastal locations, near wetlands, on bird migration routes, or other rural habitats), causing disruptions to operations.

The development and expansion of airports can be significantly restricted if they affect the habitat of endangered wildlife species. Hazards posed to wildlife include physical habitat disruption (e.g., during airport construction and operations), exposure to toxic chemicals used during airport activities, and collisions with aircraft and airport vehicles. The environmental impact processes require major airport projects to document their impacts to wildlife and plans to mitigate any hazards. For example, planning for the development of Denver/International airport involved careful assessments of its impacts on local bald eagle populations and led to mandates for several conservation actions (Dempsey et al., 1997). In a number of cases, concerns about wildlife have severely delayed or altered airport developments.

Airport-Level Wildlife Mitigations

To reduce wildlife dangers to aircraft and vice versa, many airports have programs to reduce the number of birds and other animals near runways. Strategically, these include the elimination or reduction of attractants such as food sources, open water, and cover. Vegetation management includes the types of vegetation that are planted or removed from the airport to achieve these goals, as well as activities such as grass mowing procedures at different times of the year to discourage bird settlement and nesting. It is also important to control the establishment of landfills, open dumps, or waste disposal sites on or in the vicinity of airports to reduce bird numbers (FAA, 2007a). Appropriate fencing can prevent larger ground-dwelling animals from entering the airport property.

Tactically, airport operators have various temporary ways to scare away birds and other animals that settle on the airport. These include use of loud sounds (such as from gas cannons and firecrackers, but these can cause noise annoyance to humans!), animal distress calls played through a loudspeaker located on a ground vehicle, and use of birds of prey or dogs by trained handlers. Research has also been conducted into the efficacy of chemical repellants to repel birds from standing pools of water on airports. Finally, trapping/removing or shooting a small number of animals that cause persistent hazards might be considered appropriate under certain conditions if all other methods fail. Federal, state, and municipal permits may be required before such actions are taken. For further details on wildlife impact mitigations, see Cleary and Dolbeer (2005).

6.7 Environmental Legislation and Review Processes

Environmental Legislation

Various national laws protect the environment. In the United States, the EPA is responsible for the establishment and enforcement of U.S. protection standards across all areas and industries, but it sometimes delegates these responsibilities to other agencies, such as the

FAA in the case of aviation. As of 2012, the laws relevant to environmental impacts of aviation include the National Environmental Policy Act (NEPA), Aircraft Noise Abatement Act, Clean Air Act, Clean Water Act, Endangered Species Act, National Historic Preservation Act, and Wild and Scenic Rivers Act. Similar legislation exists in other parts of the world, such as Europe's Environment Action Program.

Environmental Review Processes

Environmental review processes are designed to assist policy makers in evaluating the need for, potential environmental impacts of, and suggested mitigations of proposed actions that may significantly affect the environment. This includes many airport development activities such as changes to the airport infrastructure, airspace, and operational procedures. An overview of the U.S. NEPA process is used as an example of the steps typically required.

The NEPA process is triggered when an airport sponsor applies for federal funding for a specific project. Airports have several resources to guide their compliance with NEPA requirements. These include the following:

- FAA Order 1050.1E "Environmental Impacts: Policies and Procedures" that provides instructions for implementing NEPA to all FAA lines of business (FAA, 2006a)
- FAA Order 5050.4B"National Environmental Policy Act (NEPA) Implementing Instructions for Airport Actions" that focuses on airport action for compliance with NEPA and Special Purpose Laws (FAA, 2006b)
- "Environmental Desk Reference for Airport Actions" that integrates information on NEPA and Special Purpose Laws applicable to airport actions (FAA, 2007b)

FAA Order 1050.1E lists 18 environmental impact categories that airports need to consider, including: air quality; coastal resources; compatible land use; construction impacts; Department of Transportation Act "Section 4(f)"; farmlands; fish, wildlife, and plants; floodplains; hazardous material, pollution prevention and solid waste; historical, architectural, archeological, and cultural resources; light emissions and visual impacts; natural resources, energy supply and sustainable design; noise; secondary (induced) impacts; socioeconomic impacts, environmental justice and children's environmental health and safety risks; water quality; wetlands; and wild and scenic rivers. A "programmatic environmental assessment" against these potential impact areas is often conducted at the system or master-planning phase of the project to provide an initial, broad-brush assessment. This concept is called strategic environmental assessment (SEA) in Europe, Australia, Canada, and some other countries. A scoping session is also conducted to learn about the concerns of citizen groups and federal, state, and local agencies that have oversight responsibility on specific environmental values. For this purpose, citizens and agency review committees are formed. These are used to identify critical environmental values and issues relevant to the proposed action and to agree upon a program for addressing them. To address concerns in a timely manner, a systematic schedule for review committees is established for critical junctures in the study. The formal NEPA process may require three general levels of analysis: categorical exclusion, environmental assessment, and environmental impact statement. These are discussed in turn.

Categorical Exclusion

A proposed activity may be categorically excluded from a detailed environmental analysis if it meets certain criteria that a federal agency has previously determined as having no significant environmental impact. A number of agencies have developed lists of actions that are normally categorically excluded from further NEPA evaluation. For the FAA, the list may include the following:

- Building or extending aircraft operating area fencing or jet blast facilities
- Building, repairing, or extending an existing airport's aprons, loading ramps, taxiways, or taxi lanes provided they have only on-airport impacts
- Building, maintaining, moving, or repairing roads, if the action does not permanently reduce the level of service to unacceptable levels
- Extending, filleting, grooving, marking, rebuilding, resurfacing, or strengthening existing runways or runway surface areas
- Installing or upgrading airfield lighting; conducting landscape maintenance and vegetative and erosion control measures as long as they do not spread invasive species or attract wildlife hazardous to aviation
- Installing or upgrading nonradar equipment
- Installing vegetation, berms, or sound walls to reduce noise, provided they do not attract wildlife hazardous to aviation

In all cases, no "extraordinary circumstances" can exist for a categorical exclusion to be permitted and the onus is on the applicant to prove this is the case.

Environmental Assessment

At the second level of analysis, a written environmental assessment (EA) is prepared to determine whether or not the undertaking would significantly affect the environment. This would be required for any item that does not meet the criteria for a categorical exclusion, including the following:

- Changes involving new facilities or operations
- Land acquisition
- New runways
- Major runway strengthening or extension
- Impacts to prime and unique farmland or waters and wetlands

The format and contents of the EA documentation are prescribed in detail (FAA, 2006b). This includes the following key sections: purpose of and need for the proposed action; detail on the proposed action; alternatives to the proposed action; parts of the environment affected; description of the impacts to those parts of the environment; mitigations to the impacts identified; a cumulative impacts analysis; and a list of the agencies and people consulted in the analysis.

Drafts of the EA will typically be distributed to interested parties (including the public) for comment ahead of a final EA document being delivered to the FAA responsible official, who then determines whether to issue a finding of no significant impact (FONSI) for the proposed action (potentially as long as certain mitigations are enacted), or requiring an environmental impact statement to be prepared.

Environmental Impact Statement

An environmental impact statement (EIS) must be prepared if the EA results in a determination that the environmental consequences of a proposed federal undertaking may be significant. A federal agency may choose to prepare an EIS without first preparing an EA if it anticipates that an undertaking may significantly impact the environment, or if a project is environmentally controversial. An EIS is a more detailed evaluation of the proposed action and alternatives than the EA. Sections include an executive summary; table of contents; description of purpose of and need for the proposed action; the alternatives considered, including the “no-action” alternative; parts of the environment affected; description of the environmental impacts; a cumulative impact summary; mitigation alternatives for the impacts identified; list of preparers; list of recipients; index; comments; response to comments; appendices; and list of unavailable or incomplete information relevant to the application.

The public, other federal agencies, and outside parties may provide input into the preparation of an EIS, scoping the impacts that should be assessed and then commenting on the draft EIS when it is completed. After reviewing the final EIS, the federal agency will prepare a public “record of decision” that will approve or block the proposed activity. This record of decision includes a brief description of the airport sponsor’s proposed action and summaries of the necessary federal actions that must be completed before the airport sponsor may begin the proposed action; the alternatives considered; information needed to address resources protected under special purpose laws or airport legislation; mitigation measures in an approved EIS; and a description of any changes to the mitigation in an approved EIS.

Note that the public and representatives of responsible agencies have an important role in the NEPA process. They particularly provide input on what issues should be addressed, actively participate in public meetings, and provide comments on the draft and final versions of NEPA documents.

6.8 Summary

This chapter has discussed the range of environmental impacts that airport planners, designers, operators, and managers need to understand given the growing importance and complexity of environmental issues. Noise has historically been the dominant environmental concern for airports, but other factors such as air quality, climate change, water quality, and wildlife impacts now also must be carefully considered. Numerous mitigation alternatives are available to airports and those in the key areas of operational procedures, technologies, and market-based measures have been discussed. Environmental review processes designed to assist policy makers in evaluating the need for, impacts of, and mitigations of proposed airport actions have also been outlined. Finally, it is seen that tradeoffs between environmental impact areas and between environment and other operational performance metrics present significant challenges, but airports can manage them by having a solid understanding of the issues and effective mitigation strategies to address them.

Exercises

6.1. Consider a major commercial airport in your region. What is its history of public concerns with its noise? Are other environmental impacts now becoming more important? How has the airport operator responded politically? What have the airport and airlines done to mitigate effects?

6.2. Find, on the web, some EISs or noise studies for airport projects. Examine their noise contours. What levels of noise do they reflect? How many people are affected? How does the airport intend to mitigate these effects? Do the noise contours suggest that flight paths could or should deviate from extensions of the runway centerlines to avoid populated areas? Reflect on and discuss your findings.

6.3. Obtain a recent version of the INM noise model (available through the FAA) and exercise it. Evaluate it as a user. Do you feel that it provides you with the kind of information you might require as an airport planner? As a resident of the community? As a local political leader?

6.4. Explore air quality issues for an airport of interest. As a benchmark, first identify the kinds of air quality controls on automobiles or factories that prevail in its region. Then identify the kinds of controls or mitigations affecting airport sources. Do you think that these are compatible? If not, what might be more reasonable?

6.5. Examine the water quality and wildlife issues for some airport of interest. What are the major environmental concerns, if any? How do the airport operator and the local communities deal with them?

6.6. Find a major airport with an environmental performance plan. Critically assess it against the issues and mitigations discussed in this chapter.

6.7. Discuss some of the tradeoffs that an airport of choice has to deal with, both between environmental impact areas (e.g., minimization of noise vs. minimization of air quality or climate impacts) and between environment and other performance metrics, such as throughput. Do these tradeoffs differ at different airports?

6.8. Discuss how public and political perceptions of environmental impacts vary between stakeholders and between countries. Are some impacts considered more important in some countries than others? Why?

References

ACI, Airports Council International (2009) *Guidance Manual: Airport Greenhouse Gas Emissions Management*, Montreal, Canada, www.aci.aero.

ACRP, Airport Cooperative Research Program (2008) *Summarizing and Interpreting Aircraft Gaseous and Particulate Emissions Data*, ACRP Report 9, Transportation Research Board, Washington, DC, www.trb.org.

ACRP, Airport Cooperative Research Program (2009a) *Guidebook on Preparing Airport Greenhouse Gas Emissions Inventories*, ACRP Report 11, Transportation Research Board, Washington, DC, www.trb.org.

ACRP, Airport Cooperative Research Program (2009b) *De-icing Planning Guidelines and Practices for Stormwater Management Systems*, ACRP Report 14, Transportation Research Board, Washington, DC, www.trb.org.

ACRP, Airport Cooperative Research Program (2011a) *Handbook for Considering Practical Greenhouse Gas Emission Reduction Strategies for Airports*, ACRP Report 56, Transportation Research Board, Washington, DC, www.trb.org.

ACRP, Airport Cooperative Research Program (2011b) *The Carbon Market: A Primer for Airports*, ACRP Report 57, Transportation Research Board, Washington, DC, www.trb.org.

ACRP, Airport Cooperative Research Program (2011c) *Optimizing the Use of Aircraft Deicing and Anti-icing Fluids*, ACRP Report 45, Transportation Research Board, Washington, DC, www.trb.org.

ACRP, Airport Cooperative Research Program (2012) *Handbook for Evaluating Emissions and Costs of APUs and Alternative Systems*, ACRP Report 64, Transportation Research Board, Washington, DC, www.trb.org.

Barrett, S. R. H., Britter, R. E., and Waitz, I. A. (2010) "Global Mortality Attributable to Aircraft Cruise Emissions," Environmental Science and Technology, 44(19), pp. 7736–7742, DOI: 10.1021/es101325r.

Cleary, E. and Dolbeer, R. (2005) *Wildlife Hazard Management at Airports: A Manual for Airport Personnel*, 2d ed., FAA, Washington, DC, www.faa.gov.

Dempsey, P., Goetz, A., and Szyliowicz, J. (1997) *Denver International Airport: Lessons Learned*, McGraw-Hill, New York.

EC, European Commission (2005) *Reducing the Climate Change Impact of Aviation*, COM 459, Brussels, Belgium, ec.europa.eu.

EC, European Commission (2011) *Flightpath 2050: Europe's Vision for Aviation*, Brussels, Belgium, EUR 098 EN, doi: 10.2777/50266, ec.europa.eu.

EPA, Environmental Protection Agency (2000) *Preliminary Data Summary—Airport Deicing Operations (Revised)*, Office of Water, Report EPA-821-R-00-016, Washington, DC, www.epa.gov.

EPA, Environmental Protection Agency (2001) *National Pollution Discharge Elimination System Permit Program*, Office of Water, Washington, DC, www.epa.gov.

EPA, Environmental Protection Agency (2002) *Source Water Protection Practices Bulletin: Managing Aircraft and Airfield De-icing Operations to Prevent Contamination to Drinking Water*, EPA 816-F-02-018, Washington, DC, www.epa.gov.

EPA, Environmental Protection Agency (2011) *The Green Book Nonattainment Areas for Criteria Pollutants*, Washington, DC, www.epa.gov.

EUROCONTROL (2008) *Challenges of Growth 2008: Summary Report*, Brussels, Belgium, www.eurocontrol.int.

FAA, Federal Aviation Administration (2006a) *Environmental Impacts: Policies and Procedures*, FAA Order 1050.1E, Washington, DC, www.faa.gov.

FAA, Federal Aviation Administration (2006b) *National Environmental Policy Act (NEPA) Implementing Instructions for Airport Actions*, FAA Order 5050.4B, www.faa.gov.

FAA, Federal Aviation Administration (2007a) *Hazardous Wildlife Attractants on or Near Airports*, Advisory Circular 150/5200-33, Washington, DC, www.faa.gov.

FAA, Federal Aviation Administration (2007b) *Environmental Desk Reference for Airport Actions*, Washington, DC, www.faa.gov.

Fidell, S. and Silvati, L. (2004) "Parsimonious Alternative to Regression Analysis for Characterizing Prevalence Rates of Aircraft Noise Annoyance," *Noise Control Engineering Journal, 52*(2), pp. 56–68.

GAO, General Accounting Office (2000) *Aviation and the Environment—Results from a Survey of the Nation's 50 Busiest Commercial Service Airports*, GAO/RCED-00-222, Washington, DC, www.gao.gov.

IATA, International Air Transport Association (2009) *A Global Approach to Reducing Aviation Emissions*, Montreal, Canada, www.iata.org.

IATA, International Air Transport Association (2011) *IATA Airport, ATC and Fuel Charges Monitor*, Montreal, Canada, www.iata.org.

ICAO, International Civil Aviation Organization (2005) *Procedures for Air Navigation Services—Aircraft Operations (PANS OPS) Volume I—Flight Procedures*, Doc 8168, Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2006) *Airport Economics Manual*, Doc 9562, 2d ed., Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2007a) *ICAO Environmental Report 2007*, Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2007b) *Airport Air Quality Guidance Manual*, Doc 9889, Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2007c) *Guidance on Aircraft Emissions Charges Related to Local Air Quality*, Doc 9884, Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2008a) *Environmental Protection/Annex 16 to the Convention on International Civil Aviation: Volume 1 Aircraft Noise*, 5th ed., Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2008b) *Environmental Protection/Annex 16 to the Convention on International Civil Aviation: Volume 2 Aircraft Engine Emissions*, 3d ed., Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2008c) *Guidance on the Use of Emissions Trading for Aviation*, Doc 9885, Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2010a) *ICAO Environmental Report 2010*, Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2010b) *Guidance on the Balanced Approach to Aircraft Noise Management*, Doc 9829, Amendment 1, Montreal, Canada, www.icao.int.

ICAO, International Civil Aviation Organization (2010c) *Policies on Charges for Airports and Air Navigation Services*, Doc 9082/6, Montreal, Canada, www.icao.int.

IPCC, Intergovernmental Panel on Climate Change (1999) *Aviation and the Global Atmosphere*, J. E. Penner, D. H. Lister, D. J. Griggs, D. J. Dokken, and M. McFarland (eds.), Cambridge University Press, Cambridge, U.K., www.ipcc.ch.

IPCC, Intergovernmental Panel on Climate Change (2007) *IPCC Fourth Assessment Report: Climate Change 2007 Synthesis Report*, R. K. Pachauri and A. Reisinger (eds.), IPCC, Geneva, Switzerland, www.ipcc.ch.

Kar, R., Bonnefoy, P., and Hansman, R. (2010) *Dynamics of Implementation of Mitigating Measures to Reduce CO_2 Emissions from Commercial Aviation*, Master's thesis, Massachusetts Institute of Technology, Department of Aeronautics & Astronautics, Cambridge, MA, http://hdl.handle.net/1721.1/56268.

Mahashabde, A., Wolfe, P., Ashok, A., Dorbian, C., He, Q., Fan, A., Lukachko, S., et al. (2011) "Assessing the Environmental Impacts of Aircraft Noise and Emissions," *Progress in Aerospace Sciences, 47*(1), pp. 15–52, doi:10.1016/j.paerosci.2010.04.003.

Nelson, J. P. (2004) "Meta-analysis of Airport Noise and Hedonic Property Values: Problems and Prospects," *Journal of Transport Economics and Policy, 38*(1), pp. 1–28.

NRC, National Research Council (2002) *For Greener Skies: Reducing Environmental Impacts of Aviation*, Committee on Aeronautics Research and Technology for Environmental Compatibility, Aeronautics and Space Engineering Board, www.nap.edu.

NSTC, National Science and Technology Council (2010) *National Aeronautics Research and Development Plan*, Washington, DC, http://www.whitehouse.gov.

RAND Corporation (2009) *Near-Term Feasibility of Alternative Jet Fuels*, Santa Monica, CA, www.rand.org.

Schäfer, A., Heywood, J., Jacoby, H., and Waitz, I. (2009) *Transportation in a Climate-constrained World*, MIT Press, Cambridge, MA.

Schultz, T. (1978) "Synthesis of Social Surveys on Noise Annoyance," *Journal of the Acoustical Society of America, 64*(2), pp. 377–405.

Simaiakis, I., Balakrishnan, H., Khadilkar, H., Reynolds, T. G., Hansman, R. J., Reilly, B. and Urlass, S. (2011) "Demonstration of Reduced Airport Congestion Through Pushback Rate Control", *9th USA/Europe ATM Research & Development Seminar (ATM2011)*, Berlin, Germany.

Smith, M. (1989) *Aircraft Noise*, Cambridge University Press, Cambridge, U.K.

WHO, World Health Organization (1999) *Guidelines for Community Noise*, B. Berglund, T. Lindvall, and D. Schwela (eds.), Geneva, Switzerland, www.who.int.

[1]INM is planned to be merged in the FAA's Aviation Environmental Design Tool (AEDT) and enhanced with capabilities from the Model for Assessing Global Exposure to the Noise of Transport Aircraft (MAGENTA).

[2]In 2010, the U.S. EPA-recommended VSL was \$5.8 m ± 2.6 m.

[3]The IPCC was awarded the 2007 Nobel Peace Prize "for their efforts to build up and disseminate greater knowledge about man-made climate change, and to lay the foundations for the measures that are needed to counteract such change."

CHAPTER 7

Organization and Financing

The institutional, organizational, and financial characteristics of airports are changing rapidly, stimulated in large part by airport privatization and airline deregulation. Many different "models" of airport ownership and management exist. The traditional model (outside the United States) that places airport management in the hands of a central bureaucracy in the national government does not usually meet the needs of large airports in a fast-changing industry. Most new models center on the concept of the airport authority, a corporate entity owned by government or private investors or a combination of the two, which acts as an autonomous and flexible airport operator.

Airports must contend with legal, financial, planning, public communications, administrative, human resource, environmental, engineering/technical, commercial, and operational issues. Their evolving organizational structures are designed to carry out this common set of functions and activities. Structures become increasingly pyramidal as airports grow. They can be complex when the airport operator is responsible for a multi-airport system or when the operator engages in extensive activities outside the core airport business.

Concerns regarding the potential abuse of the quasimonopolistic position that airports enjoy in serving origin and destination traffic have led to widespread regulation of airport user charges. The focus has been on regulating aeronautical charges through target rates of return, price caps, and restrictions on the annual rate of increase in unit charges. The treatment of nonaeronautical revenues plays a major role in determining the size of aeronautical user charges.

Airport capital investments can be financed in many different ways, ranging from grants from national governments to revenue bonds issued and serviced by airport operators. The available alternatives depend on the size of the airport and on national laws, economic conditions, and practices. The ability of airport operators to obtain favorable terms for the financing of major projects depends in large part on assessments performed by credit rating agencies. These agencies have developed a set of common criteria that they use for this purpose.

7.1 Introduction

This chapter presents an overview of the institutional, organizational, and financial characteristics of airports around the world. Familiarity with these characteristics is critical to understanding many aspects of the diversity of airports as organizational and economic entities. This material also complements or provides useful background for other chapters in this book, notably the ones on user charges (Chap. 8) and demand management (Chap. 12). It is worth noting at the outset that changes in the organizational and economic aspects of airports have been dramatic since the late 1980s. This is an area where the "landscape" is evolving rapidly.

The contents of the chapter are as follows. Section 7.2 presents a brief survey of the arrangements that exist in various parts of the world regarding the ownership and management of airports. These arrangements increasingly emphasize autonomous management, typically in the form of an airport authority or similar independent corporate entity, as well as participation by private investors in the ownership of airports. These developments, along with the growing complexity of airport operations, underlie the parallel trend toward organizational structures that deviate significantly from traditional forms, as described in Sec. 7.3. As airports move toward operating in many ways like private sector entities, their revenues and balance sheets are coming under increasingly close scrutiny. Airports are in most cases natural monopolies when it comes to originating or terminating passengers. There is justifiable concern that, in the absence of appropriate economic regulation, pricing practices will take advantage of this monopolistic position. Airport economic and financial practices are therefore being subjected to regulatory and/or legal constraints in a number of countries. Section 7.4 describes briefly some types of constraints, along with relevant examples. Finally, Sec. 7.5 identifies the alternative ways in which airport capital investments can be financed and notes, once again, the considerable diversity of international practices in this respect.

7.2 Ownership and Management of Airports

Several alternative "models" of airport ownership and management are in use around the world (Doganis, 1992). Some of these have proved more effective than others. Before reviewing them, it is necessary to establish a context for the discussion.

The term *airport operator* refers to the organizational or corporate entity that has management control of an airport, that is, the right and responsibility for developing and maintaining the airport's infrastructure, as needed, and for operating the airport on a daily basis. Any one of the following four types of entities may act as an airport operator:

1. A branch of a national government
2. A branch of a local or state/regional government

3. A company specializing in airport management, acting as a contractor to a national, state, or local government
4. A corporate entity, often known as an *airport authority*, created especially for the purpose of acting as an airport operator

Whereas the first two types were dominant in the past, the last two, especially the fourth, are becoming increasingly common worldwide when it comes to the busiest airports. This fourth type of airport operator is also the one primarily associated with the global movement toward partial or full airport privatization. Note, however, that an airport authority may be fully government-owned, as has been a common practice in the United States.

The term "privatization" is not a particularly accurate description of the changes that are taking place internationally in airport ownership and management. The privatization of major commercial airports typically involves the transfer to a corporate entity for a long period of (usually) 20 to 50 years of (1) responsibility and management control for developing, maintaining, and operating the airport and (2) rights to residual income from the airport property, that is, to any profits that may be generated there. Thus, the national, regional, or local government that grants the transfer remains, *in principle*, the true owner, because the rights to the airport property will revert back to it once the airport operator's contract expires. (In fact, this true owner may retain some critical prerogatives and regulatory controls.) For most practical purposes, however, it is the airport operator that acts as the property's owner and everyday decision maker during the contract period. The airport operator also may subcontract any set of responsibilities for the entire airport or for parts of the airport to other organizations. For these reasons, the *shareholders of the airport operator* will be treated here as being the effective *owners* of the airport. Note that these shareholders can, in general, be government or private entities or both. Airport ownership, in this sense, can then rest with the following:

- National government
- Regional or state governments
- Local government
- Government- or privately owned corporate entity
- Individual private investors
- Any combination of the above

When a partnership of government and private interests own an airport operator, the strategies pursued by the operator, especially in the long term, often depend on who holds the majority stake. Another aspect with major implications for the governance of the airport is the type of private ownership involved, that is, whether private shareholding is limited to a small number of partners (corporate or otherwise) or extended to the general public through a tender of publicly traded shares ("free float").

Of the large number of owner/operator combinations that can be identified from the two lists, most existing arrangements are consistent with one of six models, A through F, described briefly as follows.

A. *Owned by a combination of national, regional, and/or local governments and operated by a branch of the national government*: Until the 1990s, model A was by far the most common around the world and is still widely used, especially at secondary airports outside the United States and in developing countries. Typically, a branch of the national civil aviation authority (CAA) or of the ministry of transportation is responsible for the management and operation of all or most airports in a country, and it appoints civil servants to carry out these functions at each airport. A few countries even have a Ministry of Airports—or other cabinet level agency—dedicated to the administration and management of their national airport system.

B. *Owned by a combination of national, regional, and/or local government and managed and operated by a branch of a local or regional government*: Model B is met in countries with a strong tradition of decentralized administration and regional autonomy. Many of the most important airports in the United States, including some the busiest airports in the world, are operated by departments in city or county governments. Chicago/O'Hare, Denver/International, and Miami/International belong to this category. For instance, the City and County of Denver Department of Aviation operates Denver/International, and the Los Angeles Department of Airports [formally, Los Angeles World Airports (LAWA)] operates Los Angeles/International, along with the nearby airports of Ontario, Van Nuys, and Palmdale. Similarly, the State of Hawaii Department of Transportation operates 16 airports, including the Honolulu International Airport. Interestingly, these U.S. airports are also among the most "privatized" in the world, in that they tend to contract out most of the activities and functions of an airport operator. Ernico et al. (2012) thoroughly cover this topic and explain why there are currently no privatized large commercial airports in the United States.

C. *Owned by a combination of national, regional, and/or local government and, possibly, of private interests and operated under a management contract by a publicly or privately owned company*: The objective here is to obtain expert airport management, responsive to local conditions and cognizant of best practices elsewhere. Several national, local, and/or regional governments have contracted with well-known airport operators to implement arrangements consistent with model C. Examples include the contracts that the British BAA had in the 1990s in the United States with the cities of Indianapolis and Harrisburg to manage their airports, and the management agreements that Fraport (the operator of Frankfurt/International) has at the airports of Cairo, Riyadh, and Dakar.

D. *Owned by a combination of national, regional, and/or local government and managed and operated as an autonomous airport authority*: Model D applies to some of the busiest airports in the world. It is also the model that initially provided extensive experience with the operation of autonomous airport authorities, proving their advantages and effectiveness (Example 7.1). The best-known example is the Port Authority of New York and New Jersey (PANYNJ), which operates, among other facilities, the three main airports in the New York metropolitan area and, as its name suggests, is jointly owned by the states of New York and New Jersey. Similarly, the Commonwealth of Massachusetts owns Massport, the airport authority that operates Boston/Logan, and the City and County of San Francisco own the authority that operates San Francisco/International.

Example 7.1 Because of the importance of model D and its successful application at several locations in the United States, this example outlines a typical set of the basic terms for these autonomous airport authorities, which are fully owned by state and/or local government. The provisions are based on those in effect for Massport.

- The authority manages and operates the airport (or group of airports) and all related property, facilities, and services. It may also have parallel responsibility for other transportation facilities and services. In the case of Massport, in addition to Boston/Logan, Worcester Airport, and Hanscom Field, a reliever airport near Boston, the authority manages and operates the seaport of Boston and several other secondary facilities.
- A board of directors appointed by the shareholders, that is, the state or local government, acts as the authority's governing body. For Massport, the board consists of seven members, appointed to revolving 7-year terms by the governor of Massachusetts. The governor normally appoints one board member every year.[1] Board members receive only nominal compensation.
- The authority operates on a not-for-profit basis and is exempt from taxes. If conditions permit, it makes "voluntary contributions" to its shareholders, in lieu of taxes. In the case of Massport, the airport has been generating an operating surplus for many years. It allocates part of this surplus to financial assistance to a number of Boston area cities and towns.
- The authority may acquire land and property at fair value, as necessary for the development and operation of its facilities, and may undertake construction projects and other enhancements to its property and facilities.
- The authority may issue tax-exempt bonds against future airport revenues[2] to finance necessary capital investments.

Provisions in the same spirit as those described in Example 7.1 govern the operation of autonomous airport authorities in other regions of the world. There are, however, important differences from country to country regarding the corporate and legal characteristics of the authority, its responsibilities, and its empowerment.[3] The makeup of the ownership of airport authorities is often more varied than in the United States. Some airport authorities outside the United States are also far less local in character. For example, Swedavia is a state-owned company that operates and develops 14 airports in Sweden, including Stockholm/Arlanda and Stockholm/Bromma. AENA and ANA, the counterparts of Swedavia in Spain and Portugal, respectively, are responsible for 46 (AENA) and 8 (ANA) airports, including the very busy airports of Madrid, Barcelona/El Prat, and Lisbon.[4] The Dublin Airport Authority (formerly Aer Rianta) operates the airports of Cork and of Shannon, in addition to Dublin, as well as airport shopping centers and other commercial activities worldwide.

E. *Majority owned by a combination of national, regional, and/or local governments with private minority shareholders, with some shares possibly traded publicly; managed and operated as an autonomous airport authority*: Because of the presence of private interests in their ownership, model E (and model F) airports require thoughtful contractual and regulatory arrangements to balance and protect the interests of the general public and those of the private investors. The Athens International Airport (AIA) Authority, operator of Athens/Venizelos, provides an example. The Greek government owns 55 percent of the shares and a group of private companies, led by Hochtief, best known as a large construction company, owns the other 45 percent.[5] To create the AIA, the Greek Parliament had to enact a lengthy special law laying out the terms and conditions of the AIA contract. In an example in which the public-private partnership was formed differently, the national government originally owned 100 percent of the shares of the Airports Company of South Africa (ACSA). Subsequently, Rome's Airport Authority (Aeroporti di Roma) obtained 20 percent following an international tender.

The second and third busiest airports in Europe, Paris/CDG and Frankfurt/International are other examples of model E. They differ from the previous ones in that some of the privately held minority shares of the airport operators (ADP and Fraport, respectively) are traded in stock exchanges ("free float").

F. *Owned fully or in the majority by private investors, with some shares possibly traded publicly; operated as an autonomous airport authority*: The former British Airports Authority (BAA) is by far the best-known example of model F. It was the first airport authority to be privatized in 1987, with 100 percent of its shares traded publicly in the London Stock Exchange.[7] The BAA originally owned and operated the three main London airports and four others in the United Kingdom. In 2006, a consortium led by Grupo Ferrovial, a Spanish construction company, bought all the shares for the then-enormous sum of £10.1 billion (about $20 billion at the time). The BAA then continued as an example of model F, as it remained 100 percent privately owned, but its shares no longer traded publicly (see also Sec. 7.4). Brussels, Copenhagen, Rome, Sydney, and numerous others are additional examples of airports owned in the majority by private investors but not traded in stock exchanges. Auckland (New Zealand), Vienna, and Zurich are also model F airports but with some publicly traded shares.

Table 7.1 provides information on the ownership and operators of a sample of European airports, including some of the busiest ones. The second column classifies these airports along the lines of the models A through F. Note that models D, E, or F apply in all cases but one. This reflects the dominance as of 2012 of the "autonomous airport authority" concept for managing and operating large commercial airports. Yet the example of the Greek airports (other than Athens) indicates that in 2012 many national government agencies still managed and operated regional and secondary airports. Moscow offers another interesting example: Moscow/Domodedovo is 100 percent privately owned, while Moscow/Sheremetyevo is 100 percent government-owned. Note also that a number of model E and F airports have publicly traded shares (indicated as "free float"). Overall, more than 150 airports worldwide, including many of the busiest ones outside the United States, are currently privatized to some extent, that is, examples of models E and F. No examples of model B (very common and important in the United States) or model C (quite common in developing countries) could be found among European airports.

Airport	Type	Operator	Ownership Distribution
Amsterdam Schiphol	D	Schiphol Group	Dutch govt. 69.5%, city of Amsterdam 20.3%, city of Rotterdam 2.2%, ADP 8%
Berlin/Schönefeld	D	Flughafen Berlin Schönefeld, GmbH	German govt. 26%, state of Brandenburg 37%, state of Berlin 37%
Brussels	F	Brussels Airport Co.	Macquarie Airports 75%, Belgian govt. 25%
Copenhagen	F	Copenhagen Airport	Macquarie Airports 53.7%, Danish govt. 39.2%, free float 8.1%
Dublin	D	Dublin Airport Authority	Irish govt. 100%
Frankfurt/ International	E	Fraport AG	State of Hesse 31.5%, city of Frankfurt 20.1%, Lufthansa 9.9%, Artio Global Investors 10%, free float 28.5%
Greece (except Athens)	A	Ministry of Transport	Greek govt. 100%
Lisbon	D	ANA	Portuguese govt. 100%
London/Heathrow	F	BAA plc	Ferrovial 62%, Cdp du Quebec 28%, Baker Street Investment 10%
London/Gatwick	F	Gatwick Airport Ltd.	Global Infrastructure Partners (GIP) 100%
Madrid	D	AENA	Spanish govt. 100%
Manchester	D	Manchester Airport Group	Council of City of Manchester 55%, 9 Borough Councils 45%
Milan (Linate + Malpensa)	D	SEA	City of Milan 84.5%, region of Lombardy 14.5%, various 1%
Moscow/ Domodedovo	F	East Line Group	East Line Group 100%
Moscow/ Sheremetyevo	D	Joint Stock Company	Russian govt. 100%
Munich	D	Flughafen München, GmbH	German govt. 26%, state of Bavaria 51%, city of Munich 23%
Paris (de Gaulle + Orly)	E	Aéroports de Paris (ADP)	French govt. 52.4%, institutional investors 29.1%, Schiphol Group 8%, free float 8.5%, ADP employees 2%
Prague	D	Joint Stock Company	Czech govt. 100%
[illegible]	[illegible]	[illegible]	[illegible]

Vienna	E/F	Flughafen Wien AG	Province of Lower Austria 20%, city of Vienna 20%, company employees 10%, free float 50%
Zurich	F	Flughafen Zürich AG	Canton of Zurich 33%, city of Zurich 5%, unique AG and private shareholders 62%

Source: ACI—Europe, 2010.

TABLE 7.1 Ownership and Operators of a Sample of European Airports in 2009–2010

A word should also be said about *global airport operators*, a new type of airport operator that has emerged since the 1990s as a result of the airport privatization trend. Depending on how one counts, there were as many as 100 such operators worldwide as of 2012. Fraport is a prominent example and one of the largest such companies. It started out as the operator of Frankfurt/International. As of 2012, it held majority stakes at the airports of Antalya, Turkey (51 percent); Lima (70 percent); Varna and Burgas, Bulgaria (each 60 percent); and minority stakes ranging from 10 to 35 percent at the airports of Delhi, Xian, St. Petersburg, and Hanover. It also held management contracts at the airports of Cairo, Jeddah, Riyadh, and Dakar. The total number of passengers at all these airports (including Frankfurt) was close to 200 million. Fraport's global initiatives ("External Activities and Services") had total revenues of about €500 million in 2011 and contributed 31.7 percent to its EBITDA (Fraport, 2012). Interestingly, global airport operators, which in the 1990s were primarily offsprings of large European airports, now have their home bases spread all over the world. Important operators are based in Istanbul (TAV), Vancouver (Vantage), Bengaluru (GMR), Kuala Lumpur (Malaysia Airports Holdings), and Sydney (MAp Airports Ltd.), to name but a few.

In conclusion, model A can be viewed as the "traditional" form of airport ownership and management. For the most part, it will not work well at the busiest airports for many reasons, including lack of flexibility, large centralized bureaucracies, limited responsiveness to local issues, limited control over airport-specific finances, and few incentives to increase revenues or reduce costs. Model B at first glance seems to have similar disadvantages. However, the model has been mostly successful in the United States because of the way it is practiced. The American model B airports, although in theory operated by municipal or regional governments, are in practice among the most "privatized" in the world, because of their extensive reliance on outsourcing for the provision of facilities and services, as well as for planning and development functions. Another critical reason for these success stories is that airport budgets and finances are treated (as legally required) as independent of those of the local and/or regional governments that operate them. Airport revenues cannot be used to subsidize other city or regional services.

Model C can be appropriate for cases where there is lack of expertise in modern airport management. Thus, it applies best to either secondary airports in developed countries or to major airports in some developing nations.

The autonomous airport authority is the common element in models D through F. As airports become busier, more complex, and more important to local and national economies, these corporate entities have become increasingly commonplace, even in developing countries. An advantage of the autonomous airport authority is that it can accommodate any form of ownership: it can be government owned or privately owned or mixed. It is an institutional device that has proved largely successful in partially insulating airports from political interference and in promoting effective management.

7.3 Organizational Structures

Airport organizational structures necessarily reflect the diversity in ownership, management arrangements, and size that characterizes airports worldwide. They can only be reviewed in general terms with the objective of identifying some discernible patterns. A useful approach is to look at how organizational structures tend to evolve as the size of an airport increases.

Operators of modern commercial airports of any reasonable size must contend, to a greater or lesser extent, with a full spectrum of legal, financial, planning, public communications, administrative, human resource, environmental, engineering/technical, commercial, and operational issues. The organizational structure of even secondary airports must somehow reflect all of these areas of activity. This gives rise to the most generic of organizational charts of an airport operator shown in Fig. 7.1. The policy source in Fig. 7.1 depends on which of the models A through F identified in Sec. 7.2 applies. In the case where the airport operator is an autonomous airport authority—a common trait of models D, E, and F—the policy source is the authority's board of directors. In the case of models A and B, it is the leadership of the government agency or department to which the airport's management reports. Finally, in the case of model C, policy initiatives may come from some combination of the owner of the airport (typically a government entity) and its operator (the board of the airport management contractor)—with the dominant source depending on the level of autonomy granted to the management contractor.

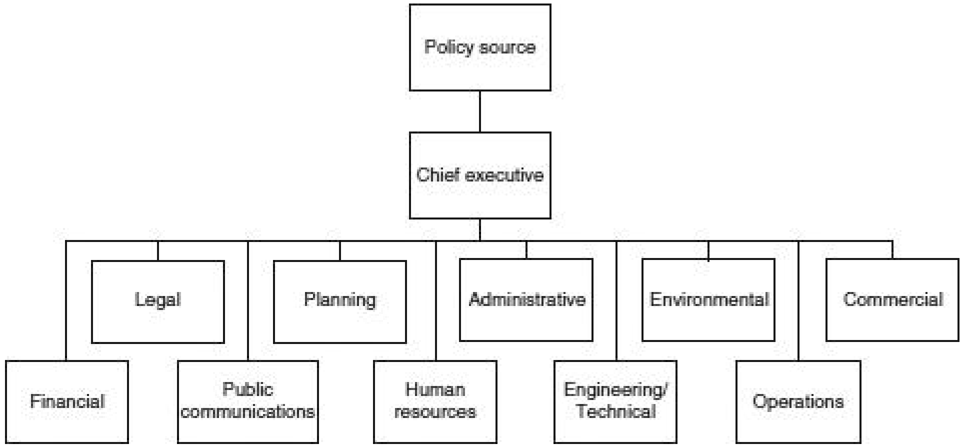


FIGURE 7.1 A generic organizational chart for an airport.

In larger airport organizations a sharp distinction often exists between *staff* and *line* units. Staff units support the chief executive in managing the airport. Line units, by contrast, are the ones that carry out the day-to-day tasks associated with the operation and serviceability of the airport's facilities. Staff units may be small but, because of their role in decision making, are often influential in determining an airport's economic and operational performance and in mapping its future course. Line units, by contrast, may employ hundreds—and, in some instances, literally thousands—of workers, especially when the airport operator is heavily involved in ground handling. The six units[8] on the left side of Fig. 7.1 clearly have a primarily staff role, whereas "commercial" and "operations" are primarily line units. "Operations" typically encompasses a wide range of functions that includes public safety (security, firefighting, emergency medical) as well as passenger handling and ramp handling (see Chap. 8). The two other units in Fig. 7.1, "environmental" and "engineering/technical," have both staff and line functions. They are sometimes organized into an "infrastructure and environmental affairs" department, a staff unit responsible for technical support on facility development and environmental issues, and an "engineering and maintenance" department, a line unit responsible for carrying out facility improvements and maintenance.

Depending on the circumstances, the relative importance of the organizational units varies, as does the depth and complexity of the organization chart. Secondary airports operated by autonomous or semiautonomous entities are usually structured along the "flat," single-tier lines of the organization chart of Fig. 7.1. In many cases, only one or two persons may staff some of the organizational units (e.g., legal and public communications).

At larger airports, the flat organization chart of Fig. 7.1 may take the more pyramidal shape of Fig. 7.2, which clearly distinguishes between staff and line responsibilities. The position of deputy director is a typical feature of more pyramidal airport organizations. It is designed to relieve the executive director from the task of overseeing daily airport operations and allow him/her to devote more time to strategic management. As noted previously, "handling" is typically a part of the operations department but may appear as a separate line unit ("handling department") in the organization chart of some of the airports that engage heavily in this activity.[9] A trend has also emerged in recent years toward giving more organizational visibility to public safety activities, because of their growing importance.

FIGURE 7.2 A two-tier organizational chart showing staff (support) units and line (operating) units.

In the case of airports operated by national, regional, or local governments (models A and B), the staff functions are typically performed by departments or "offices" located at the administrative headquarters of the governmental agency involved (e.g., a city hall), whereas the line units are positioned at the airports themselves. The staff units may support several airports simultaneously, possibly an entire national system.

More elaborate organizational structures can be found at many large airports operated by autonomous airport authorities. Figure 7.3 shows one that might be appropriate for an authority that operates several major airports[10] and can be viewed as an extension of Fig.

7.2. It indicates how line units must be replicated at each airport operated by the authority. The authority's headquarters provide shared staff support to all the airports.

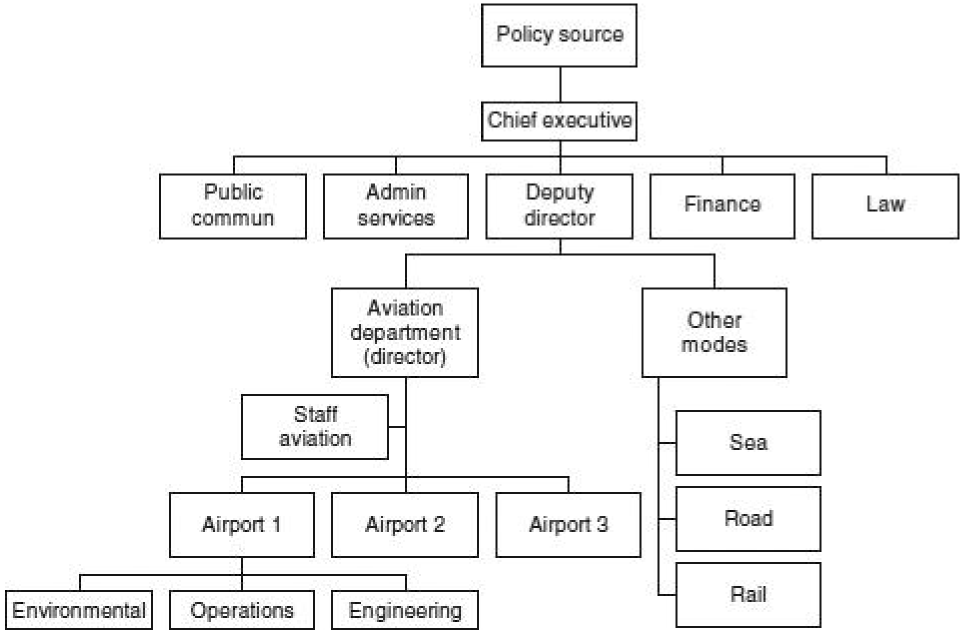


FIGURE 7.3 Organizational structure for a multi-airport authority. (*Source*: Wiley, 1986.)

At the highest level of airport size and organizational complexity, a number of mostly European and Pacific Rim airports are increasingly adopting, explicitly or implicitly, the profile of conglomerates of transportation, commercial, retail, and technical services. This shift to the role of provider of a hub of services implies a marked change in organizational structure. The traditional line and staff units of airport operators lose some of their centrality or are complemented on an equal footing by new units.

Example 7.2 Figure 7.4 shows the generic organizational structure[11] of the Massachusetts Port Authority (Massport). The staff (or "supporting") units and the line (or "operating") units are listed in columns to make the chart more compact.

FIGURE 7.4 Organizational structure of Massport.

Figure 7.5 shows the structure of the aviation department at Massport. Note that the director of aviation has a staff support unit dedicated to aviation headed by the director of aviation planning and development. Public safety and security has additional visibility and resources through a distinct line unit that reports directly to the director of aviation.

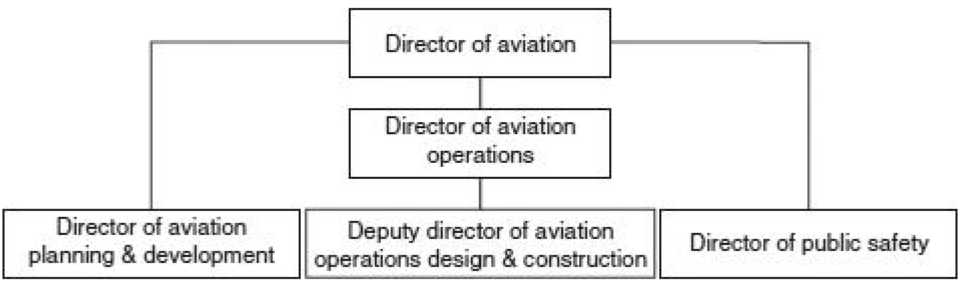


FIGURE 7.5 Aviation department of Massport.

Figure 7.6 identifies the airside responsibilities of the aviation operations department and shows how they are distributed among the functions of airside maintenance, airside operations, and airside construction.

FIGURE 7.6 Aviation operations department of Massport.

Example 7.3 The Amsterdam Airport Schiphol (AAS) Authority provides an interesting example of a provider of a hub of services. It has evolved into the "Schiphol Group," whose organizational structure in the early 2000s is shown in Fig. 7.7. In a visionary statement that anticipated a global trend, the Schiphol Group articulated in 1999 its objective of going "beyond the development of an efficient transport hub" and working toward realizing the concept of the "Airport City."

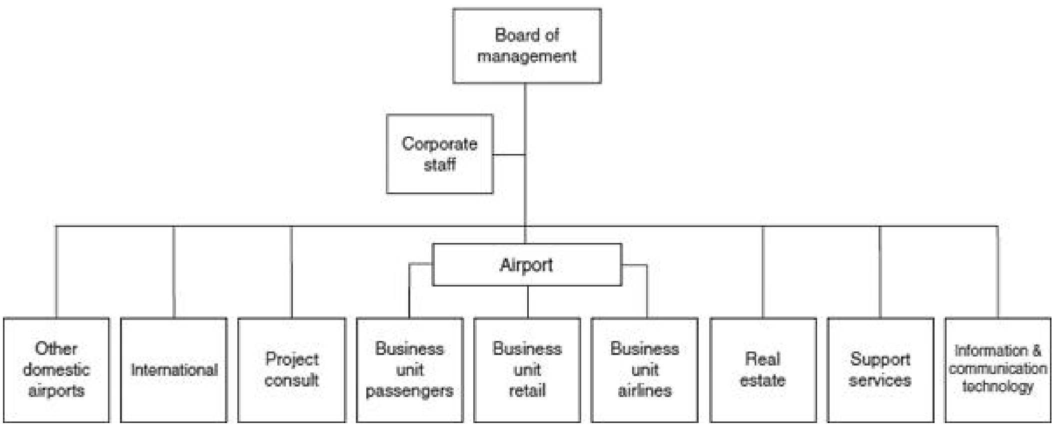


FIGURE 7.7 Organizational structure of the Schiphol Group.

The Airport City offers its passengers and visitors, but also the Schiphol based companies (airlines, distribution companies, and logistic and business service providers) 24-hours-a-day services in the field of shops, hotels and restaurants, multimedia services, corporate business and conference facilities, and recreation and relaxation… The Schiphol Group is particularly strong in the integration of all these activities into a coherent and durable entity that is closely knit with its geographical surroundings. This knowledge is also internationally in demand. Schiphol Group is therefore increasingly marketing the Airport City concept internationally. (Schiphol Group, 1999)

Note that of the nine units shown at the bottom of Fig. 7.7, only the middle three focus directly on the operation and management of Schiphol Airport. Beginning from the left, the "other domestic airports" unit is responsible for the small airports of Eindhoven, Lelystad, and Rotterdam. Schiphol International, BV, is a corporation overseeing development and management contracts at New York/Kennedy and at Brisbane, Australia. Schiphol Project Consult, BV, is another corporation specializing in project management and consulting at Schiphol itself, elsewhere in the Netherlands, and abroad. Schiphol Real Estate, BV, on the right, is one of the largest real estate agencies in the Netherlands and active in many aspects of real estate management and development at Schiphol and at other airports internationally.

Of the three units responsible for airport operations at Schiphol, the Business Unit Retail aims at the development and implementation of commercial activities, especially shops, restaurants, and hotels, for passengers and visitors. The organization charts for the other two units, Business Unit Airlines and Business Unit Passengers, appear in Figs. 7.8 and 7.9. These are responsible for all services and facilities offered to airlines, passengers, and visitors. Business Unit Airlines is also involved in marketing and account management vis-à-vis the airlines and third-party ground handling companies (see Chap. 8). Note that each of these two units is a large and complex organization in its own right.

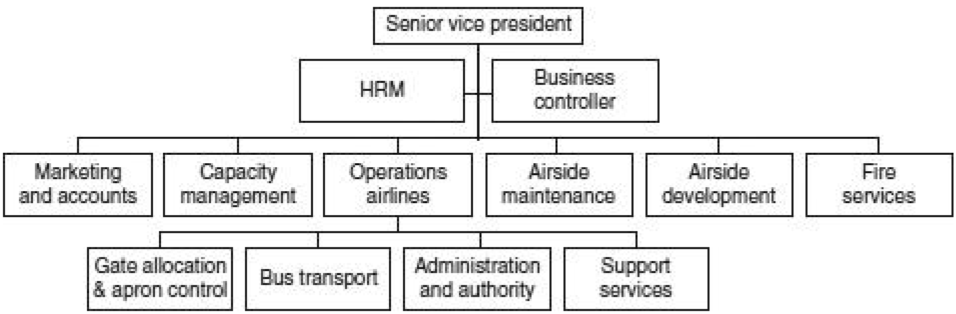


FIGURE 7.8 Business Unit Airlines, Schiphol.

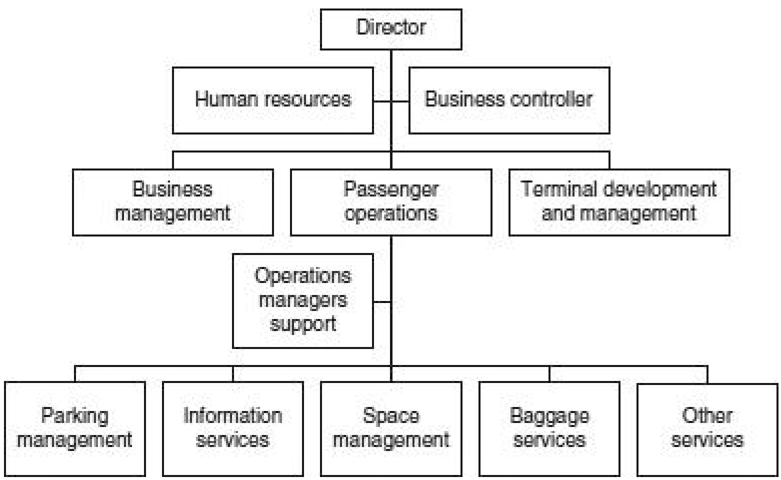


FIGURE 7.9 Business Unit Passengers, Schiphol.

Finally, the two rightmost units in Fig. 7.7 acquire and provide products and services in their respective areas for the entire Schiphol Group. Among the responsibilities of Schiphol Support Services, for example, are vehicle acquisition and management, utility services, airport security, access control and identification services, etc., that may be required by facilities managed and operated by the Schiphol Group.

By 2010, the structure shown in Fig. 7.7 had been consolidated without, however, altering its basic concept. Four "business areas" were established (Schiphol Group, 2011). Roughly speaking, Business Unit Airlines and Business Unit Passengers were merged into the "Aviation" business area. Because of its ever-growing importance, Business Unit Retail was elevated to a separate business area called "Consumers." "Real Estate" also became one of the four business areas. The three units on the left—Other Domestic Airports, International, and Project Consult—were merged to become the "Alliances and Participations" business area. The Support Services unit and Information and Communications Technology unit were absorbed, as appropriate in each case, by the four business areas.

7.4 Regulatory Constraints on Airport User Charges

Concerns regarding the potential abuse of the quasimonopolistic position that airports enjoy in serving origin/destination traffic have motivated widespread monitoring and considerable regulation of airport user charges. The airport privatization movement, in particular, has led to growing awareness of the need for regulatory safeguards to protect the public interest from potential conflicts with the goals of private airport investors. For example, private investors may be less willing than publicly held entities to invest in capacity expansion projects that may entail large capital expenditures. They may also be more likely to increase airport user charges at an airport from year to year, if given the latitude to do so.

A global consensus exists about the need to "contain" prices charged for aeronautical facilities and services, if necessary through regulatory measures. The ICAO Council has stated that airport operators may recover the *full cost, but no more*, of aeronautical facilities and services (ICAO, 2009). Full cost includes the cost of operations, maintenance, and management and administration, as well as interest on capital investment, depreciation of assets, and, when conditions permit, a fair return on investment. By contrast, the nonaeronautical revenues of airports are largely—but, in some countries, not fully—unregulated.

The extent of monitoring and regulation of aeronautical charges varies widely around the world. In the great majority of countries, charges are reviewed by a central government agency (e.g., Ministry of Transport, or CAA, or a competition authority) usually on an annual basis or upon request by airport operators or airlines. The review typically concerns: the rate of increase/decrease of charges relative to other cost indices; the "reasonableness" of the charges, particularly in comparison to what neighboring countries are charging for similar airports and services; and any special circumstances, such as large capital expenditures by the airport operator that may justify unusually large changes. In some cases (especially in the case of model A of Sec. 7.2) the government itself, not the airport operator, sets the charges and revises them periodically.

A number of countries, however, especially in Europe, Oceania, and increasingly in Asia, have set up specialized government bodies that oversee the regulation of airport charges. The best known and most active of these is the U.K. CAA. The regulatory philosophies espoused by these bodies may differ. For example, whereas the CAA has consistently practiced strict regulation of charges at BAA airports, the Australian Competition and Consumer Commission (ACCC) has adopted since 2002 a "probationary light-handed approach," choosing to intervene only in cases of alleged abuse of the charge-setting responsibilities of airport operators. New Zealand has essentially abandoned formal regulation of airport charges: the Ministry of Commerce can ask, as deemed necessary, for ad hoc reviews of charges by the National Competition Commission.

The three most common approaches used by regulators of airport aeronautical charges internationally are (1) specifying targets (or upper limits) on the rates of return on investment that airports may earn; (2) mandating upper limits ("caps") on the unit rates that an airport operator may charge in any particular year; and (3) restricting the annual rate of increase of unit charges. For example, the CAA regulators of the BAA specified until recently an approximately 7.5 percent target rate of return on net capital assets and placed limits on the annual rate of increase of the BAA's aeronautical charges. In other cases (e.g., Vienna) the limits on the annual rates of increase in aeronautical charges have been tied to traffic growth rates—the higher the growth rate, the lower the limit. A variety of other regulatory schemes, in a similar general spirit, are in effect at many major airports around the world, especially the ones with some degree of privatization. Overall, the setting of airport user charges continues to be an area of controversy and a constant irritant in the relationship between airport operators and airport users, leading to numerous heated disputes over the years and occasionally to litigation. This is not surprising in view of the vagueness of existing regulatory guidelines on this subject at the international level.

Example 7.4 A famous example of charges regulation is the formula used by the U.K. CAA to limit increases in BAA's aeronautical charges. At the time of BAA's full privatization in 1987, the Monopolies and Mergers Commission determined that the BAA's three London airports (Heathrow, Gatwick, and Stansted) would be regulated under a formula that specifies that the annual rate at which an airport can increase its *aeronautical* charges cannot exceed (RPI − x) percent, where RPI is the annual retail price index for the United Kingdom in the previous year.[12] Note that the intent of the (RPI − x) formula is to promote efficiency in airport operations. When x is 0 percent or greater, any improvements in airport financial performance must come from traffic growth and higher productivity, not increases in the real prices charged.

It has been the CAA's responsibility to review the value of x every 5 years. In 2000, Heathrow, Gatwick, and Stansted were restricted to (RPI − 3) percent, (RPI − 1) percent, and (RPI + 1) percent, respectively, that is, in the latter case $x = -1$ percent. This meant that changes in aeronautical charges at Heathrow and Gatwick had to be, at most, 3 and 1 percent, respectively, below RPI. Because the RPI in 1999 was about 2.9 percent, this also meant that aeronautical charges at Heathrow had to be reduced not only in real-price terms (as the increase was limited to 3 percent below RPI) but in current-price terms, as well.

Table 7.2 summarizes the history of *x* in the regulatory formula for Heathrow and Gatwick between 1998 and 2013. Note that the CAA allowed the BAA to increase aeronautical user charges at rates higher than the RPI (negative values of *x*) after 2003 at Heathrow and after 2008 at Gatwick. This was in recognition of the very large capital expenditures that were necessary at the two airports to increase capacity and improve their poor level of service. For the 2008–2009 fiscal year, for which the CAA made special provisions, Heathrow was allowed to increase aeronautical charges by up to 15.6 percent above RPI!

Period	Heathrow (%)	Gatwick (%)
1998–2003	+3.0	+1.0
2003–2008	–6.5	0
2008–2009	–15.6	–8.2
2009–2013	–7.5	–2.0

TABLE 7.2 The Value of *x* in (RPI − *x*) for London/Heathrow and London/Gatwick, 1998–2013

"Single Till" versus "Dual Till"

A critical question for regulators is whether aeronautical charges should be affected in any way by airport revenues on the nonaeronautical side. This is an issue that has generated a great deal of heated controversy. It is at the center of the debate outside the United States concerning the relative merits of the *single-till* and the *dual-till* approaches to setting aeronautical airport user charges. The difference between the charges that airlines would pay under each of these two approaches can be large. There is a parallel (milder) dilemma in the United States concerning the use of the *residual* versus the *compensatory* approach for setting aeronautical user charges (see as follows).

Airlines have argued in favor of the single-till approach. Under single till,[13] the nonaeronautical earnings of the airport may greatly influence the size of the aeronautical fees that the airport operator may charge.[14] In effect, the revenues from nonaeronautical services (which are usually highly profitable) end up offsetting a large portion of aeronautical costs under single till. This, in turn, means that the airport needs to collect less revenue from aeronautical charges in order to reach its regulated rate of return (or any other economic target specified) on aeronautical investment. Simply put, if in a particular year the airport needs revenue of $100 to meet its regulated economic return from the aeronautical facilities and services it provides and nonaeronautical activities contribute $40 toward meeting this target, the airport will need to raise only the remaining $60 from aeronautical charges. Thus, in the great majority of cases, the single-till approach will lead to reduced charges for aeronautical services. Nonaeronautical activities end up subsidizing, in a way, aeronautical ones. For example, the proceeds from duty-free sales will help reduce the landing fee to be charged under this scheme.

In contrast, airport operators generally support the dual-till approach. Under dual till, the aeronautical side of the airport's business is treated separately from the nonaeronautical one. The airport operator will set charges on the aeronautical side to achieve the regulator-specified economic targets solely though revenues from aeronautical facilities and services. At the same time, the airport operator is largely free to seek the maximum possible profit on the nonaeronautical side. Thus, in most cases, the dual-till approach will result in higher aeronautical user charges and higher overall profits for the airport operator than under the single-till approach. Full cost recovery plus a fair return on investment is achieved on the aeronautical side, while the profits from the unregulated nonaeronautical facilities and services accrue to the airport operator, as well.

The dispute regarding single till versus dual till is largely unresolved at this time. Numerous economic papers have argued each point of view—Czerny (2006) provides a good starting point. In popular terms, some of the principal arguments of airlines and others who support single till include the following:

- The traveling public is the eventual beneficiary of lower airline costs that result from single till, as some or most of the savings are passed on to consumers through fare reductions in a competitive system.
- Nonaeronautical activities exist and thrive at airports because of the passengers that airlines bring in. Without the airlines, no revenue or profit would be derived from these activities. Airlines are therefore entitled to some of the resulting benefits by paying lower user charges. In fact, the more price incentives airlines have at an airport, the more traffic and nonaeronautical business they will generate.
- Airport operators have monopoly power regarding nonaeronautical services and often apply monopolistic pricing to these services. They should not be allowed to benefit from such practices by retaining all resulting profits, as they can do under dual till.

Some of the main arguments on the opposite side include the following:

- Airlines are already protected from monopolistic pricing through regulatory caps on the rate of return that airport operators can seek from aeronautical facilities. There is no reason why airlines, which are for-profit organizations, should not be charged for the fair cost of the facilities and services they require.
- Only aeronautical revenues should be subject to economic regulation; the commercial side of the airport business should not be a factor in such regulation.

- At congested airports and at hubs they dominate, airlines are unlikely to pass on to their passengers a significant part of the cost savings they achieve through single till.
- Single till leads to charges for aeronautical facilities and services below the true costs; this results in economically inefficient use of airport resources, especially at congested airports.
- The profits generated from commercial activities at airports reflect the premium location for such activities that airports provide, not monopolistic pricing.

Example 7.5 A stark example of the effects of single till is provided by the history of landing fees at BAA's London/Heathrow airport. Until the British Airports Authority was privatized in 1987, the landing fees at London/Heathrow were among the highest in the world—and easily the highest in Europe. By 2000 they had declined to rank among the lowest charged by any major airport, including those in developing countries. For example, in 2000 the landing fee paid by a Boeing 747 operating during the peak period of the day in the busy summer season was about $660. This was a small fraction of the landing fee that the same airplane would have paid that year at practically any other major airport in Europe or in the Pacific/East Asia region. London/Heathrow is one of the world's busiest international airports and one of the most congested. Access to it is highly valued by airlines.

The explanation for this paradox lies in the combination of the single-till approach and the (RPI - x) formula prescribed by the BAA's regulators. With the highly profitable nonaeronautical activities at London/Heathrow contributing a growing fraction of total revenues, the share of revenue to be raised from aeronautical sources steadily diminished. Landing fees thus reached their low, by any criterion, levels. It has been estimated that if the BAA operated under a dual-till system and if it sought a 7 percent return on net assets from its revenue centers, landing fees at Heathrow airport would have been 35 percent higher in the late 1990s (Warburg Dillon Read—UBS, 1999).

Because of the subsequent changes in x, shown in Table 7.2, landing fees at Heathrow and Gatwick rose significantly after 2003 (see Chap. 12, Example 12.5). However, the single-till system kept them low compared to those of comparable major European airports.

Residual versus Compensatory

The approximate counterparts to the single- and dual-till approaches in the United States are known, respectively, as the *residual* and *compensatory* systems. The compensatory system is straightforward conceptually: The airlines pay charges that are sufficient to recover the full costs to the airport operator of the facilities and services that the airlines use. Under the residual system—and in a manner analogous to single till—airlines pay only the difference between the total revenue target and the revenues from all nonaeronautical and other sources.[15]

There is, however, an important difference between single-till and the residual system. To benefit from a residual system, airlines in the United States have to take on a financial risk. They sign long-term use agreements with the airport operator under which they underwrite the service of debt issued by the airport. The airlines essentially agree to cover any shortfall that may occur in servicing this debt. Under the compensatory system, by contrast, the airport operator assumes the full financial risk associated with servicing its debt.

Airlines that sign airport use agreements under the residual system are called *signatory airlines*. Among other rights, they can collectively approve or reject plans submitted by the airport operator for new buildings and other capital investments. Signatory airlines may enjoy significant savings in airport charges by paying only for residual costs. For example, airlines paid $0.51 per thousand pounds of aircraft landed weight at Los Angeles/International in 1992, the last year in which that airport used the residual system. In 1993, when the airport switched to a compensatory system, the rate tripled to $1.56 per thousand pounds. At the busiest residual airports, airline fees can be exceptionally low. In 1999, for example, Honolulu charged no landing fee because the residual revenues covered all airfield costs!

From the viewpoint of the airport operator, the preferred choice between a compensatory and a residual system of charges depends on the local circumstances. Generally, secondary airports and hubs dominated by one or two airlines (e.g., Cincinnati, Minneapolis/St. Paul, Atlanta, Dallas/Ft. Worth) are better off with a residual system because of their strong dependence on a few airlines to generate traffic. At the opposite end, operators of busy, mostly origin/destination airports with no dominant carriers (New York/Kennedy, New York/LaGuardia, Boston/Logan, Los Angeles/International, Seattle/Tacoma) will prefer a compensatory system. In general, as an airport becomes busier and less dependent on connecting traffic and on solely one or two airlines to generate traffic, it will be more likely to adopt the compensatory system. Several U.S. airports that fall somewhere in the middle of the spectrum have adopted so-called “hybrid” systems, which apply only a specified portion of nonaeronautical revenues (as opposed to 100 percent for residual systems and 0 percent for compensatory ones) to the computation of aeronautical charges.

7.5 Financing Capital Investments

Because of the large amounts of required capital, financing of large-scale airport projects is always a central concern of airport owners and operators. It is also a *constant* concern, as such projects are commonplace at fast-growing airports around the world. Most national governments continue to provide sizable grants and other financing in various forms to their airports. However, in the case of major airports, the *relative importance* of government financing of airport capital programs, as measured by the *share* of capital contributed, has been diminishing steadily in recent years. Two factors are responsible for this trend. First, as the number, size, and complexity of major airports grow, government resources are insufficient to meet the need for capital funds. Alternative sources must be sought. Second, air transportation has become a “mature industry” in many regions of the world and governments increasingly expect it to become self-sustaining, including paying—directly or indirectly—for its infrastructure needs, that is, for airports and air traffic management.

The typical sources of financing available to airports can be classified into the following broad categories.

Outright Government Grants

Outright government grants are still the most common type of airport financing in many countries. National, regional, and local governments worldwide recognize the direct and indirect benefits that economies derive from air transportation. They usually assign high priority to the financing of airport capital investments. In the United States, for instance, the federal government has been providing since 1946 sizable annual grants for airport development under a series of federal assistance programs, according to a changing set of fund distribution guidelines. The current program, called the Airport Improvement Program (AIP), has been operating since 1982. It distributed roughly $3.5 billion annually to airports from 2008 to 2011. However, most of the AIP funding does not go to the busiest airports. For example, in 2010, the 35 busiest airports handled close to 75 percent of all passengers but received less than 20 percent of AIP funds, or roughly $600 million out of a total of $3.5 billion. As these same airports spent about $6.3 billion on capital investment projects in 2010, the share of AIP funds was less than 10 percent, reflecting the trend mentioned earlier.[16]

The European Union also has several programs that provide funding for airport development in EU Member States and EU Candidate Countries. For example, the European Regional Development Fund and the EU Cohesion Fund made available €1.85 billion (roughly $2.3 billion) for airport development during 2007–2013.

Special-Purpose User Taxes

A common way to finance airport projects involves special-purpose user taxes imposed by national, regional, or local governments to finance local airport projects. These are taxes whose proceeds accrue *directly* to *specific* airport operators. A good example is the Passenger Facility Charge (PFC) that the Federal Aviation Authority (FAA) may authorize for certain airport operators in the United States (see Chap. 8). Revenues from the PFC at each approved site must be spent on projects or programs that enhance safety, security, capacity, and noise mitigation at that site. By contrast, another passenger departure tax limited to *international* flights from the United States is intended to raise funds for the entire national airport and air traffic management system and, as such, cannot be considered as targeted to financing any specific local airport projects.[17]

Low-Cost Loans from International or National Development Banks

A number of international and national development banks and funds specialize in financing major infrastructure projects through low-cost loans. This funding is generally intended primarily for airports in developing countries or in economically weaker areas of a region or country. The World Bank (International Bank for Reconstruction and Development), the African Development Bank, the Inter-American Bank, the European Bank for Reconstruction and Development (EBRD), and the European Investment Bank (EIB) are examples of international institutions that have been active in this sector. A number of countries have also established government-owned export credit agencies that may play a similar role. A typical example of the terms offered is a 1997 EIB loan of more than $1 billion for the construction of the new airport at Athens/Venizelos. The loan was for a 20-year term with a 6.05 percent annual interest rate (a low rate at the time) and a 7-year "grace period" on payments. Some loans to developing countries carry even more favorable terms, often requiring little more than repayment of the principal over an extended period of time.

Operating Surpluses

A fast-growing number of the busiest and economically strongest airports in the world have now reached the enviable position of generating sufficient economic surpluses to pay directly for small- and medium-size capital improvement projects without assistance from any form of external financing. This eliminates interest payments and the administrative and overhead costs associated with external funding.

Loans from Commercial Banks

Many large commercial banks, notably several Japanese ones, have been active in providing short- and medium-term (3- to 10-year) loans for airport capital projects. Such loans are attractive to qualified airport operators because of the flexibility they provide and their ready availability. However, when it comes to large-scale airport projects, only relatively small fractions of the financing requirements are typically covered in this way because of the higher interest costs involved.

General-Obligation Bonds

General-obligation bonds for financing airport capital improvements may be issued by national, regional, or local governments. As the name suggests, these bonds are secured through the full taxing power of the governmental entity involved. Should airport revenues prove insufficient to cover obligations to bondholders, taxpayers at large must cover the shortfall. Interest paid by these bonds is typically exempt from taxes of the issuing authority.[18] As a result of being tax exempt and highly secure, general-obligation bonds can usually be sold at low interest rates. They are thus particularly attractive to airport operators. However, governments are reluctant to finance through general-obligation bonds capital projects at airports with a large user base and revenue base. In the United States, where laws strictly limit the total debt that a local or regional government may secure through general-obligation bonds, the financing of airport projects through general-obligation bonds is progressively giving way to general *airport* revenue bonds (GARBs).

General Airport Revenue Bonds

GARBs may be issued by airports that are in a position to service debt entirely through their own revenues. As the name suggests, GARBs are serviced through the entire pool of revenues of the airport operator, not just those from a specific facility (e.g., a terminal). The critical difference from general-obligation bonds is that the taxpayers do not back up GARBs. The interest rates that these revenue bonds bear may thus be significantly higher than those of general-obligation bonds.

The rates depend critically on how secure these bonds are judged to be by credit rating agencies (see as follows). An important relevant parameter is the level of "coverage," essentially the ratio of net airport revenues to debt service requirements in any particular year. The higher the coverage, the lower the interest rate. The portfolio of long- and short-term airport use agreements that the airport operator has secured with the airlines may also be important in determining the ability of an airport to raise financing through revenue bonds. A summary of the factors used to assess the credit position of an airport is provided at the end of this section. As of 2011, no airport in the United States had defaulted on its GARB debt in the previous 60 years!

National law in many countries does not permit airports to issue revenue bonds. In others, including several European countries, such bonds constitute a relatively novel and unexplored way to finance airport capital improvements. In the United States, GARBs are the most popular means of airport financing. Some economically strong airports are, in fact, able to secure revenue bonds solely against their own earning power, without requiring airline guarantees. In a slight variation, the Port Authority of New York and New Jersey and the Massachusetts Port Authority, both of which own and operate seaports and other facilities in addition to airports, issue revenue bonds against earnings of the whole authority (not just the airport). Other airports, however, choose or are forced to secure GARBs against long-term airline use agreements under which the carriers commit to cover any shortfall in debt service. In return for backing up the bond issues, the airlines pay only for residual costs and obtain other important rights regarding airport management, operations, and planning (see Sec. 7.4).

Private Financing against Specified Rights to Airport Revenues

Private financing is fast becoming one of the principal means of financing airport capital improvements in both developed and less developed countries. This type of financing can take many different forms. In one of the most common arrangements, the airport operator signs a BOT (build, operate, and transfer) contract with a private group that undertakes to finance all or part of a development project against specified rights to its future revenues. This may involve just a single facility (e.g., a multistory automobile parking garage) or a complex (e.g., a new passenger terminal and supporting facilities) or, in a few instances, an entire airport. In this last case, the private group may become the airport operator, as well, for an agreed period of time.

In another arrangement, common in the United States, an airline may issue *special facility revenue bonds* (SFRBs) to finance the development of a major airport facility, such as a terminal building, for its own use. In contrast to GARBs, SFRBs are solely the responsibility of the airline and therefore entail significantly higher risk. For example, the bankruptcy of American Airlines in 2011 led to a downgrading to nearly "junk" status of the portfolio of SFRBs (amounting to more than $1 billion) for airport development that the airline had previously issued.

Examples 7.6 and 7.7 demonstrate that the financing of very large airport projects typically comes from a combination of sources and that the mix of funding depends strongly on local conditions.

Example 7.6 Construction of the Athens/Venizelos airport required capital of roughly $2.4 billion at the 1996 rate of exchange. This was obtained as follows:

- 47 percent ($1,128 million) from the European Investment Bank in the form of a low-cost loan under the terms described earlier in this section
- 15 percent ($360 million) from a consortium of commercial banks in loans at market rates
- 12 percent ($288 million) from a Greek national airport development fund endowed through a special-purpose user tax imposed on all passengers departing from the then-existing airport at Athens/Hellenikon
- 11 percent ($264 million) from grants from the European Union under the Second EU Convergence Program
- 7 percent ($168 million) from grants from the Greek State, including the value of the land parcel on which the airport was built
- 6 percent ($144 million) from share capital contributed by the shareholders (55 percent Greek State, 45 percent consortium of German companies)
- 2 percent ($48 million) from secondary debt taken on by the shareholders.

Thus about $720 million, or 30 percent of total financing, was essentially grant money (outright government or EU grants and a special-purpose user tax), whereas another 47 percent was obtained through a low-cost loan from the EIB.

Example 7.7 The financing of the construction of the Denver International Airport in the late 1980s and of Massport's large-scale airport, seaport, tunnel, and bridge modernization program centered on Boston/Logan offers an interesting contrast that reflects the different perceptions vis-à-vis the economic prospects of the two airports at the respective times the projects were undertaken.

Much of Denver's financing relied on public obligations. It came from: a federal grant of $500 million, specially authorized by the U.S. Congress; $400 million from PFC revenues at Denver; and contributions from the City of Denver in the form of $900 million in general obligation bonds, $300 million from the sale of the old Denver Stapleton Airport, and a $400 million commercial bank loan.

In sharp contrast to Denver, Massport avoided public obligations. Its financing relied on: $152 million from various federal grants of national (not specially targeted) scope; $420 million from Massport internal funds; $545 million from PFC revenues; $847 million from a series of GARB issues; and $939 million from third-party (public/private) development ventures.[19] Neither Massachusetts taxpayers nor any local governments incurred any debt service obligations for the Massport projects.

As already noted, the ability of airport operators to obtain favorable terms for the financing of capital investments depends in large part on their credit rating by specialized companies in this field, such as Moody's Investors Services, Standard and Poor's, and Fitch ICBA. Each of these companies has its own rating methodology for airports, but the factors they take into consideration are common to all. The list of these factors is instructive because it summarizes the attributes that also define an airport's overall economic prospects:

- Market strength (geographic location; regional economic characteristics, such as demographics, disposable income, etc.; origin/destination versus hub)
- Air traffic characteristics (air traffic forecast, range and market share of airlines at the airport; strength and commitment of these airlines to the airport)
- Physical infrastructure (utilization of existing facilities; need for new facilities; control of the gates by airport operator)
- Management and operations (cost recovery method and its adequacy to meet the airport's needs; contractual terms in airline agreements, concession contracts, etc.)
- Financing (existing debt burden; share of debt secured by general revenues, PFC, airlines; cash reserves)
- General context (political climate; environmental concerns, and disputes)

It is worth noting that airports have generally enjoyed higher credit ratings than airlines in Europe and the United Sates ever since the 1990s, especially during periods of economic crisis.

Exercises

7.1 Review the arguments for and against the single- and the dual-till approaches summarized in Sec. 7.4. How valid are these arguments? Can you think of any additional ones not on the list?

7.2. The six factors used by credit rating agencies to assess airports are listed at the end of Sec. 7.5. The list includes examples of specific criteria considered for each factor. Review these examples and identify the conditions under which each criterion would contribute to the assessment in a positive or a negative way. For example, an unstable political environment is obviously a negative. Several other cases in the list are not as simple.

7.3. Select an airport with which you are familiar and perform a qualitative assessment of creditworthiness using some of the criteria listed at the end of Sec. 7.5.

References

ACI-Europe, Airports Council International-Europe (2010) *Airport Ownership Report 2010*, Brussels.

Czerny, A. I. (2006) "Price-Cap Regulation of Airports: Single-Till versus Dual-Till," *J. of Regulatory Economics, 30*, 85–97.

Doganis, R. (1992) *The Airport Business*, Routledge, London, U.K.

Enriquez, R. (2002) "The Mexican Southeast Airport Group Privatization Case," S.M. thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Ernico, S., B. Boudreau, D. Reimer, and S. Van Beek (2012) *Considering and Evaluating Airport Privatization*, ACRP Report 66, Transportation Research Board, Washington, DC. www.rand.org

Fraport (2012) *Annual Report 2011*, Frankfurt a/Main, Germany.

ICAO, International Civil Aviation Organization (2009) *Policies on Charges for Airports and Air Navigation Services*, ICAO Doc. 9082/6, ICAO, Montreal, Canada.

Schiphol Group (1999) *Profile and Organizational Chart*, Schiphol Airport, Amsterdam, the Netherlands.

Schiphol Group (2011) *Annual Report 2010*, Amsterdam Schiphol Airport, Amsterdam, the Netherlands.

Warburg Dillon Read—UBS (1999) *Airports Review 1999*, Warburg Dillon Read—UBS, London, U.K.

Wiley, J. (1986) *Airport Administration and Management*, Eno Foundation for Transportation, Westport, CT.

---

[1] Massachusetts governors are elected to 4-year terms. Thus, governors elected to one term normally manage to appoint a majority of members (four of seven) only in the last year of their tenure. The intent of this appointment system is to insulate, as much as possible, the board of directors of Massport from political changes and pressures and thus ensure some level of continuity and stability in their policies.

[2] For a discussion of revenue bonds, see Sec. 7.5.

[3] For instance, airport authorities outside the United States rarely have the right to issue revenue bonds.

[4] As of 2012, the Spanish and Portuguese governments were considering privatization of some or all of these airports.

[5] The composition of the private group has changed significantly since its formation in 1996.

[6] Aeroporti di Roma sold its shares to ACSA in 2005, reportedly at a significant profit.

[7]The British government retained one "golden share" in BAA, which gave it decision-making power in fundamental policy matters.

[8]In practice, the names of these units differ from location to location; generic labels are used here to describe their function.

[9]*Handling* refers to the services provided to aircraft on the apron ("ramp"), such as loading and cleaning, and to passengers or cargo in airport buildings (see Chap. 8). Fraport's handling unit employed some 9000 people in 2011.

[10]The chart is patterned after the organizational structure of the Port Authority of New York and New Jersey at the time and is still valid conceptually today.

[11]Some specifics of Figs. 7.4 through 7.6 have changed over time; most importantly, there is much greater organizational emphasis on airside and landside security.

[12]The formula is actually somewhat more complex, but the (RPI $-$ x) part captures its essence; for details, see Enriquez (2002).

[13]Till = a drawer for storing money in a bank.

[14]The computation of these fees must, of course, also consider the constraints imposed by the economic regulator, such as any limits on the rate of return on capital invested in aeronautical facilities.

[15]"Other" sources may include aeronautical revenues raised from general aviation and from airlines that have not signed long-term airport use agreements ("nonsignatory" airlines).

[16]Ticket taxes paid by passengers at these airports contribute a very large portion of the proceeds of the Airport and Airway Trust Fund, from which AIP is primarily funded.

[17]In fact, only a minuscule fraction of the funds raised in this way is allocated back to the international airports where the funds were collected.

[18]In the United States, interest from general-obligation bonds issued by local or state authorities is exempt from both federal and state taxes.

[19]Of the total $2.9 billion, approximately $310 million were invested in seaport, tunnel, and bridge construction and improvements and the remainder in airport and airport-related capital improvement projects.

CHAPTER 8

User Charges

Airport operators derive most of their revenues from a wide variety of user charges. Practices differ considerably regarding what user fees are imposed and what rules are applied to compute these charges. The desirable attributes of any system of charges at an airport include transparency, adequate cost recovery, reasonableness, promotion of efficiency in the use of airport resources, and flexibility. A well-designed system is essential to achieving financial credibility and to being able to obtain funding for capital projects.

The process of developing a system of user charges is complex. It requires specification of policy guidelines, definition of revenue centers and cost centers, development of a detailed cost base, allocation of costs to revenue centers, development of a pricing methodology, and consultation with airport users.

Existing policy guidelines at the international level are rather vague and often lead to disputes between airport operators and airport users. In general, airport operators may recover the full cost of facilities and services, including the cost of operations, maintenance, management, and administration, as well as interest on capital investment, depreciation of assets, and, when conditions permit, a fair return on investment. Almost universally, the average-cost pricing method is used to compute unit charges. The method is simple and flexible, but it becomes problematic when it comes to pricing congested facilities.

User charges are classified into aeronautical and nonaeronautical. The various types of charges in these two categories are reviewed and explained. Major airports currently derive roughly as much revenue from nonaeronautical charges as from aeronautical ones. The latter were far greater than the former until the late 1980s. Comparing in a fair way the size of user charges at different airports is a difficult task because of the numerous factors that influence these charges, many beyond the control of airport operators. Accounting practices may play an important role in this respect. An example is the use of historical cost versus replacement cost in computing charges for depreciation of assets.

An area of significant international controversy concerns handling charges. Airport operators should strive to create a competitive environment for the provision of handling services.

8.1 Introduction

Airport operators derive most of their revenues from a variety of charges they impose on users of airport services and facilities.[1] Growing pressure to achieve economic self-sufficiency and, when permitted by the regulatory environment, profitability has led to the development of systems of charges that cover every aspect of airport-related activities.

User charges are also part of the direct operating costs that airlines and other aircraft operators face. As such, these charges are carefully reviewed and often criticized by airport users. They are the source of endless controversy, with disputes occasionally taking on an international dimension and involving national governments. An additional benefit of reviewing user charges—as is done in this chapter—is that through them one can also enumerate and become familiar with all the essential and ancillary services and facilities that are typically provided at an airport.

The contents of the chapter are as follows. Section 8.2 introduces the topic of airports as economic entities through a discussion of revenue centers and cost centers and outlines the methodology typically used to establish systems of user charges. This is an area in which international practices vary widely, partly because there is only limited guidance from such organizations as the International Civil Aviation Organization (ICAO) or from bilateral or multilateral aviation agreements (see Sec. 8.3). Sections 8.4 and 8.5 identify and describe in some detail the various types of aeronautical and nonaeronautical charges, respectively, and the rationale behind them. Section 8.6 discusses the size of the revenues that airports derive from these various sources in both absolute and comparative terms. Section 8.7 summarizes some of the difficulties one invariably encounters in attempting to compare charges across airports. It cautions against any facile comparisons. Section 8.8 concentrates on the often-controversial subject of ramp and passenger handling charges and explores the reasons for disputes regarding the provision of these services. Section 8.9 presents a detailed example of the computation of the most common—and still most important—airport charge, the landing fee. This also illustrates the application, advantages, and disadvantages of the average-cost pricing approach, which is widely used by airports throughout the world. A related topic concerns the use of historical cost at some airports and of current (or replacement) cost at others in computing the depreciation costs charged to airport users every year (see Sec. 8.10).

8.2 Cost and Revenue Centers

The development of a system of charges for airport facilities and services is one of the most fundamental tasks facing an airport operator. Growing pressure to achieve economic self-sufficiency and, when permitted by the regulatory environment, profitability, has led to the development of sophisticated systems of charges that cover every aspect of airport activ-

ities. This is not simply a matter of good accounting practices. A well-designed system of charges is also essential to achieving financial credibility. It is critical to an airport operator's ability to obtain funding for capital projects from banks and investors. The desirable attributes and general structure of such charging systems are outlined in this section.
A system of charges should ideally have the following attributes:

1. *Transparency*: Transparency encompasses several characteristics. First, the system should be *simple*, so that prospective airport users can readily understand how much they will be charged and what services and facilities they will be paying for. Moreover, it should be supported by adequate *documentation*, containing the data and explaining the line of reasoning used in developing the price structure. The documentation should also demonstrate a reasonable relationship between airport costs and prices charged. Finally, the charges should be *defensible* in the legal sense. They should not contravene national and international statutes and conventions or violate bilateral or multilateral agreements.
2. *Adequate cost recovery*: The prices charged for airport facilities and services should meet, with high probability, the airport operator's cost-recovery objective. This objective varies considerably across airport operators. Depending on the regulatory and economic environment in which the airport operates, some will act as profit maximizers, others will seek full recovery of costs—including, possibly, a fair return on investment—whereas a third group may seek only partial cost recovery, in instances when it is unrealistic to expect full recovery.
3. *Reasonableness*: User charges should be "reasonable" regarding several criteria. In absolute terms, they cannot be prohibitively high for the types of users that an airport wishes to attract. For example, it would be untenable for an airline to pay airport user charges that amount on average to, for example, 25 percent of the revenues of its flights. In relative terms as well, the amounts charged should not be out of line with those charged by comparable airports in the same or in neighboring countries. Reasonableness should also extend across different segments of aviation (e.g., foreign airlines should not be charged for a higher fraction of airport costs than their use of the airport warrants).
4. *Promotion of efficiency in airport use*: The fees charged to a user should be closely related to the true costs associated with that user. If a particular class of users is systematically subsidized or undercharged, these users will utilize the facility to a greater extent than economic efficiency would dictate. In developing a system of user charges that promotes economically efficient use, the cost of

delays at the airport must definitely be considered. Unfortunately, this is rarely the case in practice, as discussed later in this chapter and especially in Chap. 12.

5. *Flexibility*: Airport pricing systems must be flexible, so that user charges can be easily modified in response to change. Airports operate in a highly dynamic environment. As airline deregulation spreads worldwide, the pace of change, if anything, accelerates (see Chap. 2). Regulatory restrictions on how much charges can be increased or decreased from year to year, as well as certain long-term use agreements with major airlines, may severely constrain the ability of airport operators to adjust flexibly and dynamically to change. For example, because of regulatory restrictions in several European countries, a number of major airports could not modify their user charges sufficiently following the 1999 abolition of duty-free shopping for intra-European Union passengers and suffered serious economic losses as a result.

The development of a system of charges that satisfies all of the aforementioned criteria is obviously not an easy task. To carry out the process successfully, a number of steps are necessary, each requiring considerable effort on the part of the airport operator. The principal steps are as follows:

1. *Specification of general policy guidelines*: Top decision makers in the organization must participate in this step. It requires addressing a number of fundamental questions, such as: What are the overall financial objectives of the airport operator? What fraction of total airport revenues will be derived from aeronautical and from nonaeronautical sources? What will be the principal means of financing capital investments? What is the level of cost recovery that will be sought from each of the airport's revenue centers (see Step 3)? Will the airport consider congestion costs in developing its pricing structure?

2. *Development of a detailed cost database by item and by cost center*: Many airports still cannot fully account for their costs because they lack an adequate database. The ICAO recommends that airport costs be classified in matrix form by *item* and by *area of service* as shown in Table 8.1 (ICAO, 1989). Areas of service are often referred to as *cost centers*. In practice, the ICAO classification scheme may often be insufficiently detailed for the needs of busy and complex modern airports. Note, for example, that the list of areas of service does not include "aircraft stands," the area where ramp handling is provided. Aircraft stands are among an airport's most costly elements to develop, maintain, and operate, as well as major generators of airport revenues. Thus, numerous variations of the ICAO classification scheme are in use today. All these schemes provide an ac-

counting methodology for allocating item costs among areas of service/cost centers. For example, of the total amount paid for wages and benefits, what portion should be allocated to the aircraft movement areas, to the passenger terminal facilities, and to each of the other cost centers? This provides the basis for setting user charges in Steps 3 through 5 (see as follows) that reflect the true costs of the services and facilities provided. Note that in some cases (e.g., administrative overheads), the allocation of cost items to areas of service may be a difficult exercise.

By Item	By Area of Service (Cost Center)
Direct personnel costs (wages and benefits)	Aircraft movement areas (runways, taxiways, taxilanes)
Depreciation and amortization	Hangar and maintenance areas
Interest	Air traffic control and communications
Supplies and externally provided services	Meteorological services
Administrative overheads	Firefighting, ambulance, and security services
Other expenses	Passenger terminal facilities
Taxes	Cargo terminal facilities
	Other facilities and services

Source: ICAO, 1989.

TABLE 8.1 Classification of Airport Costs by Item and by Area of Service as Recommended by the ICAO

3. *Definition of revenue centers*: Most airports identify a small set of revenue centers, that is, groups of services and/or facilities, which are lumped together for the purpose of setting revenue targets and collecting user charges. One reason for defining such revenue centers is that certain activities and services at an airport can naturally be grouped together and it is convenient to apply a single charge, or a small number of charges, for each such grouping. For example, all aircraft movement areas (runways, taxiways, taxilanes) are typically grouped into a single revenue center, the airfield. A single charge, the landing fee, is then collected to cover all associated facility and service costs. A second, equally important reason is that airport operators often choose to put in effect different pricing policies for

different revenue centers. For example, most major airports aim only at recovering full costs from the airfield, but may attempt to extract the maximum possible profit ("charge what the traffic will bear") from "commercial concessions."

4. *Allocation of costs among revenue centers*: The airport operator must allocate the costs associated with each of the areas of service ("cost centers") to the revenue centers. This may be easy to do when a cost center is contained entirely within a single revenue center, so that all the associated costs can be immediately allocated to that revenue center, or it may require considerable work when a cost center overlaps several revenue centers simultaneously. Table 8.2 lists, as an example, one possible set of cost centers and revenue centers defined at a typical airport in the United States.[2] Note that all the costs associated with the "airfield area" cost center will be recovered through the "landing/traffic area" revenue center. However, costs associated with the "passenger terminal area" must be apportioned among "concessions," "airline leased areas," "other leased areas," "parking and car rentals," and possibly others.

Cost Centers	Revenue Centers
Airfield area	Landing/air traffic area
Passenger terminal area (including gates)	Airline leased areas (passenger and cargo)
Hangars	Other leased areas and rental land
Cargo terminals	Concessions
Other buildings and facilities	Parking and car rentals
General and administrative	Utilities
Miscellaneous	Fuel
	Miscellaneous

TABLE 8.2 Cost Centers and Revenue Centers at a Typical Airport in the United States

5. *Compute unit charges*: Computation of unit charges is often referred to as the *pricing step*. Steps 2 through 4 have defined by this point a set of revenue centers and have associated a cost with each of them. Moreover, the airport operator's policy guidelines (see Step 1) have specified cost-recovery targets for each revenue center. Given a demand forecast, it is then possible to compute what charges should be imposed on each unit of demand to achieve each revenue center's economic target. At many airports this computation is performed through a simple

cost-averaging approach, that is, essentially by simply dividing the revenue target by the amount of projected demand to compute a charge per unit of demand (see Sec. 8.9). However, more sophisticated approaches, such as marginal-cost pricing, may also be used, especially in the case of congested airports. It is also possible to apportion the costs associated with each revenue center to each of the various categories of users and thus develop different unit charges for each user category.

6. *Establish a framework for interacting with users*: Finally, it is essential to ensure the participation of airport users in the steps just described. Consultation with the users should be an integral part of the process, with regularly scheduled meetings and well-defined procedures for handling user comments and complaints.

This six-step process is dynamic and iterative. Policy guidelines, cost databases, cost assignment to cost centers and revenue centers, and pricing methodologies will necessarily be reviewed and revised over time in response to ever-changing conditions and to criticism from and the requirements of airport users.

8.3 Guidelines and Background for the Setting of User Charges

As might be inferred from the previous section, international practices regarding aeronautical and nonaeronautical user charges are enormously diverse, because individual airports or national government agencies must determine their own policies, cost bases, and revenue targets. This diversity has led to widespread conflict between, on the one hand, airlines and other airport users and, on the other, airport operators and civil aviation agencies. At the root of the problem is the fact that the regulatory and statutory framework of international civil aviation lacks specificity when it comes to the subject of user charges and, more generally, airport economics.

The November 1944 treaty of the Chicago Convention on International Civil Aviation, on which much of the international economic and technical regulatory framework of air transportation rests, pays scant attention to the issue of airport user charges.[3] Only Article 15, Chapter II, of the treaty mentions the subject of user charges. It stipulates that "uniform conditions" must be provided for use of airport facilities by aircraft of all Contracting States and that user charges "against international air services" must be *nondiscriminatory*. The treaty does not suggest any specific methodologies on which the computation of user charges should be based, nor does it refer to the notion of the "reasonableness" of user charges, which is often at the core of current disputes. It does, however, invite states to publish the charges they impose and to communicate them to the ICAO.

The so-called Bermuda 2 Agreement of 1977 between the United Kingdom and the United States was the first to spell out a fundamental principle that is used widely today in computing airport user charges (Doganis, 1992): User charges may reflect, but shall not exceed, the *full cost* to the charging authorities of providing appropriate airport and air navigation facilities and services, and may provide for a *reasonable rate of return* on assets, after depreciation.

This principle was adopted, for all practical purposes, by the ICAO 15 years later (ICAO, 1992) and echoed in the ICAO's *Policies on Charges for Airports and Air Navigation Services* (ICAO, 2009). This influential document, henceforth referred to as *ICAO Policies on Charges*, contains the ICAO's guidelines on the subject and stipulates that

- International users must bear their "full and fair" share of the cost of the airport.
- The "full cost" of the airport and its essential supplementary services should include the "cost of capital and depreciation of assets, as well as the costs of maintenance, operation, management and administration."
- "Airports may produce sufficient revenues to exceed all direct and indirect operating costs (including general administration, etc.) and so provide for a reasonable return on assets at a sufficient level to secure financing on favorable terms in capital markets for the purpose of investing in new or expanded airport infrastructure and, where relevant, to remunerate adequately holders of airport equity."

Note that the last statement acknowledges the fact that many airports seek to become self-sufficient economically and, where feasible, return a profit to their owners, be they governments or private entities.

The policies mentioned in *ICAO Policies on Charges* offer useful guidance on developing sound systems of user charges. They invite airports to maintain full financial records that provide a satisfactory basis for determining and allocating the costs to be recovered, as well as to publish their financial statements. Consultation with users should precede any significant change in charges or charging systems. Advance notice of at least 4 months before the proposed implementation date is recommended. The aim is to seek at an early stage user views on the changes and, if possible, to reach consensus on them. Failing an agreement, the airport operator should be free to impose the changes. Users should be notified at least 1 month in advance of the date when a final decision will take effect.

Additional aspects of the *ICAO Policies on Charges* are worth mentioning because they point to some of the relevant issues of contention on an international scale. The document offers valuable perspective by stating that "experience gained worldwide indicates that where airports and air navigation services have been operated by autonomous entities their overall financial situation and managerial efficiency have generally tended to improve . . .

The Council, therefore, recommends that, where this is economically viable and in the best interest of providers (airports and ANSPs) and users, States consider establishing autonomous entities to operate their airports or air navigation services."

A related issue concerns government subsidies. The ICAO Council does in fact encourage them, when appropriate, by urging that "in shaping policies toward airport finances, States should consider the broader economic implications of airports," such as their contribution to industrial development, cultural exchanges, tourism, etc. Indeed, practically every nation in the world subsidizes to some extent, directly or indirectly, the development of infrastructure at its airports. In a number of instances, the cost of airport *operations* may receive national subsidies as well.

The *ICAO Policies on Charges* also discourage the widespread practice, especially in developing countries, of reducing overtly or covertly airport charges to domestic carriers and compensating for this by overcharging foreign carriers. Note that this practice is in direct conflict with the principle of nondiscrimination in the Chicago Convention treaty. The ICAO Council states that an airline's capacity to pay "should not be taken into account until all costs are fully assessed and distributed on an objective basis." It is only at that stage that a state may decide to recover less than its full costs from an airline in recognition of other local, regional, or national benefits. In other words, governments may decide to subsidize national carriers by reducing the airport user charges they pay at domestic airports, but this should not be done at the expense of foreign carriers.[4]

The ICAO also expresses concern regarding the "proliferation of charges on air traffic." "Airport users should not be charged for facilities and services they do not use." Indeed, a number of nations have been using charges on civil aviation to pay, in part, for such items as radar installations primarily intended for military purposes or capital investments into new airports unrelated to the ones that the airlines paying the charges are using.

8.4 The Various Types of Airport User Charges

User charges at airports are classified into two categories: *aeronautical* and *nonaeronautical*. As the names suggest, the former are charges for services and facilities related directly to the processing of aircraft and their passengers and cargo, whereas the latter refer to charges related to the numerous ancillary commercial services, facilities, and amenities that are often available at an airport. This section identifies and discusses the principal types of aeronautical user charges.

Landing Fee

The *landing fee* is the most universal type of aeronautical user charge. It is the fee that aircraft pay for use of the airfield, that is, of the runway and taxiway systems of an airport.

The costs it covers include capital costs, operations and maintenance costs, and the cost of providing such services as firefighting, snow plowing, and security on the airfield.

In the overwhelming number of instances, the landing fee is computed with reference to the weight of the aircraft. Typically the maximum takeoff weight (MTOW) is used for this purpose,[5] but some airports, especially in the United States, use the maximum landing weight[6] (MLW) instead. The amount to be paid is derived in one of the following ways:

- In direct proportion to the weight (the most common, by far, approach in practice)
- As a fixed charge up to a specified weight threshold plus an amount that is directly proportional to any weight above that threshold
- In proportion to the weight of the aircraft but with a changing rate per unit of weight for different ranges of weight (e.g., \$*x* per ton up to 50 tons and \$*y* per ton for any weight above 50 tons)

All these alternatives are consistent with the *ICAO Policies on Charges*. They provide considerable flexibility to airport operators, who can develop landing fee "formulas" that allow them to encourage the presence of certain types of aircraft at their airports.

Example 8.1 In 2010, Athens International Airport (AIA) charged a landing fee according to the formula:

$$\text{Landing fee} = (\text{unit rate per ton}) \times (\text{MTOW in tons})$$

for aircraft with MTOW up to 120 tons. However, for aircraft with MTOW exceeding 120 tons, the landing fee was computed on the basis of the formula:

$$\text{Landing fee} = (\text{unit rate per ton}) \times (\text{MTOW in tons}) \times (120/\text{MTOW in tons})^{0.4}$$

Note that the intent is to have landing fees for wide-body aircraft (which, as a rule, have MTOW greater than 120 tons) increase less than linearly with aircraft weight. For example, with the unit rate of € 8.21 in force in 2010, a Boeing 737-400 with MTOW of 68 tons paid a landing fee of € 558, whereas the fee for an Airbus 340-600 with MTOW of 365 tons was € 1920. That is, the ratio of the respective landing fees was 3.44, whereas the ratio of the MTOW of the two aircraft is 5.37. This reflects the airport's effort to attract more wide-body aircraft.

Despite such flexibility, there is still considerable room for improving current practices. For one, the relationship between the weight of an aircraft and the costs associated with its operation on a runway and taxiway system is not particularly strong. Charging according to weight is, in fact, more related to ability to pay than to the true cost caused by an aircraft's operation on the airfield. Even more important in view of current conditions at many major airports, charges that are solely weight-based do not take congestion-related costs into consideration. A growing number of important airport operators are therefore examining alternative landing fee schemes that encourage efficient use of congested airfields. Such approaches are discussed in detail in Chap. 12.

An ongoing controversy regarding landing fees concerns the practice of charging different landing fees to aircraft depending on flight origin. Sometimes domestic flights or flights to/from certain nations receive preferential treatment by paying lower landing fees. The *ICAO Policies on Charges* are unambiguously negative about this practice, which cannot be justified on technical grounds. Nonetheless, it is used by a number of airports, including some of the busiest in the world—in effect, subsidizing domestic operations or certain airline routes. The European Commission has issued over the years requests to a number of EU airports to abandon this practice.

Terminal Area Air Navigation Fee

Many airports charge a fee for the cost of *terminal area* air traffic management (or "air traffic control" or "air navigation") services and facilities provided to arriving and departing aircraft, including the cost of runway and taxiway lights, airport radar, instrument landing systems, and other landing and traffic control aids (see Chap. 13). A terminal area covers a volume of airspace that typically extends to a radius of 50 to 80 km around a major airport or a multi-airport system.[7] The airport operator collects the air navigation fee on behalf of the air navigation service provider (ANSP), which in most cases will be a national civil aviation authority or similar body. In some instances, however, the airport operator may have paid for part or all of the cost of the local ATM facilities and equipment. Revenues from the air navigation fee may then be shared proportionately between the airport operator and the civil aviation authority. The terminal area air navigation fee is usually collected as part of the landing fee. This fee does not exist at airports in the United States, where most ATM costs are paid through taxes and fees on the tickets of air passengers (see the section Passenger Service Charge later).

Aircraft Parking and Hangar Charges

Aircraft *parking and hangar charges* are imposed for the use of contact and remote apron stands and, if applicable, hangar space. Parking and hangar charges are typically proportional either to the weight of the aircraft or to its dimensions. At many airports, there is no parking charge for "normal" use of a stand, that is, for occupancies of less than a specified amount of time. This time may vary by type of stand—with a stricter limit sometimes applicable to contact stands. Typical time limits for free aircraft parking range from 2 hours to as many as 6 hours. At contact stands, there may be an additional charge for use of the "aviobridges" (or "jetways") and at remote stands for use of mobile staircases. However, mobile staircases are more often paid for as part of the handling charges (see as follows).

Airport Noise Charge

Noise charges have spread rapidly in recent years (see Chap. 6). The charge often varies by time of day, with considerably higher amounts typically applied to nighttime operations.

Noise charges are usually collected as part of the landing fee. However, some airport operators choose to collect them separately, in order to demonstrate both to airport users and, especially, to the airport neighbors their commitment to addressing noise-related concerns.

The original purpose of noise charges was to cover the costs of mitigation measures that many airport operators have been forced to adopt. These range from installation of noise-monitoring equipment around the airport to noise insulation of public-use buildings and private homes. A more recent parallel use is as a demand control mechanism that penalizes the noisiest aircraft and offers a discount or rebate to the least noisy ones. For purposes of the noise charge, aircraft are typically subdivided into a small number of categories according to their noise characteristics, with a different charge applied to each category. The categories were initially along the lines of the classification of aircraft into Stage 1 through 4 in the United States (or Chapters 1 through 4 of the ICAO, see Chap. 6). However, more elaborate subdivisions are increasingly being defined at the national or regional level. Several European airports, for example, subdivide Chapters 3 and 4 aircraft into several different categories according to their noise characteristics and charge them at different rates.

The *ICAO Policies on Charges* are somewhat ambiguous on the subject of noise-related charges, reflecting the tension between less developed countries, many of whose airlines operate older and noisier aircraft, and developed ones, which place a premium on noise mitigation. The *ICAO Policies on Charges* state that such charges "should be levied only at airports experiencing noise problems," that they "should recover only the costs of noise alleviation," and that they "should be non-discriminatory and not prohibitively high for the operation of some aircraft"—obviously referring to the Stage/Chapter 1 and 2 aircraft still operated by many airlines in developing nations.

Emissions-Related Charges

A number of airports (e.g., Bern, Geneva, Stockholm, Zurich) have imposed charges related to pollutants emitted by aircraft engines that affect local air quality (LAQ). For instance, in 2010 Stockholm charged about $7/kg of estimated emitted NO_x. The imposition of emission-related charges will probably become more widespread in future years, primarily in Europe, in response to expressed concerns about the impact of airports on LAQ (see Chap. 6).

Passenger Service Charge

Passenger service charges are also known as *terminal service fees* and are intended to cover costs related directly to the use of passenger buildings. Their application and method of collection varies considerably from country to country. For example, in the United States, no such fee is usually applied to domestic passengers (who account for more than 90 percent of total traffic), as they mostly utilize passenger buildings (or parts thereof) that are operated by the airlines themselves under long- or short-term leases. In such cases the airport operator has no claim to any costs to be recovered. However, U.S. airports apply a charge to all *international* passengers to cover the costs of federal inspection services (FIS), such as immigration, customs, and health. Moreover, in cases where international passengers utilize terminals whose space and gates are shared by several (U.S. and non-U.S.) airlines, a passenger service charge for the general use of the terminal may be applied, in addition to the FIS-related passenger charge.

The collection of the passenger service charge is not a trivial matter. In the great majority of cases today, the airlines, rather than individual passengers, pay this fee to the airport operator. The airlines then simply add the relevant amount, typically as a separate identifiable item, to the price of their tickets to recover the fee from passengers. The amount paid by the airline to the airport is computed on the basis either of the actual number of passengers on each flight or of a previously agreed "estimated number" of passengers per flight. However, there still exist cases (e.g., international passengers departing from some airports in developing countries) in which the airport operator collects the fee directly from the individual passengers. This is sometimes due to the airlines' refusal to have anything to do with fees deemed arbitrary or excessive. This method of collection can be annoying to airport passengers and causes delays and queues.

Site-specific taxes on passenger tickets, whose proceeds accrue *directly* to airport operators, constitute another form of passenger service charge. A prominent example is the Passenger Facility Charge (PFC) that the Federal Aviation Administration (FAA) is empowered to authorize at airports in the United States. To obtain authorization to collect a PFC, an airport operator must submit to a complicated application and review process. The PFC applies to both domestic and international passengers. Revenues from the PFC at each approved site must be spent on projects or programs related to enhancing safety, security, capacity, and noise mitigation at that site.

Many countries impose taxes on air passenger tickets that are not site-specific. These constitute a more general form of taxation and should not be confused with the airport-specific passenger charges. They are intended to support the development of aviation infrastructure at large or, in some cases, to simply augment a government's overall tax revenues. An example is the passenger ticket tax[8] that the United States has used for many years to maintain the Aviation Trust Fund (ATF), which supports improvements in aviation infrastructure, in general. In contrast to the PFC, this passenger ticket tax should *not* be viewed as an airport passenger service charge.

Cargo Service Charge

In a manner entirely analogous to the passenger service charge, many airport operators impose a fee per ton of freight (or other agreed measurement unit) to cover the cost of cargo processing facilities and services provided by the airport. This fee is typically collected from the carriers, which, in turn, may pass on this cost to their customers. Cargo service charges are not imposed at airports in the United States.

Security Charge

Aviation security services are provided at essentially every commercial airport in the world, and a corresponding charge is almost universally collected. Following the events of September 11, 2001, airport security arrangements everywhere have become much tighter and are updated and modified continuously, in response to any incidents or perceived new threats. Not coincidentally, security-related expenditures for new equipment and personnel have increased at a very rapid pace, with airlines and their passengers shouldering most of the costs through direct or indirect charges. At airports where a security charge is not explicitly identified as a separate fee, it is usually collected as part of the general-purpose passenger service charge. The *ICAO Policies on Charges* indicate that security should be a "State responsibility," that "authorities may recover security costs but no more," and that "users requiring additional security services may be charged additionally for them." These have indeed become the guidelines under which security services are generally provided and paid for. The service provider is typically the national police or other government security agency. However, in many cases, especially at busy international airports, the national government (or the regional or local government in some nations) may relegate responsibility for the provision of security services either to the airport operator (which hires special personnel for this purpose) or to a specialized third-party contractor. In all circumstances, eventual responsibility and oversight rests with national, regional, or local government authorities, as the case may be. The security charge itself is typically imposed on a per-passenger basis, collected by the airport operator from the airlines, and distributed among the providers of the security services according to local circumstances.

Airlines requesting additional security services beyond the standard ones are usually charged an additional amount. An example is departures by U.S. and non-U.S. carriers from overseas airports for destinations in the United States. These flights are required by the U.S. government to exercise additional security precautions, such as passenger interviews or intensified screening, and are typically charged additional security fees at overseas airports. In other cases, some flights may be designated as "high risk," either by the airport operator or by the airport's national government. Such high-risk flights are often charged for the cost of the extra security measures to which they are subjected, even though the airlines performing these flights may not have requested the extra services in the first place.

Two other types of aeronautical charges—ground handling charges and en route air navigation fees—deserve mention because of their typically significant size. Unlike the aeronautical charges already described, airport operators are not usually the recipients of these charges.

Example 8.1 (continued) It is becoming increasingly commonplace to compare the "total aeronautical charges per passenger" that airports impose. This, supposedly, makes it possible to identify "expensive" airports (but see Sec. 8.7 for the pitfalls associated with such comparisons). An example is given in Table 8.3 for AIA in 2010.

Charge for	Euros
Landing (prorated per departing pax)	5.08
Aircraft parking (prorated per departing pax)	1.31
Passenger terminal facilities	12.16
Security	5.00
Loading bridge (prorated per departing pax)	1.41
Ground power (prorated per departing pax)	0.33
Baggage handling system	1.92
CUTE computer system	0.32
Check-in desk	0.20
Airport development fund (ADF)	
EU destinations	12.00
non-EU destinations	22.00
Total, EU destinations	39.73
Total, non-EU destinations	49.73

*Assumption: 110 passengers (75 percent load factor), 60-minute stay, contact gate at Main Terminal Building.

TABLE 8.3 Aeronautical Charges and Taxes per Departing Passenger on a B737-400,* Athens International Airport, 2010

To prepare an estimate of total aeronautical charges per passenger, it is necessary to make several assumptions concerning the reference aircraft, the number of passengers on it, the time when the aircraft operates (as charges may vary according to time-of-the-day at some airports), etc. For instance, Table 8.3 assumes that the reference aircraft is a Boeing 737-400 operating at a 75 percent load factor and using a gate of the Main Terminal Building of AIA for 60 minutes. Note that some of the charges shown, such as the landing charge and the aircraft parking charge, are imposed on a per-aircraft basis (and must be subsequently "prorated" by dividing by the assumed number of passengers on the aircraft) whereas others are imposed on a truly per-passenger basis. In the particular case of AIA, the total of these "passenger-based" charges—primarily consisting of the passenger terminal facilities charge, the security charge, and the airport development fund (ADF) charge—is far greater than the "aircraft-based" charges.[9] This is consistent with a general worldwide trend toward more emphasis on "passenger-based" charges.[10]

The method of collection of these charges also varies across airports. All the charges in Table 8.3, except for the last ("ADF charge"), are paid by each airline to the airport operator. The airlines then pass on these charges to their customers through the prices they charge for their base airfares. The ADF charge, on the other hand, is essentially a tax imposed by the Greek State on air tickets to pay for past and future airport development in Athens and elsewhere in that country. It is therefore added as a separate item to the tickets of passengers using AIA and is denoted explicitly as one of the taxes, which are added to the airfare.

Ground Handling Charges

The diverse and important category of *ground handling charges* is subdivided into *ramp handling charges*, that is, charges for handling services on the apron ("ramp"), and *traffic handling charges* for services that are provided within the passenger or cargo buildings. Traffic handling services are sometimes further broken down into *passenger handling* and *cargo handling*. The loading and unloading of aircraft, as well as the sorting, bundling, and delivery of baggage to retrieval carousels and other devices, are all considered part of ramp handling. Only a relatively small (and diminishing) number of airport operators are currently providing ground handling services. Instead, these services are typically provided by airlines (self-handling or from one airline to another) or by specialized companies that are licensed for this purpose by airport operators. As a result, the recipients of ground handling charges are usually entities other than airport operators. Handling costs to the airlines are often comparable in size to the *total* of all the other aeronautical charges mentioned previously. Because of their importance and the considerable controversy that sometimes surrounds them, ground handling services and charges are discussed more extensively in Sec. 8.8.

En Route Air Navigation Fee

This last type of aeronautical charge is not an airport fee *per se*, as it pertains to the provision of ATM services in en route airspace, that is, outside an airport's terminal airspace. Typically, the proceeds from this fee go to national[11] civil aviation authorities (formally, ANSP) or similar bodies, which are generally responsible for ATM facilities, equipment, and operations in en route airspace.

Charges for en route air navigation have increased rapidly over the years. They now constitute an important part of airline costs outside the United States, whereas they were rather insignificant up to the late 1970s. The *ICAO Policies on Charges* stipulate that the costs to be taken into account when determining ATM user charges "should include only those related to services and facilities approved under the relevant Regional Air Navigation Plan[12] of ICAO." After recognizing that attribution of en route air navigation costs to users is a difficult task, the ICAO recommends that en route charges should take into account the distance flown and the aircraft weight, the latter "in less than direct proportion." Most countries around the world have complied with these guidelines by now. They have adopted approaches for computing en route air navigation charges that are usually identical to or variations of the "EUROCONTROL formula." This formula computes a fee for each flight in direct proportion to the number of *service units* that flight incurs. Specifically, the number of service units, *n*, incurred by an aircraft with a MTOW of *T* tons flying a great-circle distance of d km[13] in the en route airspace of a EUROCONTROL Member State is given by the expression

$$n = \frac{d}{100} \times \sqrt{\frac{T}{50}} \tag{8.1}$$

(As an example, an aircraft with a MTOW of 200 tons flying a great-circle distance of 300 km in French en route airspace is charged for 6 service units in France.) Each Member State imposes its own unit charge per service unit.[14]

The United States is one of only a few countries that do not impose en route air navigation charges. The cost of ATM facilities and services is paid for primarily through the Aviation Trust Fund (funded for the most part by the passenger ticket tax and the passenger segment tax—see earlier discussion of the passenger service charge) and, to a lesser extent, through general taxes.

8.5 Nonaeronautical Charges

Airport revenues are also derived from *nonaeronautical charges*. These are also often referred to informally, and somewhat imprecisely, as *commercial revenues*. The importance of nonaeronautical revenues has been growing steadily over the years, as operators of both major and secondary airports have increasingly directed their efforts toward maximizing income from nonaeronautical sources. Many of the busiest airports in the world currently derive more revenue from nonaeronautical than from aeronautical charges. Profit margins from this sector are also typically much larger than those on the aeronautical side. In fact, the *ICAO Policies on Charges* recommend, with unusual candor, that "full development of revenues of this kind be encouraged"[15] (ICAO, 2009)—one of the few points on which airports and airlines seem to be in full agreement.

The reasons for the enormous success of commercial activities at airports are several and quite obvious. First, large numbers of travelers and their greeters pass through busy airports each day, creating a huge potential market. Second, air passengers come, on average, from the more affluent strata of society and include a large fraction of "high-end" business travelers. Third, many air travelers find themselves with lots of free time on their hands in airport terminals—often more than an hour in the case of departing and transfer passengers (see Chaps. 15 and 16). Moreover, the duty-free and tax-free shops, which are typically available to international passengers, make airports particularly inviting places (to some) to spend money. Passengers of certain nationalities are culturally attuned to buying presents for their relatives and friends at home when returning from trips abroad—and the final airport stop before they get home is a particularly convenient place for doing so.[16] The many sources of nonaeronautical revenue can be classified into a few major categories described briefly as follows.

Concession Fees for Aviation Fuel and Oil

Suppliers of aviation fuel and oil at an airport pay a fee, typically an agreed percentage of gross revenue, to the airport operator. In some cases, the airport operator may itself buy the fuel and resell it to the aircraft operators. The *ICAO Policies on Charges* recommend that fuel concession fees be treated in the same way as "nondiscriminatory, aeronautical charges," that is, that only a modest return be sought from such concessions, in view of the importance of fuel and oil costs to the airlines and other aircraft operators.

Concession Fees for Commercial Activities

Concession fees include fees for the operation of duty-free shops, retail shops, bars and restaurants, bank and currency exchange branches, newsstands, game arcades, etc., on airport premises, mostly in passenger buildings. These fees are charged either on a fixed-rent basis for space provided or, very often, on a variable-rent basis. The latter may involve a fee ranging from 20 percent to as high as 60 percent (!) of gross sales, usually supplemented by a guaranteed minimum annual level of revenue for the airport operator.

Revenues from Car Parking and Car Rentals

Car parking and car rentals are fast-growing sources of revenues for airports around the world. They are often the largest generators of nonaeronautical revenues at airports in the United States. Arrangements regarding automobile parking facilities and services vary considerably across airports. In many cases the airport operator itself will build, manage, and operate the car parking facilities. As an alternative, the airport operator may build the facilities but then contract the management and operation of car parking to a specialized operator. A third, increasingly common arrangement involves a BOT ("build, operate, and transfer") agreement with a contractor who undertakes to finance, construct, manage, and operate the car parking facility over a specified period of time, typically of the order of 10 to 25 years.

Similarly, a variety of arrangements are in place regarding the provision of car rental services at airports. For example, space for the stationing of rental cars may be collocated with the regular car parking facilities, or provided at remote locations on the airport's premises, or relegated to locations outside the airport property.

Rental of Airport Land, Space in Buildings, and Assorted Equipment

A wide range of possibilities is included under the category of rentals. The most obvious are revenues derived from space rented to airlines for offices and "VIP passenger" lounges, as well as from facilities and equipment rented to shippers, freight forwarders, etc. Revenues from advertising space may also be significant. Airport property and land may be rented for the development of aircraft maintenance and repair hangars for the airlines and for fixed-base operators (FBOs). Major airports also make often-complex arrangements for the development of hotels, office buildings, and even shopping centers on their property, thus generating sizable rental and concession revenues.

Fees Charged for Airport Tours, Admissions, etc.

Once a significant source of revenue, fees charged for tours and similar activities have now become entirely secondary as revenue sources at all but a few locations.

Fees Derived from Provision of Engineering Services and Utilities

Revenues from services and utilities provided by the airport to airlines, concessionaires, and other users are rapidly growing at major airports, in contrast to the previous category.

Nonairport Revenues

The term *nonairport revenues* refers to income from an airport operator's activities that are not directly related to either aeronautical or ancillary (i.e., nonaeronautical) services at the operator's principal airport(s). As such, nonairport revenues could plausibly be treated as a distinct category from "nonaeronautical revenues." Examples of such activities (see also Chap. 7) include provision of consulting, educational, and training services to other airport operators; management contracts for operating terminal buildings, technical departments, or even entire airports elsewhere; real-estate ventures and holdings; subsidiaries operating duty-free shops, hotels, or restaurants; equity investments into various, mostly travel-oriented commercial ventures; and acquisition of shares in other airports, in connection with various privatization schemes (see Chap. 7). A number of major airport operators, especially in Europe and increasingly in Asia (but not in the United States), are becoming deeply engaged in such off-airport activities.

8.6 Distribution of Airport Revenues by Source

One of the more striking developments in the history of airports worldwide has been the dramatic growth of nonaeronautical revenues during the 1980s and 1990s. Growth of these revenues exceeded by a significant margin growth in aeronautical revenues during that period. By 2000 the total size of nonaeronautical revenues at the busiest airports in the world was about equal, on average, to that of aeronautical revenues. However, this trend was arrested (with many exceptions, especially in developing regions of the world) during the first decade of the twenty-first century; mainly because of the various shocks that the world's economies suffered between 2000 and 2010, the balance between aeronautical and nonaeronautical revenues stayed roughly constant overall. In the United States, in fact, aeronautical revenues grew faster than nonaeronautical ones during that decade.

Example 8.2 Table 8.4 provides an approximate breakdown of the revenues of Massport at the Boston/Logan airport in 2010. Note that approximately 43 percent of the revenues came from aeronautical charges, 34 percent from nonaeronautical ones, and 23 percent from nonoperating revenues. About half of the nonoperating revenue is derived from the PFC. Almost the entire amount of $23 million of "other" nonoperating revenue was raised through a newly instituted customer facility charge (CFC), which is a special charge of $6 on car rentals intended to support the development of a consolidated car rental facility at the airport. The CFC illustrates a new way of funding certain types of facilities at airports. Remarkably, the single largest revenue source at Boston/Logan was automobile parking and car rentals. As can be seen from Table 8.4, the combined revenue from these sources, plus the $23 million from the CFC, amounted to nearly 30 percent of total revenue in 2010—nearly double the amount contributed by landing fees despite the fact that Boston/Logan charges airlines for the full cost of the airfield (compensatory system, see Chap. 7). This underscores the importance of the automobile in the economics of airports in the United States. Another noteworthy point is the roughly equal (56 vs. 44 percent) split in the contributions of aeronautical and nonaeronautical revenue to total operating revenue in 2010. Overall, Boston/Logan generated nearly 90 percent of the total revenues of Massport in 2010, the remainder coming from Boston's seaport and a number of other transportation facilities.

Source	Revenue ($ US million)	% of Partial Categories	% of Total Revenues
Operating revenues			
Aeronautical revenues			
Terminal rental charges	87	20	16
Landing fees	92	21	16
Cargo and other hangar rentals	23	5	4
Others	41	10	7
Total aeronautical revenues	243	56	43
Nonaeronautical revenues			
Land and nonterminal facility leases	5	1	1
Terminal concessions	23	5	4
Rental cars	29	7	5
Automobile parking	118	27	21
Other	16	4	3
Total nonaeronautical revenues	191	44	34
Total operating revenues	434	100	77
Nonoperating revenues			
Interest income	15	12	3
Grant receipts	30	24	5
PFC	59	46	11
Other	23	18	4
Total nonoperating revenues	127	100	23
Total revenues	561		100

TABLE 8.4 Sources of Revenue for Boston's Logan International Airport, 2010

The pattern suggested by Example 8.2 and Table 8.4 for Boston/Logan is not atypical. Tables 8.5 and 8.6 show, respectively, the combined revenues and expenses in fiscal year (FY) 2010 at the 35 busiest airports—defined as "large hubs"—in the United States, as reported to the FAA through Form 5100-127 that all U.S. airports are required to complete each year (FAA, 2012). Revenues and expenses are each broken down into "operating" and "nonoperating" categories, with operating revenues further subdivided into "aeronautical" and "nonaeronautical." Total revenues amounted to $13.842 billion and total expenses to $12.320 billion, with operating revenues (operating expenses) accounting for 78 percent

(81 percent) of these totals. Note that 59 percent of the operating revenue was derived from aeronautical sources and 41 percent from nonaeronautical. This is in significant contrast to the situation in 2000, when aeronautical and nonaeronautical revenues each accounted for almost exactly 50 percent of operating revenues and underscores the fact that, in the first decade of the twenty-first century, there has been a reversal in the United States of the trend of the previous 20 years toward an ever-increasing share of nonaeronautical revenues. The four largest single sources of revenue in FY 2010 were landing fees, terminal rental charges, automobile parking, and the PFC. As in Example 8.2, the combined revenue from parking and car rentals exceeded the revenue from landing fees! Another noteworthy point is that the 35 large hubs that accounted for almost 75 percent of total U.S. passenger traffic in 2010 received $626 million, or less than 20 percent of the approximately $3.5 billion in airport grants that were distributed by the federal government that year, in connection with the Airport Improvement Program (AIP), which subsidizes airport capital projects and is reviewed and approved on an annual basis by the Congress. This reflects the emphasis, in the United States, on directing federal aid to smaller airports, based on the assumption that larger airports are better able to fund capital projects through the surpluses they generate.

Source	Revenue ($ US million)	% of Partial Categories	% of Total Revenues
Operating revenues			
Aeronautical revenues			
Terminal rental charges	2,897	27	21
Landing fees	2,243	21	16
Cargo and other hangar rentals	339	3	3
Others	842	8	6
Total aeronautical revenues	6,321	59	46
Nonaeronautical revenues			
Land and nonterminal facility leases	224	2	2
Terminal concessions	1,066	10	8
Rental cars	778	7	5
Automobile parking	1,753	16	13
Other	609	6	4
Total nonaeronautical revenues	4,450	41	32
Total operating revenues	10,771	100	78
Nonoperating revenues			
Interest income	227	9	2
Grant receipts	626	21	4
PFC	1,915	62	14
Other	253	8	2
Total nonoperating revenues	3,071	100	22
Total revenues	13,842		100

TABLE 8.5 Sources of Revenue for the 35 Airports with the Largest Number of Passengers in the United States in 2010

Category	Revenue ($ US million)	% of Partial Categories	% of Total Revenues
Operating expenses			
Personnel costs	2,684	27	22
Communications and utilities	648	6	5
Contractual services	2,245	22	18
Other	1,251	13	10
Subtotal	6,828		55
Depreciation of assets	3,171	32	26
Total operating expenses	9,999	100	81
Nonoperating expenses			
Interest expense	2,321	99	19
Other	34	1	0
Total nonoperating expenses	2,355	100	19
Total expenses	12,354		100

Table 8.6 Expenses at the 35 Airports with the Largest Number of Passengers in the United States in 2010

On the side of expenses (Table 8.6), depreciation and interest expenses accounted for about 45 percent of the total, underlining the point that major airports are very capital-intensive facilities. Contractual services also accounted for almost one-fifth of total expenses, an indication of the emphasis on "outsourcing" at U.S. airports. Note that the net surplus (= total revenue - total expenses) of the 35 large hubs in 2010 was roughly $1.5 billion, after paying for interest costs. It is this surplus that enables these airports to largely self-finance their more routine capital expenditures.

In related statistics, the 35 large hubs had a total indebtedness at the end of FY 2010 of $62 billion, raised $8.7 billion from bond issues during the year and had capital expenditures and construction in progress worth a total of $6.3 billion—all very large numbers, indicative of the size and importance of the airport industry in the United States.

Analogous patterns vis-à-vis the relative size of aeronautical and nonaeronautical revenues can be observed outside the United States. In Asia, for instance, aeronautical and nonaeronautical charges contributed almost exactly equally (51 vs. 49 percent) to the total operating revenues of major airports in 2009 (ATRS 2012), but there were significant differences within this group: some airports like Singapore, Seoul/Incheon, and Hong Kong derived well above 60 percent from nonaeronautical sources, whereas others lagged far behind in this respect. It is difficult, however, to compare directly airport revenues and expenses across different parts of the world. Unlike the United States, financial performance data for European and Asian airports are often not readily available. In addition, accounting practices and the classification and definitions of revenues and expenses vary widely from airport to airport. Subject to these *caveats*, Table 8.7 provides information on the distribution of revenues for a set of 190 airports that served about 80 percent of European airport passengers in 2009 (ACI-Europe, 2010). The classification used is significantly different from that of Table 8.5 and far less detailed when it comes to aeronautical revenue, where airline-related charges include landing fees, aircraft parking charges, etc. and passenger-related charges consist of passenger service charges, security charges, etc. Overall, aeronautical and nonaeronautical revenues are roughly equal (53 vs. 47 percent, respectively). In contrast to the United States (see Table 8.5), “retail concessions” plus “food and beverage” account for 33 percent of total nonaeronautical revenues, almost twice the share of “car parking” plus “rental car.” The large “other” category of nonaeronautical revenues includes revenues from provision of utilities and from off-airport activities (consultancies, off-airport real estate, management contracts, etc.). Note that depreciation of assets is not included in the listing of expense categories.

	Revenue ($ US million)*	% in Category	% of Total Revenues
Aeronautical revenues			
Passenger-related (primarily passenger terminal and security charges)	11,870	61	32
Aircraft-related (primarily landing and aircraft parking fees)	7,590	39	21
Total aeronautical revenues	19,460	100	53
Nonaeronautical revenues			
Retail concessions (includes duty/tax-free shopping centers, restaurants/bars, currency exchange, etc.)	4,740	28	13
Food and beverage	850	5	2
Car parking	2,370	14	7
Rental cars	510	3	2
Property income/rents	3,050	18	8
Advertising	340	2	1
All others (engineering services, utilities, refueling, consultancies, off-airport work, security services to third parties)	5,080	30	14
Total nonaeronautical revenues	16,940	100	47
Total revenues	36,400		100

*Converted from euros at $1.40 per euro.

Source: ACI Europe's *Economics Report 2010*, 2010.

TABLE 8.7 Estimated Total Operating Revenues at 190 *European Airports* in 2009

It should also be noted that Table 8.7 does not include additional revenue of approximately $2.5 billion raised by certain airport operators (e.g., Düsseldorf, Hamburg, Milan, and Vienna) mostly in Austria, Germany, and Italy, through their ground handling operations (see Sec. 8.8).

8.7 Comparing User Charges at Different Airports

It is extremely difficult to compare fairly the magnitude of user charges at different airports. In every case one has to understand well the numerous factors that affect the setting of these charges. One must also look carefully at the detailed description of the charges and at what each charge pays for. This is a perfect example of a case where one must truly read the "fine print." Facile comparisons typically lead to erroneous conclusions. Unfortunately, many publications are replete with those.

Some of the factors that significantly influence the magnitude of user charges at an airport are summarized next (this is far from an exhaustive list).

Government Funding

As noted in Sec. 8.3, practically all governments recognize the national, regional, and local benefits generated by airports. However, the extent to which they provide direct or indirect funding to airports varies greatly around the world. Obviously, the higher this kind of funding, the lower the user charges will need to be. Government funding can take the form of: direct grants-in-aid (mostly for capital improvement projects); special-purpose taxes whose proceeds go directly to airports; general funds benefiting aviation as a whole; and preferential tax treatment. One must also carefully distinguish between funding based on taxes or fees that governments impose specifically on aviation users and funding that comes from general tax revenues. The former essentially redistributes user payments, with governments deciding how to allocate the funds among alternative aviation-related programs. In the United States, for example, the AIP, which provides grants to airports for capital investments, is funded from user fees and taxes collected from passengers and aircraft operators through the Airport and Airway Trust Fund. However, funds derived from general tax revenues amount to subsidies of airports and other aviation-related services. For instance, the tax exemptions enjoyed by general-obligation bonds and revenue bonds that finance airport capital projects can be viewed as indirect subsidies. Such direct and indirect subsidies are still commonplace in most countries.

Coverage of Charges and Quality of Services Offered

The services that airport users pay for through any particular type of user charge may vary significantly from airport to airport. An example is the landing fee. At many European and Asian airports this fee includes a substantial charge for terminal airspace ATM services. By contrast, in the United States these services are paid for through a ticket tax and from general tax funds; a charge for ATM services is *not* included in the landing fee. Obviously, the quality of facilities and services offered (as measured by level of comfort, absence of delays, reliability, etc.) also varies immensely from airport to airport.

Volume of Traffic

Unit charges (e.g., the landing fee per ton of aircraft weight) are typically computed by averaging costs. This means that the unit charge is derived essentially by dividing the total costs to be recovered by the number of units of demand at each cost center of the airport (see Sec. 8.9). Most airport facilities and services are characterized by decreasing marginal costs (at least up to a point) as demand increases.[17] These economies of scale mean that airports with large volumes of traffic enjoy an advantage regarding user charges. Airports in the United States have some of the lowest landing fees in the world, partly for this reason.

Size and Timing of Capital Expenditures

Airport charges are affected to a great extent by the size of capital expenditures. Large investments into airport infrastructure (airfield, terminal buildings, supporting facilities and equipment, ground access infrastructure) will naturally result in larger airport charges. Similarly, recent capital expenditures will usually generate larger depreciation costs and, consequently, larger charges. By contrast, older facilities may already be fully or mostly depreciated and will not contribute substantially to the size of user charges.

Characteristics of Traffic

Certain types of traffic are inherently less costly to accommodate and process than others. Airport unit costs are therefore strongly influenced by the composition and characteristics of the traffic. For example, passengers on dense domestic "shuttle" routes, for example, New York–Boston or Frankfurt–Munich or Madrid–Barcelona, can be processed very efficiently with minimal space requirements for passenger buildings, because of short dwell times, simplicity of check-in, few checked bags, etc. (see Chaps. 15 and 16). Once again, U.S. airports enjoy a major advantage in this respect because of the large volume of domestic traffic that moves through them.

General Cost Environment

The general cost environment within which an airport operates is the most obvious and, possibly, the most important factor affecting user costs. The costs of construction, equipment, energy, and technology vary greatly from country to country. The cost of acquiring and maintaining technologically advanced airport equipment is often extremely high in developing nations. National labor regulations and practices deserve special mention, as personnel costs often are the dominant component of airport expenditures. Western European airports, for example, typically operate in environments with high pay scales, generous health and other benefits (vacation days, paid holidays, etc.) for their workers, strong labor unions, and limited flexibility in task assignments.

Accounting Practices

Accounting practices, in themselves, can also be important in determining the size of user charges. An example, discussed in detail in Sec. 8.10, is depreciation schedules based on historical cost versus those based on current cost. Airports utilizing current-cost accounting charge much higher depreciation costs than historical-cost accounting airports.

Treatment of Nonaeronautical Revenues

Finally, airport and/or national policies regarding the treatment of nonaeronautical revenues are also very important. In the most obvious example, at airports utilizing single-till and residual systems (see Chap. 7), the airlines pay only the difference between the airport costs attributed to them and the revenues that the airport collects from commercial activities and from all other airport users. Hence, charges to the airlines and to other aeronautical users appear to be smaller at these airports. By contrast, airlines pay for the full cost of the facilities and services they use, including a fair return on investment, at compensatory and dual-till airports. Chapter 7 discusses these points in more detail.

In conclusion, and in view of all these difficulties, only some qualitative, general statements can be made regarding the relative magnitude of airport user charges internationally. For example, the highest landing fees and passenger terminal charges, as of 2011 (and over the previous two decades), can be found in Japan and in some Western European countries. However, many airports in Eastern Europe and in developing countries also have surprisingly high aeronautical charges, typically reflecting low levels of traffic and the large investments of capital needed to upgrade their often-obsolete facilities. High charges, unfortunately, tend to hinder further the growth of traffic. By contrast, aeronautical charges are among the lowest in the world at U.S. airports, when it comes to domestic traffic.[18]

At the same time, the "ranking" of airports regarding the magnitude of their aeronautical charges tends to be highly unstable over the years. For example, London/Heathrow and London/Gatwick moved from being among the airports with the highest landing and passenger service charges in the world in the early 1980s to the ranks of modestly priced airports in the late 1990s and early 2000s, as a result of increases in productivity, strict regulation, and the application of a single-till approach (see Chap. 7). However, they moved back to being "expensive" airports in the 2010s, because of regulatory changes and large capital expenditures.

ICAO Document 7100 (ICAO, annual) and other surveys of user charges at the world's major airports (see, e.g., IATA, monthly) provide the data and documentation necessary for updating conclusions of this type. In any event, readers should be aware that user charge comparisons—no matter how popular in the media, professional publications, and (even) regulatory proceedings—are often of dubious value for the many reasons described in this section.

8.8 Ground Handling Services

Ground handling services are typically subdivided into ramp handling and traffic handling, as noted in Sec. 8.4. The former are essentially airside services provided mostly on the apron ("ramp") and the latter are landside services at passenger buildings and in and around cargo terminals. The providers of handling services are referred to simply as "handlers."

Table 8.8 lists the principal ground handling services at airports. Note that traffic-handling services do not typically include passport control, immigration, customs, security control, and health inspection—functions generally performed by government organizations or their contractors. It should also be emphasized that several different handlers may provide the ramp and traffic handling services for any single flight. Moreover, at many airports, the principal ramp or traffic handlers may not be involved at all in the provision of some of the services listed in Table 8.8. Notable examples are catering transport and aircraft fueling, which may be provided directly by food caterers and by aviation fuel and oil concessionaires, respectively.

Traffic Handling Services	Ramp Handling Services
Passenger handling	Baggage handling and sorting
Ticketing	Loading and unloading of aircraft
Check-in	Interior cleaning of aircraft
Boarding supervision and services	Toilet service
Executive lounge/"club" operation	Water service
Cargo and mail handling	Passenger transport to/from remote stands
Some information services	Catering transport
Preparation of various handling and load-control documents	Routine inspection/maintenance of aircraft at the stands
Various supervisory or administrative duties	Aircraft starting, marshaling, and parking
	Aircraft fueling
	Aircraft deicing

Table 8.8 Principal Ground Handling Services at Airports

At every airport, one or more of the following types of provider may offer ground-handling services to any specific airline:

- The airport operator.
- The airline itself ("self-handling").
- Another airline ("third-party handling").
- An independent operator (not an airline) specializing in ground handling. These independent handling companies are sometimes referred to as *fixed-base operators* (FBOs).

The market share of each of these four types of handlers varies greatly across regions of the world. In the United States, airport operators at major airports do not perform ground handling. By contrast, in Austria, Germany, and Italy, not only do many airport operators offer the entire range of ground handling services, but they also enjoyed, until the late 1990s, government-sanctioned monopolies for the provision of all or some of these services. In general, the number of airport operators that are providers of ground handling services has been gradually diminishing, as these operators have difficulty competing with specialized ground handlers. However, a significant number of airport operators in Europe,[19] Asia, and Africa are still heavily involved in ground handling.

Similar differences exist regarding handling performed by the airlines. Whereas self-handling, as well as third-party handling, is routine at practically every North American airport—and indeed constitutes the most common mode of ground handling—such arrangements are either prohibited at many airports elsewhere or require stringent and time-consuming review and approval procedures. Moreover, in a number of countries only the national airline has the right to self-handle and to provide third-party services.

Rules and practices can be even more diverse when it comes to independent handlers. Typically, these handlers must receive a concession (often in response to a competitive tender) to operate at an airport and must demonstrate that they satisfy a set of requirements, such as technical know-how, security clearances for their personnel, access to adequate insurance, special vehicle permits, etc. The airport operator will usually also impose a set of service standards that the handler must meet. The independent handler may also sign a concession contract that provides for payment of a fixed or a variable fee to the airport operator. Airlines that perform third-party handling are often required to pay similar fees to the airport operator and adhere to similar service standards.

Overall, the typical situation at North American airports is that at least one and usually several airlines provide third-party services while, at the same time, several independent FBOs offer these services, as well. However, practices on payment of fees by the third-party airline handlers and/or the independent handlers to the airport operator vary considerably. Payment of a fee to the airport operator by airlines and independent handlers (where such handlers are allowed) for the right to provide ground-handling services is standard practice outside North America.

The issue of competition in the provision of handling services has generated considerable controversy. Several regulatory authorities have tried to ensure that any—not self-handling—airline will have more than one options when choosing a provider of handling services at any airport (e.g., the airport operator and an airline providing third-party handling) on the assumption that competition, regarding both cost and quality of service, will be beneficial to the airlines and their passengers, in the long run. For example, in the late 1990s, the European Union issued a Directive mandating competition for the provision of ground handling services at EU airports, which required that: (1) self-handling be permitted and facilitated as of January 1, 1998, at airports with more than 1 million passengers or 25,000 tons of freight per year; (2) third-party and/or independent handling be permitted as of January 1, 1999, at airports with more than 3 million passengers or 75,000 tons of freight per year (Deutsche Bank, 1999). Characteristically, this Directive met with considerable resistance from some EU Member States, with some major airports (e.g., Frankfurt/International) requesting and eventually being granted time extensions in implementing it.

This reflects the fact that the financial rewards of ground handling operations may be high, especially in cases where airport operators or national airlines enjoy monopoly rights. Another reason for the contentiousness that surrounds ground handling is its labor intensiveness. Traditional airport operators with heavy involvement in handling employ large numbers of personnel. For instance, at their peaks, the ground-handling branches of Frankfurt/International and of Rome/Fiumicino were employing about 9000 and 3500 people, respectively. Downsizing these handling activities can thus be painful in human, economic, and political terms.

It is difficult to estimate the "typical" charge for handling any particular type of aircraft because reliable data are hard to come by and many service providers treat such data as confidential. It is also standard practice to offer heavy discounts to airlines with large traffic volumes. Undoubtedly, large differences also exist in the size of handling charges at low-cost and high-cost airports. As a rough indication, reported costs for handling services for a routine turnaround of an aircraft in 2010 ranged from $800 to $4000 for narrow-body aircraft, such as a B737 or A320, and from $3000 to $13000 for a B747 on an intercontinental flight. Note that, with the exception of airports that provide ground handling services, the revenues from ground handling charges do not appear in the balance sheets of airport operators as they accrue to entities (airlines or independent ground handling companies), which are different from the airport operator.

A related issue concerns the ideal number of competitors for the provision of ground handling services at an airport. Because significant economies of scale are associated with such services, having too many competitors runs the risk of instability (as some of the competitors may be driven out of the market) and of deterioration in service quality due to pressure to reduce costs. An anecdotal rule of thumb, whose origin and factual basis are hard to trace, states that a handler providing a full range of services requires a volume of about

4 to 5 million passengers per year to have a commercially viable operation. Accordingly, an airport with 15 million passengers per year would ideally award a total of at most three permits to third-party and/or independent handlers to operate on its premises.

The number of companies that specialize in ground handling services has increased considerably over the years, mirroring the growth of air traffic. Several of these companies have an international presence at many airports. Some are owned in part or fully by a major airline. Because ground handling can be very profitable under the right conditions, there have been a number of instances in which tenders for handling concessions at major airports have given rise to legal challenges by unsuccessful bidders contesting the validity of the outcomes.

8.9 Landing Fee Computation: Average-Cost Pricing

The *average-cost method* is by far the most commonly used method to compute the unit charges imposed at the various revenue centers of an airport. Economists have widely criticized its use at airport facilities that are congested—as a growing number of them currently are (see the end of this section). Despite this deficiency, the application of average-cost pricing is almost universal at both uncongested and congested airports worldwide. The method is very simple conceptually and also makes it easy to update the unit charges from year to year—two of the main reasons for its popularity.

To apply average-cost pricing, one needs to specify the *revenue centers* of an airport and their associated *cost centers*. Once this has been done, the following three-step procedure is applied to each of the airport's revenue centers separately, or, if desired, even to specific subelements of the revenue centers.

Step 1. At the beginning of any defined time period—typically the beginning of a fiscal year—the airport operator determines a *revenue target, X*, for the revenue center in question. This target is set in accordance with the airport's overall economic policies. For example, *X* may be equal to the *full costs* of the facilities and services, including a fair return on investment, when a dual-till system or compensatory system is in use (see Chap. 7). Or it may be less than the full costs, if airport policy calls for preferential treatment of the users of certain types of facilities or if it is decided that full-cost recovery would lead to unreasonably high charges. Or, finally, it may be equal to the *residual cost* to be raised from the revenue center, as is often the case at airports in the United States that use a residual system or those outside the United States using the single-till system[20] (see Chap. 7).

Step 2. For the same time period, typically the next fiscal year, a forecast of demand, *Y*, is prepared for the same revenue center. Different units of demand would apply to different parts of the airport. For example, "number of tons of aircraft weight" expected to land at the airport might be the natural units in which to forecast demand for the runway and taxiway system, whereas "number of passengers" might be appropriate for computing the passenger service charge for use of passenger buildings.

Step 3. The charge, Z, per unit of demand for the time period of interest is then determined by simply dividing X by Y, that is, $Z = X/Y$.

Example 8.3 illustrates a typical application of the average-cost pricing method to the computation of the landing fee in the United States.

Example 8.3 Table 8.9, based on simplified and modified actual figures, shows the computation of the landing fee rate for FY 2010 at a hypothetical Airport AP, operated by a state-owned airport authority in the United States. Airport AP uses a compensatory system of charges. Thus, revenues from the landing fee are expected to cover the full costs of the airfield of AP, that is, the part of the system of runways, taxiways, apron taxilanes, and service roads that is accessible to all aircraft and airport surface vehicles. A separate fee is charged for use of aircraft stands, as well as for any parts of the airfield dedicated exclusively to specific airport users.

	Item	Amount, $
A.	Capital cost of public part of airfield at beginning of FY	354,339,888
B.	Public aircraft facilities—Depreciation	14,173,596
C.	Public aircraft facilities—Interest on	15,308,339
D.	Equipment—Depreciation	457,413
E.	Equipment—Interest on	330,168
F.	Snow-removal services	2,540,000
G.	Maintenance and operations	28,228,906
H.	Administration	16,670,916
I.	Allocated portion of estimated tax liability	3,578,719
J.	Prior year adjustment to projection	(4,545,064)
K.	Annual cost of airfield facilities in FY (= B through J)	76,742,993
L.	Projection of scheduled air carrier landed weight (000 lb)	21,200,000
M.	Landing fee per 1000 lb of landed weight for FY (= K/L)	3.62

TABLE 8.9 Computation of Unit Rate to be Charged for Landing Fees at Airport AP, FY 2000

Item A in Table 8.9 represents the *net value* (or *remaining value*) after depreciation, at the beginning of FY 2010, of the airfield facilities.[21] The capital investments depreciated are only those made by *the airport authority itself*. Any grants or other "free" funding received from any other sources do not enter the computation of the landing fee. In this particular case, AP had received by FY 2010 approximately $124 million in federal grants for airfield capital improvements under the AIP. This amount is *not* included in item A.

Item B is the amount of depreciation on item A taken in FY 2010. AP uses a 25-year, straight-line depreciation schedule. The amount charged for item B is then exactly 4 percent of item A. This schedule of depreciation of airfield facilities, which is standard for U.S. airports, is particularly favorable to airport users and results in low depreciation costs over time. Item C covers the amount of interest, primarily on tax-exempt revenue bonds (see Chap. 7), that AP must pay in FY 2010 on the outstanding portion of funds borrowed in the past to finance capital investments for airfield improvements. Items D and E are entirely analogous to B and C, respectively, and refer to airfield equipment such as fire trucks and runway inspection vehicles. A 10-year, straight-line depreciation schedule is applied. Item F refers to the amount budgeted for the airfield snow-removal contract that AP signs each year.

Items G and H, which together add up to more than 50 percent of the total cost shown on line K for FY 2010, cover the variable costs of managing, operating, and maintaining the airfield. The cost of personnel, including health and other benefits and employee contributions to retirement funds, is a principal component of both of these items. Because the airport authority owns and operates other facilities in addition to Airport AP, care must be taken so that G and H include only (1) administration, operation, and maintenance costs related *directly* to the airfield of AP, and (2) an appropriate allocated portion of the *overhead administrative costs* for the operation of the authority as a whole.

The airport authority that owns and operates AP is exempt from taxes. Instead, each year it makes a contribution, in lieu of taxes, to various local municipalities most affected by the airport's negative externalities. The amount of the contribution is determined by an agreement between AP and the municipalities, which is updated periodically. The contribution represents, in a sense, partial compensation to the communities for the costs associated with being neighbors of the airport. Item I is the portion of the authority's total contribution, which is drawn from airfield revenues. Note that the amount involved ($3,578,719) is equal to only about 1 percent of the relevant net assets (item A). This reflects the prevailing philosophy in the United States: The airfield is considered to be a public-use facility that is made available to airlines and other airport users at the lowest possible cost. Landside commercial facilities, such as concessions in passenger terminals, generally contribute a far more substantial percentage of revenues to the funds distributed in lieu of taxes.

Item J is an adjustment for the over-recovery or under-recovery of the revenue target during the previous fiscal year. The credit in this case resulted from underestimation of the actual air carrier weight (item L) that landed at AP in the previous fiscal year (FY 2009) and the consequent over-recovery of the revenue target in FY 2009.

The bottom three lines of Table 8.9 constitute the three-step average-cost pricing procedure outlined earlier in this section, with items K, L, and M corresponding, respectively, to the quantities X, Y, and Z. Item L is the amount most subject to uncertainty. Any errors in forecasting the total weight of aircraft landing in FY 2010 will be adjusted for through item J in FY 2011. Note, as well, how easy it is to update Table 8.9 from year to year. All one needs to do is update the database on the various cost items and forecast the level of traffic during the next fiscal year.

A committee representing the principal carriers and regional airlines at Airport AP meets on a regular basis with airport authority officials to review data like those shown in Table 8.9. This is standard practice at all U.S. airports. During these meetings, the discussion and any airline criticism typically center on items G, H, and I.

The application of average-cost pricing to other parts of the airport is entirely analogous to the one described in Example 8.3. After the revenue centers and associated cost centers are defined, all that changes from one application to another is the type of costs one is concerned with and the units in which demand is measured. The application can also be easily extended to residual systems (see Chap. 7), which will usually reduce the amounts that airlines will pay for landing fees and other aeronautical charges.

As can be seen from the detailed example, average-cost pricing gives no consideration whatsoever to the possible presence of congestion at an airport. In fact, instead of discouraging potential users from operating at congested airports, average-cost pricing works in exactly the opposite direction. By simply dividing the revenue target, X, by the demand base, Y, to determine the unit cost, Z, average-cost pricing *reduces* the amounts charged for access to a busy airfield—or to any other busy airport facility—as traffic increases, no matter how congested the facility. Access to the airport becomes less expensive, in terms of user charges, as the airport gets more and more congested. As a result, delay-related costs will grow even more rapidly than otherwise. These delay costs can be very substantial and are extremely sensitive to even small changes in traffic volumes, as discussed extensively in Chaps. 11 and 12.

Average-cost pricing is thus regarded as economically inefficient when it comes to congested facilities. *Marginal-cost pricing* is the alternative approach that economists recommend. The guiding principle of marginal-cost pricing is that "efficient use of a facility is achieved when each user pays a charge equal to the additional (marginal) cost that his/her use of the facility causes to others." This marginal cost has a short-term and a long-term component. In the case of the airfield, for example, the short-term marginal cost associated with an aircraft movement is the cost of (1) the delay to other aircraft caused by the movement and (2) the "wear and tear" to the runways, taxiways, and taxilanes (Carlin and Park, 1970). Long-term marginal costs are those associated with the need to expand the existing infrastructure and facilities to increase airfield capacity (Little and McLeod, 1972).

The application of this principle is far from simple, in practice. Long-term marginal costs cannot be estimated until a plan is developed for exactly what facilities will be built to increase capacity. Short-term, delay-related marginal costs are also difficult to compute. Despite such practical difficulties, a number of important airports around the world are adopting modifications to average-cost pricing designed to account for some of these marginal costs. This development and related practices are discussed in more detail in Chap. 12.

8.10 Historical Cost versus Current Cost

An important source of differences in the magnitude of airport user charges stems from differences in the accounting base used to estimate depreciation and amortization costs. Because of general price inflation, as well as of changes in the relative prices of goods and services, the cost in *current (nominal) prices* of replacing airport facilities and equipment at the end of their lifetime is generally much greater than the amount originally paid for them. A passenger building constructed in 1985 at a major airport of a country might have cost $150 million (in 1985 prices in that country). In 2010 it might easily cost $500 million (in 2010 prices in that same country) to construct the same facility, or one of similar size and quality. Using a depreciation schedule based on *historical cost*, that is, on the amount originally paid for a facility or a unit of equipment, may thus lead to under-recovery of the cost of replacing that facility or equipment. *Current-cost* (or *replacement-cost*) *accounting* involves the periodic (usually annual) revaluation of *net assets* (i.e., of the remaining value of assets) according to their replacement costs. Under current-cost accounting, the book value of assets is usually increased by the rate of inflation (as measured by some general or specialized price index) from year to year (see Example 8.4). The book value may also be subjected periodically to more careful revaluation, if it is believed that the replacement cost of the asset is changing at rates substantially different from the general inflation rate.

The use of a current-cost basis may add greatly to the depreciation costs charged at an airport and thus increase user fees significantly. When the BAA switched from historical-cost accounting to current-cost accounting in 1980, the amount charged to users for depreciation jumped from £25 million in FY 1980 to £51 million in FY 1980, while net assets (remaining value of assets) went from £370 million in FY 1980 to £820 million in FY 1981! In the United States, depreciation on the basis of current cost is not an accepted accounting practice and is therefore not used by airports. However, many airports around the world, especially in countries where high rates of inflation are endemic, practice current-cost accounting. Indeed, in many cases current-cost accounting better reflects the realities of the airport environment. Airport operators should consider its use at locations where this approach is permitted.

Example 8.4 Consider an airport asset and assume for simplicity that: (1) its cost at $t = 0$, the beginning of its economic lifetime, is $100; (2) the asset is depreciated over a 10-year period; (3) the replacement cost of this asset increases annually at the rate of 8 percent (which may be different from the general rate of inflation). Table 8.10 compares the depreciation schedule of this asset over its 10-year lifetime on the basis of historical cost (i.e., the cost of the asset at $t = 0$) with the depreciation schedule on the basis of current cost, assuming a straight-line depreciation method when historical cost is used. Note the increasing difference between the amounts charged for depreciation under the two schedules as the asset ages. At the end of the 10 years of the asset's lifetime, users will have paid depreciation costs of $100 under the historical-cost accounting method and $144.9 under the current-cost accounting method (in current dollars in both cases). Note that the current-cost accounting method is far more likely to generate a stream of revenues sufficient to cover the cost of replacing the asset at the end of the 10 years.

	Historical-Cost Accounting		Current-Cost Accounting*	
Age of Asset (years)	Remaining Value of Asset	Depreciation for the Year	Remaining Value of Asset	Depreciation for the Year
0	$100	$0	$100	$0
1	90	10	97.2	10
2	80	10	93.3	10.8
3	70	10	88.2	11.7
4	60	10	81.6	12.6
5	50	10	73.5	13.6
6	40	10	63.5	14.7
7	30	10	51.4	15.9
8	20	10	37	17.1
9	10	10	20	18.5
10	0	10	0	20

*Current-cost depreciation schedule assumes the replacement cost of the asset increases by 8 percent each year.

TABLE 8.10 Historical-Cost and Current-Cost Depreciation Schedules

Exercises

8.1. It is often claimed that providers of ground handling services can achieve considerable economies of scale as the volume of traffic they handle increases. Describe some of the conditions that should apply to achieve such economies.

8.2. Consider the situation in which two or more specialized companies have been authorized to provide ground handling services at a major airport. Describe some of the logistical issues that may arise out of the simultaneous presence of several handlers. Think, for example, of the allocation of aircraft stands, the location of ground handling equipment, the redundancy of equipment, etc. Could the handlers benefit from mutual cooperation and coordination under some circumstances?

8.3. Review the EUROCONTROL formula in Eq. (8.1) for computing en route air navigation (air traffic management) charges. What is the reasoning behind it? In what ways does it reflect or not reflect the true cost of air traffic management services in the en route airspace of a country?

8.4. Consider the computation of the landing fee per unit of aircraft weight at Airport AP in Example 8.3. Airport AP uses a compensatory system of charges (see Chap. 7). If instead it based charges on residual costs, Table 8.9 would have to be modified in a simple way. Indicate what new line item or items would have to be added to Table 8.9 to adjust to a residual system. Would a residual system necessarily reduce the unit charge for the landing fee?

References

ACI, Airports Council International—Europe (2010) *Economics Report 2010*, Brussels, Belgium.

Carlin, A. and Park, R. (1970) "Marginal Cost Pricing of Airport Runway Capacity," *American Economic Review, 60*, pp. 310–318.

Deutsche Bank (1999) *European Airports: Privatization Ahead*, London, U.K.

Doganis, R. (1992) *The Airport Business*, Routledge, London, U.K.

FAA, Federal Aviation Administration (2012) *Airport Financial Reports, Form 5100-127*, FAA, Washington, DC.

IATA, International Air Transport Association (updated monthly) *IATA Airport, ATC, and Fuel Charges Monitor*, Montreal, Canada.

ICAO, International Civil Aviation Organization (1989) *Airport Planning Manual, Part 1: Master Planning*, Doc. 9184-AN/902, ICAO, Montreal, Canada.

ICAO, International Civil Aviation Organization (1992) *Statements by the Council to Contracting States on Charges for Airports and Route Air Navigation Facilities*, ICAO Doc. 9082/4, ICAO, Montreal, Canada.

ICAO, International Civil Aviation Organization (2009) *Policies on Charges for Airports and Air Navigation Services*, ICAO Doc. 9082/6, ICAO, Montreal, Canada.

ICAO, International Civil Aviation Organization (annual) *Manual of Airport and Air Navigation Facility Tariffs*, ICAO Doc. 7100, ICAO, Montreal, Canada.

Little, I. M. D. and McLeod, K. (1972) "New Pricing Policy for British Airports," *Journal of Transport Economics and Policy, 6*(2), May, pp. 101–115.

---

[1]The term *airport operator* (see Chap. 7) refers throughout this chapter to any organizational entity responsible for operating an airport; it may be an airport authority, a branch of a national government, a department of a local government, etc.

[2]Numerous variations of this table are used in practice at different airports.

[3]This is not surprising, as it was hardly possible at the time to imagine the current volumes of airport activity and the economic stakes involved.

[4]The European Union has explicitly banned subsidies of this kind for EU airlines.

[5]In some instances an airport may impose certain limits on the allowable weight of the aircraft—most often because of inadequate runway length for long-range flights by larger aircraft under certain or all weather conditions. In such cases the maximum allowable weight, given the airport limitations, should be used as the basis for computing the landing fee.

[6]Because practically all airports use the average-cost approach to compute landing fees (see Sec. 8.9), the landing fee paid by any given aircraft at any specific airport will be about the same, no matter whether MTOW or MLW is used as the metric of weight.

[7]The shape of this volume as well as the altitude to which it extends depends on local conditions and national practices (see Chap. 13).

[8]This tax, first applied in 1993, was equal to 7.5 percent of the base fare in 2010. In addition, a $3.70 "segment tax" was added for each of the first two segments (or "legs") of a one-way trip.

[9]Aircraft-based charges are also often referred to, confusingly, as "airline-based" charges.

[10]In 2011, passenger-based charges accounted for more than 60 percent of total aeronautical revenues, according to a survey of 376 airports by the Airports Council International.

[11]In a very few instances, an international body may be responsible for ATM in some parts of en route airspace. An example is the EUROCONTROL organization (see Sec. 8.4), which is responsible for en route ATM over a small part of northwestern Europe.

[12]One of the roles of the ICAO is the preparation of Regional Air Navigation Plans covering the entire planet; these plans specify the number and location of air navigation facilities that are necessary to provide adequate "coverage" and services in each region and are approved by the ICAO Council.

[13] The great-circle distance, *d*, is reduced by 20 km for every landing and takeoff within the Member State by the aircraft involved.

[14] Thirty-five different countries, including all the Member States of the European Union, participated in EUROCONTROL's charging system as of 2011. For the unit air navigation charge for each State, see ICAO's Document 7100 (ICAO, annual).

[15]An earlier version encourages airport operators to "develop these revenues to the maximum extent possible" (ICAO, 1992).

[16]In a fascinating statistical nugget, the BAA found that in 1995 the average Japanese passenger spent approximately $80 at BAA airports, Scandinavians $40, U.K. citizens $16, and U.S. citizens $12.

[17]User charges at the overwhelming majority of airports do not take delays into consideration. For this reason, the increasing marginal costs due to delays are only rarely accounted for (see Chap. 12 for an extensive discussion).

[18]Charges on international traffic are much higher because of various federal taxes and fees for customs, immigration services, etc.

[19]In 2009, roughly 15 percent of European airport operators were also providers of ground-handling services at their own airports (ACI—Europe, 2011).

[20]Note that this requires estimation of the revenues from other revenue centers at the airport (e.g., from terminal buildings) that will be subtracted from the full cost of the revenue center in question (e.g., the airfield).

[21]This is the net value of the *historical cost* of the airfield; see Sec. 8.10.

PART III

The Airside

CHAPTER 9

Airfield Design

The geometric design of an airfield should provide for operational efficiency, flexibility, and potential for future growth. It should also comply with an extensive set of design standards and recommended practices, developed over the years by international and national civil aviation organizations and intended to promote a maximum level of safety.

The two most influential sets of design standards are those of the International Civil Aviation Organization (ICAO) and the U.S. Federal Aviation Administration (FAA). They are based on similar, but not identical, coding systems that classify airfields according to the most demanding type of aircraft they are designed to serve. Once the reference code of an airfield or runway has been specified, design standards can be obtained from the relevant manuals and other supporting documents.

Airfields typically account for 80 to 95 percent of the total land area occupied by an airport and affect in critical ways every facet of airport operations. The principal determinants of the size of the airfield include the number and orientation of the runways; the geometric configuration of the runway system; the dimensional standards to which the airfield has been designed; and the land area set aside to provide for future growth and/or environmental mitigation. This chapter discusses these topics in varying levels of detail.

The chapter reviews the characteristics and some advantages and disadvantages of a broad set of common airport layouts. These range from single runways, to a pair of parallel runways, to intersecting pairs of runways, to systems of three, four, or more runways. Several airports in the United States use complex, multirunway layouts to serve very large volumes of nonhomogeneous traffic.

Four common mistakes in planning and designing airfields are the following:

- Failure to provide flexibility for future expansion
- Overbuilding the airfield in its initial phases
- Lack of integration and coordination of the planning process
- Insufficient appreciation of the economic consequences of some design choices

The implications of these mistakes for the capital and operating costs of airports and their users can be serious.

The chapter also reviews many of the FAA and ICAO technical and dimensional standards for the various elements of the airfield. These include coverage for crosswinds, runway length, other runway geometric standards and obstacle clearance requirements, separation of runways from adjacent facilities and from static or moving objects, taxiway geometric standards and separation requirements, apron layouts and separation requirements, and obstacle limitation surfaces (or "imaginary surfaces") in the airspace in the vicinity of airports. The objective is to provide a summary overview of the standards and recommended practices, outline their rationale, and indicate where readers can find further information in the detailed relevant documents.

9.1 Introduction

The geometric design of an airfield critically affects every aspect of airport operations. This includes landside facilities and services, as the layout of the runway system largely dictates the general placement of the passenger, cargo, and other buildings, as well as the interfacing of airside and landside operations.

Because of the overwhelming importance of safety for aviation operations, airfield design must comply with a voluminous set of detailed standards and recommended practices developed over the years by national and international civil aviation authorities and organizations. The ICAO plays a central role in this respect. Its Annex 14 to the International Convention on Civil Aviation specifies the *standards* and *recommended practices* that have been adopted by its nearly 200 Member States (formally "Contracting States") over the years. An ICAO standard is

> any specification for physical characteristics, configuration, material, performance, personnel or procedure, the uniform application of which is recognized as necessary for the safety or regularity of international air navigation and to which Contracting States will conform (ICAO, 2009).

By comparison, a recommended practice is

> any specification … the uniform application of which is recognized as desirable … and to which Contracting States will endeavor to conform.

Note that a standard is "necessary" for safety, while a recommended practice is "desirable" ("endeavor to conform"). In practice, the design standards and practices of national civil aviation agencies everywhere are largely based on or, in most cases, identical to those specified in the ICAO's Annex 14. Member States that, for any reason, are unable to comply with an Annex 14 standard—and adopt a different one—must notify promptly the ICAO to this effect. This is referred to as "filing a difference." The ICAO publishes these differences for the information of all its Member States.

The United States has historically filed the largest number of differences. The FAA has developed and applies a set of airport design standards and recommended practices, which are similar to those of the ICAO, but also differ in some important respects. This chapter cites examples later on. For practical purposes, the FAA plays almost as important a role as the ICAO in setting airport design specifications. One reason is that the United States is still the most important air travel market in the world, with the largest volume of air traffic and with most of the busiest airports (see Chaps. 1 and 2 and Sec. 9.2). Second, the FAA and the U.S. government have traditionally invested heavily in research on aviation, including airports and air traffic management (ATM). As a result, the design standards that the FAA has adopted or updated have often preceded the ICAO adoption of identical or very similar standards.

Airport professionals should therefore be cognizant of both the ICAO's and the FAA's sets of design standards and recommendations. Annex 14 (ICAO, 2009) was first published in 1951 and has since been amended many times, usually following reports and studies by committees and panels of experts. A large number of related ICAO publications amplify on aspects of the Annex 14 and provide more detailed guidance.[1] Important examples are three multivolume manuals, the *Aerodrome Design Manual* (ICAO, current-a), the *Airport Services Manual* (ICAO, current-c), and the *Airport Planning Manual* (ICAO, current-b), all of which are updated at irregular intervals. Of special relevance to this chapter are the *Aerodrome Design Manual—Runways* (ICAO, 2006), the *Aerodrome Design Manual—Taxiways, Aprons, and Holding Bays* (ICAO, 2005), and the *Airport Planning Manual—Master Planning* (ICAO, 1987). The principal document on the FAA side is the *Airport Design Advisory Circular* (FAA, 2012). This also references numerous other related advisory circulars and federal aviation regulations (FAR), some of which are cited later in this chapter.[2] A few commercial vendors increasingly provide specialized computer-aided-design (CAD) software to support the planning and design of airfields.

Despite such support and extensive sets of guidelines, airport planners must still exercise a great deal of judgment in making critical design choices. Subject to the environmental, political, and economic constraints at each site, they must address such fundamental issues as the following:

- How much land should be acquired or reserved for a new airport?
- What should be the overall geometric layout of runways, taxiways, and aprons?
- What size of aircraft should the airfield be designed for?

- How should the construction of airside facilities be phased?

Variants of these questions must also be addressed when modifying or expanding the airfields of existing airports. Modification and expansion projects have, in fact, become the most common, by far, context for airport planning and design in view of the small number of entirely new major airports currently being built or planned anywhere in the world. Expansion and modification projects are often as complex as the design of new airports—and sometimes more so. One of the biggest challenges in this respect is developing a schedule of construction activities and a transition plan that will allow the airport to continue operations during the period of airport reconfiguration.

Four generic types of mistakes are common in planning and designing airfields:

- Failure to provide flexibility for responding to future developments
- Overbuilding for the initial stages of an airport's operation
- Adopting a hierarchical, nonintegrated approach to design that does not consider adequately the interactions among the various elements of the airport
- Insufficient appreciation of the economic implications of design choices

The first of these applies to long-range planning. Too many airports face today severe, sometimes insuperable, constraints because their original designers and planners failed to anticipate the eventual land area requirements of the airfield (see Sec. 9.4). Another example of failure to plan for flexibility is the selection and construction of airside layouts (runways, taxiways, and aprons) that make it impossible to accommodate new, larger types of aircraft without either making very expensive changes to existing facilities or having to rebuild them from scratch (see Secs. 9.5–9.8).

The second type of mistake, that is, the tendency to overbuild the airfield in the initial stages of airport operations, is in some ways the reverse of the first. For example, an airport in its early phases of development may not need the full system of runways and taxiways it has been planned for, or the eventual full length of one or more of its runways. Only part of the planned system may be sufficient for the initial phase of operations. This may mean, for instance, building only one full-length taxiway running parallel to a main runway, instead of the two parallel taxiways that may be necessary when the airport reaches full development.

Failure to adopt an integrated approach to planning for the various parts of the airfield is a third weakness encountered in practice. Airports tend to be planned and designed hierarchically, often without fully considering the interactions among their various "subsystems" (runways, taxiways, aprons, passenger and cargo buildings, service areas, etc.). On the airside for example, the highest level of the design process typically focuses on settling the configuration of the runway system, with limited analysis of what this implies for the other components of the airfield. Similarly, on the landside, passenger buildings are often designed with inadequate understanding of how they interface with the apron areas, taxilanes, and taxiways. Because of such absence of a "systems" viewpoint, taxiway and apron systems, in particular, are often inefficient and sometimes include parts that are obvious candidates to become congestion points ("hot spots") for air traffic (see Secs. 16.2 and 16.3). Airfield design also needs to consider safety-related criteria such as minimizing the number of runway crossings and reducing the likelihood of runway incursions. This can be achieved only through an integrated approach to planning and design.

Finally, the economic implications of some design choices are often not fully appreciated and analyzed. Sometimes planners make design choices that save some capital costs but greatly increase the operating costs of airport users—for example, by increasing taxiing times on the airfield. This is because planners often do not have a good grasp of the cumulative economic value of their design choices. For example, saving an average of 2 minutes of taxiing time per landing and/or takeoff may be worth tens of millions of dollars per year to aircraft operators at a busy airport with hundreds of thousands of annual operations (see Secs. 9.7 and 14.3).

This chapter both reviews the most important airfield design standards and recommended practices and provides a perspective on how these are, or should be, applied. The ICAO and the FAA use simple classification schemes to develop two-element reference codes for airports. These reference codes determine the design standards to be used at each airport. Section 9.2 explains the airport reference codes (ARCs), discusses their application, and provides relevant background and terminology for the chapter. Section 9.3 reviews "wind coverage" requirements that determine whether it is necessary to construct runways in more than one orientation at an airport. Section 9.4 offers a brief tour through progressively more complex airfield configurations, using a few important airports as examples. It shows how the requirements for separations between runways largely dictate the overall layout of the airport. It also points to some important systemic differences between airport traffic characteristics in different regions of the world. Section 9.5 provides an overview of the topic of runway length. The emphasis is on explaining the fundamental concepts and the meaning of various technical terms without going into much technical detail. Section 9.6 summarizes some of the most important design standards for runways, as they apply to major airports. Sections 9.7 and 9.8 do the same for taxiways, elements of taxiway systems, such as high-speed exits and taxilanes, and apron stands. In all cases, the principal concern centers on the practical implications for busy airports serving large commercial airplanes. Section 9.9 goes beyond airport boundaries to compare the standards that the ICAO and the FAA have developed for protecting the airspace in the immediate vicinity of airports from natural or man-made obstructions that may pause a threat to the safety of runway operations. The section describes the various obstacle limitation surfaces (or "imaginary surfaces") that form the basis for these standards.

It should be noted that the review of design standards in Secs. 9.6 through 9.9 is far from exhaustive, as it omits several topics altogether—such as the design of taxiway turns and fillets, visual aids and marks, and emergency and rescue services—and leaves out numerous details on others. Such coverage is beyond the scope of this text. Those engaged in the detailed design of airfield facilities are familiar with the voluminous materials referenced earlier and other related documents. These professionals also consult regularly with the competent government organizations and regulators, are typically employed by engineering consulting firms, and work with special-purpose airfield design software.

9.2 Airport Classification Codes and Design Standards

Reference Codes for Aircraft Classification

Both the ICAO and the FAA use simple classification schemes to develop a two-element *reference code* for each type of aircraft. The ICAO reference code (ICAO, 2009) consists of two "elements," a code number and a code letter (Table 9.1). The *code number* of any type of aircraft is determined by the *airplane reference field length*, the minimum field length that aircraft requires for takeoff at maximum takeoff weight (MTOW), sea level, standard atmospheric conditions,[3] no wind, and level runway (see Sec. 9.5). The *code letter* is determined by two physical characteristics of the aircraft: its *wingspan* and the *distance* ("*span*") *between the outside edges of the wheels of the aircraft's main gear*. When the aircraft's wingspan and outer main gear wheel span (OMG) correspond to different code letters, the aircraft is assigned the more demanding code letter. For example, the Boeing 747-800 has an airplane reference field length of approximately 3000 m, which gives it code number 4, a wingspan of 68.5 m (code letter F), and OMG of 12.7 m (code letter E). Thus, the ICAO reference code for the Boeing 747-800 is 4-F.

ICAO Code Element 1		ICAO Code Element 2		
Code Number	Airplane Reference Field Length (RFL)	Code Letter	Wingspan (WS)	Outer Main Gear Wheel Span (OMG)
1	RFL < 800 m	A	WS < 15 m	OMG < 4.5 m
2	800 m ≤ RFL < 1200 m	B	15 m ≤ WS < 24 m	4.5 m ≤ OMG < 6 m
3	1200 m ≤ RFL < 1800 m	C	24 m ≤ WS < 36 m	6 m ≤ OMG < 9 m
4	1800 m ≤ RFL	D	36 m ≤ WS < 52 m	9 m ≤ OMG < 14 m
		E	52 m ≤ WS < 65 m	9 m ≤ OMG < 14 m
		F	65 m ≤ WS < 80 m	14 m ≤ OMG < 16 m

Note: The OMG limits for code letters D and E are identical.

Source: ICAO, 2009.

TABLE 9.1 ICAO Airport Reference Code

Analogously, the FAA uses aircraft *approach speed* to determine the first element of its reference code, the aircraft *approach category*, designated by a letter between A and E (Table 9.2). The aircraft approach speed is defined as 1.3 times the stall speed in the aircraft's landing configuration at maximum landing weight (MLW). The second element is a Roman numeral (I through VI) that specifies the *design group* to which the aircraft belongs. The design group is determined by the most demanding of two of the physical characteristics of the aircraft: its *wingspan* and its *tail height*. For example, in the case of the Boeing 737-800, the aircraft's approach speed is 142 knots (FAA approach category D); its wingspan of 34.3 m and tail height of 12.6 m, both place it in design group III. Thus, the FAA reference code for the Boeing 737-800 is D-III.

FAA Code Element 1		FAA Code Element 2		
Aircraft Approach Category	Aircraft Approach Speed (AS) (knots)	Airplane Design Group	Aircraft Wingspan (WS)	Tail Height (TH) Wheel Span (OMG)
A	AS < 91	I	WS < 49 ft (WS < 15 m)	TH < 20 ft (TH < 6 m)
B	91 ≤ AS < 121	II	49 ft ≤ WS < 79 ft (15 m ≤ WS < 24 m)	20 ft ≤ TH < 30 ft (6 m ≤ TH < 9 m)
C	121 ≤ AS < 141	III	79 ft ≤ WS < 118 ft (24 m ≤ WS < 36 m)	30 ft ≤ OMG < 45 ft (9 m ≤ OMG < 13.5 m)
D	141 ≤ AS < 166	IV	118 ft ≤ WS < 171 ft (36 m ≤ WS < 52 m)	45 ft ≤ OMG < 60 ft (13.5 m ≤ OMG < 18.5 m)
E	166 ≤ AS	V	171 ft ≤ WS < 214 ft (52 m ≤ WS < 65 m)	60 ft ≤ OMG < 66 ft (18.5 m ≤ OMG < 20 m)
		VI	214 ft ≤ WS < 262 ft (65 m ≤ WS < 80 m)	66 ft ≤ OMG < 80 ft (20 m ≤ OMG < 24.5 m)

Source: FAA, 2012.

TABLE 9.2 FAA Airport Reference Code

Airport Reference Code

The *airport reference code* (ARC) of an airfield is determined by the code of the *most demanding* type of aircraft ("critical aeroplane" in ICAO terminology) that the airport is designed to serve. For instance, if the most demanding aircraft for some airport is the Airbus 340-600, classified as a 4-E aircraft in the ICAO's scheme and as a D-V in the FAA's, the airport's ARC would be 4-E or D-V according to the ICAO's or the FAA's reference codes, respectively.

It is important to note that an airport's "most demanding aircraft"—the type that determines an airport's ARC—need *not* be an aircraft that is currently using the airport. In other words, an airport can be designed to accommodate in the future aircraft types that are more demanding than the ones that have been served there in the past.

At multirunway airports, individual runways may differ in their ability to serve different types of aircraft. For example, one runway may be too short for handling takeoffs by long-range aircraft, whereas another may be sufficiently long for this purpose. In such cases, different *runway design codes* (RDCs) will be associated with different runways. For example, the long runway may have an FAA RDC of D-V, and the shorter runway an RDC of C-III. In such situations, the runway with the "highest" RDC determines the overall ARC. Thus, in our example the ARC will be D-V—the most demanding of D-V and C-III.

Practical Implications

When it comes to the first element of the ICAO reference code, note that the most common narrow-body commercial aircraft, such as the Airbus 320 and the Boeing 737, has a reference field length greater than 1800 m (Table 9.3*b*). This means that the ARC of virtually all major commercial airports has an ICAO code number of 4. At the same time, for all practical purposes the wingspan of the most demanding aircraft determines the second element of the ICAO reference code. This is because, for the existing types of important commercial jets, the OMG almost never places these aircraft in a code letter category higher than the one to which they would be assigned based on their wingspan. For example, no airplane assigned code letter D on the basis of its wingspan would be assigned code letter E or F on the basis of its OMG. It follows from these two observations that the ICAO reference code for major airports can only be 4-C (in the rather unusual case where aircraft like the Airbus 320 or the Boeing 737 are the most demanding that the airport can serve) or, far more often, 4-D, 4-E, or 4-F.

Turning to the FAA ARC, it is again true that the wingspan of the most demanding aircraft determines the airplane design group for all practical purposes (see Table 9.3). This is because the tail height never places an aircraft in a design category higher than the one to which it would have been assigned based solely on its wingspan. Note also that the wingspan thresholds that separate FAA airplane Groups I through VI from one another are *exactly* the same as the ICAO thresholds. For example, FAA Group IV aircraft have wingspans between 36 and 52 m, exactly the same as ICAO code letter D aircraft. This means that the second elements of the FAA and the ICAO reference codes for all types of aircraft correspond perfectly. The only difference is that the FAA uses Roman numerals and the ICAO uses capital letters to designate that second element. An airport with an ARC in Group V per the FAA will have code letter E per the ICAO and vice versa—certainly a desirable circumstance for airfield designers. This is pointed out in Fig. 9.1, which plots the length and wingspan of many of the most common types of current commercial jet airplanes and identifies on the right vertical axis the second code element to which they belong.

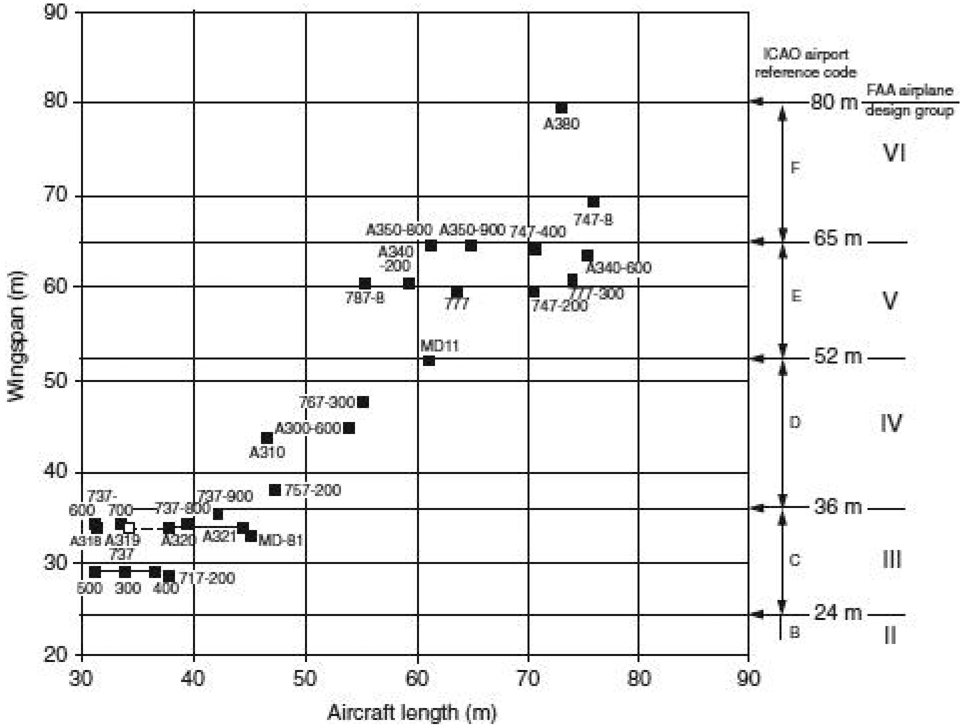


FIGURE 9.1 Length and wingspan of current types of commercial jet transport airplanes. The corresponding ICAO and FAA reference codes are indicated along the vertical axis on the right. [*Sources*: Manufacturer specifications, FAA (2012), ICAO (2006).]

The second code element largely determines many of the geometric design standards at airports, such as the required separation between a runway and a parallel taxiway or between two parallel taxiways. This is because wingspan reflects well the physical size of aircraft, especially when it comes to selecting airfield dimensions that will ensure safe operations. It follows that it makes little difference, in most instances, whether an airport is designed to FAA or ICAO standards because (1) these agencies will classify any airport in the same way, on the basis of wingspan, and (2) their dimensional standards for different wingspan categories are usually either identical or very similar. There are, however, a few significant exceptions to this statement (see Secs. 9.6–9.9).

The choice of ARC is obviously a critical decision for airport planners and operators. Building for a more demanding aircraft than necessary means incurring unnecessary capital and maintenance costs: the dimensions of runways, taxiways, and aprons and the separations between them will be larger than necessary. On the other hand, it may be even costlier to “under-design” the airport. If, at some future time, an airline wishes to initiate service with a type of aircraft that the airport is not designed to handle, this service must either be denied, or arrangements must be made to accept the aircraft under some special handling provisions, or the airport's facilities must be modified to make them compatible with the aircraft. The first two choices are unattractive in the long run, especially if the popularity of the aircraft in question increases over time. The third choice can be very expensive and disruptive if adequate provisions were not made at the outset for the possibility of future redimensioning of airfield facilities, as Examples 9.1 and 9.2 suggest.

Example 9.1 The Airbus 380, with a wingspan of 79.8 m, was the first nonmilitary airplane with an ICAO reference letter F (or FAA Group VI) designation to be widely introduced at commercial airports. Before its entry into service, Airbus Industrie surveyed 81 leading airports around the world, considered the top candidates to receive A380 service, in order to identify potential airport/aircraft compatibility problems. The survey found that the three principal problems were runway and taxiway dimensions and separations, weight effect on taxiway bridges, and the impacts of aircraft size and capacity on passenger buildings. An Airports Council International survey of 30 of these airports found an average cost of about $100 million per airport for the adjustments needed to accept the A380 at these airports, or a total of $3 billion for just these 30 airports (Airbus Industrie, 2001).

Example 9.2 The FAA, anticipating the future development of aircraft larger than the Boeing 747, published as early as 1983 design standards for aircraft in Group VI with a wingspan of up to 80 m (FAA, 1983). The ICAO did not officially publish its corresponding standards (code letter F) until 1999. Prudent planners at airports that might be used in the future by new airplanes in Group VI or Code Letter F, such as the Airbus 380, consulted the FAA standards between 1983 and 1999 in designing new airports or planning for the expansion of existing ones.

The message is clear: *design for flexibility and build in stages*. Whenever the land area available makes this feasible, the geometric design should make possible the future adjustment of the geometric characteristics of the airfield so it can serve larger aircraft than it was originally built for. To follow such a strategy, airport designers and planners should be aware of the full foreseeable range of potential aircraft sizes and associated design standards.

Table 9.3 lists some of the characteristics of many common types of commercial jet airplanes and a few turboprops. It is subdivided into parts for wide-body aircraft (see Table 9.3*a*), narrow-body aircraft (see Table 9.3*b*), and regional jets and turboprops (see Table 9.3*c*). The information provided includes the following:

	WS (m)	Length (m)	TH (m)	OMG (m)	MTOW (tons)	Passenger Seats	Range (km)	TO Field Length (m)	ICAO RC	FAA RC	FAA TDG
Airbus											
A300-600R	44.9	54.1	16.6	10.9	171.7	220–266	7,540	2320	4-D	C-IV	5
A310-300	43.9	46.7	15.8	10.9	164.0	218–240	9,600	2260	4-D	C-IV	5
A330-200	60.3	58.8	17.4	12.6	233.0	253–293	13,430	2220	4-E	C-V	6
A340-300	60.3	63.6	16.9	12.6	276.5	295–375	13,700	3000	4-E	D-V	6
A340-600	63.5	75.3	17.3	12.6	368.0	380–440	14,350	3100	4-E	D-V	6
A350-900	64.8	66.9	17.1	12.9	268.0	315–366	15,000	n.a.	4-E	D-V	6
A380	79.8	72.7	24.5	14.3	560.0	525–644	15,400	2900	4-F	D-VI	7
Boeing											
747-200B	59.6	70.6	19.3	12.4	377.8	366–452	12,700	3190	4-E	D-V	6
747-400	64.4	70.6	19.4	12.6	396.9	416–524	13,450	3200	4-E	D-V	6
747-8	68.5	76.3	19.4	12.7	448.0	467–605	14,800	3000	4-F	D-VI	6
767-200ER	47.6	48.5	16.1	10.9	179.2	181–224	11,825	2740	4-D	D-IV	5
767-300ER	47.6	54.9	16.0	10.9	186.9	218–269	11,065	2540	4-D	D-IV	5
777-200ER	60.9	63.7	18.5	12.9	297.6	301–400	14,300	3570	4-E	C-V	6
777-300ER	64.8	73.1	18.5	12.9	351.5	365–451	14,690	3200	4-E	D-V	6
787-8	60.2	56.7	16.9	11.7	228.0	242–264	15,200	2850	4-E	D-V	6
787-9	60.2	62.8	17.0	11.7	251.0	250–290	15,700	n.a.	4-E	D-V	6

Sources: Manufacturers' data, FAA (2012), ICAO (2006).

TABLE 9.3*a* Characteristics of Common Wide-Body Turbofans

	WS (m)	Length (m)	TH (m)	OMG (m)	MTOW (tons)	Passenger Seats	Range (km)	TO Field Length (m)	ICAO RC	FAA RC	FAA TDG
Airbus											
A318	34.1	31.4	12.6	8.7	68.0	107–117	5700	1828	4-C	C-III	3
A319-100	34.1	33.8	11.8	8.7	75.5	124–134	6700	2164	4-C	C-III	3
A320-200	34.1	37.6	12.1	8.7	77.0	150–164	5900	2090	4-C	C-III	3
A321-200	35.8	44.5	11.8	9.1	93.0	185–199	5950	2560	4-C	C-III	5
Boeing											
717-200	28.5	37.8	8.9	5.4	54.9	106–117	3800	1800	4-C	C-III	3
737-200	28.4	30.5	11.2	6.4	58.1	102–136	4300	2090	4-C	C-III	5
737-300	28.9	33.4	11.1	6.4	62.8	128–140	4200	2300	4-C	C-III	3
737-400	28.9	36.5	11.1	6.4	68.0	146–159	4200	2540	4-C	C-III	3
737-500	28.9	31	11.1	6.4	60.6	108–122	4450	2470	4-C	C-III	3
737-600	34.3	31.2	12.6	7.0	66.0	108–123	5650	1750	4-C	C-III	3
737-700	34.3	33.6	12.5	7.0	70.1	128–140	6230	2100	4-C	C-III	3
737-800	34.3	39.5	12.5	7.0	79.0	160–175	5670	2400	4-C	D-III	3
737-900	35.7	42.1	12.5	7.0	85.1	174–204	5000	3000	4-C	D-III	3
757-200	38.1	47.3	13.6	8.6	115.7	200–239	7200	2910	4-D	C-IV	5
MD-81*	32.8	45.0	9.0	6.2	63.5	155–172	2910	1870	4-C	C-III	3
MD-87*	32.8	39.7	9.3	6.2	63.5	130–139	4400	1860	4-C	C-III	3

*Boeing/Douglas airplane.

Sources: Manufacturers' data, FAA (2012), ICAO (2006).

TABLE 9.3*b* Characteristics of Common Narrow-Body Turbofans

	WS (m)	Length (m)	TH (m)	OMG (m)	MTOW (tons)	Passenger Seats	Range (km)	TO Field Length (m)	ICAO RC	FAA RC	FAA TDG
Bombardier											
CRJ 200ER	21.2	26.8	6.2	4.0	23.1	50	3045	1770	3-B	C-II	3
CRJ 700ER	23.2	32.5	7.6	5.7	34.0	70	3208	1676	3-B	C-II	3
CRJ 900ER	24.9	36.4	7.5	6.8	37.4	86	2950	1861	4-C	C-III	3
CRJ 1000ER	26.2	39.1	7.5	7.0	41.6	100	2843	2079	4-C	C-III	3
Dash8 (Series Q400)*	28.4	32.8	8.3	6.1	29.3	78	2522	1402	3-C	B-III	3
Embraer											
ERJ 135ER	20.0	26.3	6.8	4.1	19.0	37	2300	1610	3-B	C-II	3
ERJ 145ER	20.0	28.5	6.8	4.1	22.0	50	1900	1970	4-B	C-II	3
ERJ 170-200 (ERJ 175)	26.0	31.7	9.7	6.2	38.8	78–88	3889	2244	4-C	C-III	3
ERJ 190-100	28.7	36.2	10.3	6.9	50.3	94–114	4260	2056	4-C	C-III	3
ATR											
ATR 42-500*	24.6	22.7	7.6	4.9	18.6	42–50	1555	1100	2-B	B-III	3
ATR 72-500*	27.1	27.2	7.7	7.3	22.5	68–74	1324	1165	2-C	B-III	3

Sources: Manufacturers' data, FAA (2012), ICAO (2006).

TABLE 9.3*c* Characteristics of Common "Regional Jets" (Narrow-Body Turbofans) and Turboprops (indicated with*)

- Wingspan (WS), length, tail height (TH), and outer main gear wheel span (OMG)
- Maximum certificated structural takeoff weight (MTOW)
- Number of passenger seats
- Maximum range of the aircraft at full payload
- Takeoff field length at MTOW, sea level, standard atmospheric conditions, no wind, and level runway
- ICAO and FAA reference codes (RC)
- FAA taxiway design group (TDG), a recent grouping of aircraft types for taxiway design purposes—see Sec. 9.7

The MTOW, range, and takeoff field length shown are only indicative and may vary for the same aircraft model depending on many factors, such as engine type, or modifications and improvements that may become available over time. The number of passenger seats shown indicates a typical range for standard three-class (first, business, and economy) seating or two- or single-class configurations. The number of seats may be increased in some cases beyond the ranges shown in Table 9.3, depending on intended use of the aircraft (e.g., short hops), the seat pitch desired by the airline, and the class configuration of the airplane cabin.

Note that the ranges and takeoff field lengths in Table 9.3 apply under a set of nominal conditions (e.g., takeoff field length at MTOW from a level runway at sea level under standard atmospheric conditions and no wind). They may vary greatly on a day-to-day basis and the specific circumstances involved, such as the selected tradeoff between payload and the amount of fuel carried, the weather conditions (temperature and winds at takeoff airport, winds aloft), and the physical characteristics of the takeoff runway (see Sec. 9.5).

Runway Designation and Classification

Every runway is identified by a two-digit number, which indicates its magnetic azimuth[4] in the direction of operations to the nearest 10° (see the exception for four or more parallel runways below). For example, a runway with a magnetic azimuth of 224° is designated and marked as "Runway 22" (for 220°). Obviously, the identification numbers at the two ends of any runway will differ by 18. For instance, the opposite end of Runway 22 is designated as Runway 04, and the runway is referred to as "Runway 04/22." In the case of two parallel runways, the letters R, for right, and L, for left, are added to distinguish between the runways. Thus, Boston/ Logan has a pair of close parallel runways designated as 04R and 04L when the runways are operated to the northeast and, respectively, as 22L and 22R when operated to the southwest. The letters R, C (for center), and L are often used with three parallel runways. If an airport has four parallel runways, one pair is typically marked to the nearest 10°, with the additional indications R and L, and the other to the next nearest 10°, with the additional indications R and L. For example, Los Angeles/International has four parallel runways arranged in two close pairs, the pair 6L/24R and 6R/24L to the north of the terminal complex and the pair 07L/25R and 07R/25L to the south. This also indicates that, with operations to the east, the magnetic azimuth of the four Los Angeles runways is between 60° and 70°.

There may be deviations from these practices, if this is deemed advisable to avoid confusion and ensure safety. For instance, Seoul/Incheon opened with two close parallel runways, designated 15R/33L and 15L/33R. When a third parallel runway was later added to the west and at some distance from the two existing ones, it was designated 16/34, in preference to changing the designation of the existing runway 15R/33L to 15C/33C and calling the new runway 15R/33L. Similarly, when Atlanta opened a fifth parallel runway (see Fig. 9.7) in 2006, it was called 10/28, to avoid renaming any of the already existing runways 8L/26R and 8R/26L, to the north of the terminal complex, and 9L/27R and 9R/27L to the south. In contrast, at Dallas/Fort Worth (Fig. 9.8), the five parallel runways in the principal direction of operations are designated as 17L/35R, 17C/35C, and 17R/35L (three runways on one side of the terminal complex) and 18L/36R and 18R/36L (a close pair on the other side).[5]

For the purpose of specifying design standards, runways are also classified as noninstrument and instrument. A *noninstrument* (or *visual*) *runway* is intended for the operation of aircraft using visual approach procedures. This chapter is concerned only with design specifications for *instrument runways*, that is, runways that permit the operation of aircraft using instrument approach procedures. Instrument runways are further subdivided into nonprecision and precision. A *nonprecision approach runway* has some visual aids and, at a minimum, a navigation aid that provides directional guidance adequate for a straight-in approach, but no vertical guidance (Chap. 13). A *precision approach runway* allows operations with a decision height and visibility corresponding to at least Category I limits (see Chap. 13).

9.3 Wind Coverage

The construction of runways in more than one direction is usually motivated by the requirement to provide adequate coverage for crosswinds. Landings and takeoffs are typically conducted *into* the wind.[6] For instance, when the wind is from the north, runways with a northerly direction, if available, will be preferred over others, and landings and takeoffs will be performed in a generally south-to-north direction. When any given runway is in use, the *crosswind* is the component of the surface wind velocity vector perpendicular to the runway centerline. Its magnitude can be computed by simply multiplying the speed of the prevailing wind by the sine of the angle between the wind direction and the runway centerline.

The ICAO specifies that a runway should not be used if the crosswind component exceeds (ICAO, 2009) the following:

- 19 km/h (10.5 knots) for airplanes whose reference field length is less than 1200 m
- 24 km/h (13 knots) for airplanes whose reference field length is between 1200 and 1499 m
- 37 km/h (20 knots) for airplanes whose reference field length is 1500 m or greater, except that with poor braking action (e.g., when the runway surface is wet) the limit is 24 km/h (13 knots)

The corresponding FAA requirements (FAA, 2012) call for crosswinds not exceeding the following:

- 10.5 knots (19 km/h) for ARCs A-I and B-I
- 13 knots (24 km/h) for A-II and B-II
- 16 knots (30 km/h) for A-III, B-III, C-I through C-III, and D-I through D-III
- 20 knots (37 km/h) for all other ARCs

Naturally, these limits are conservative. For example, aircraft in FAA Groups IV, V, or VI (or with ICAO code letter D, E, or F) can maneuver with crosswinds as high as 25 to 30 knots (46–55 km/h). The actual selection of the runway(s) to be used at any given time at an airport is made by the provider of ATM services (the FAA in the United States), *not* the airport operator, taking into consideration prevailing winds (see Chap. 10). A pilot may request reassignment to a different runway on account of crosswinds.

Both the ICAO and the FAA recommend that the number and orientation of runways should be such that *crosswind coverage* (or the *airport usability factor* in ICAO terminology) is at least 95 percent. In other words, the percentage of time during which the use of a runway system is restricted because of crosswinds should be less than 5 percent. Note that the 95 percent target may still leave approximately 18 days per year without crosswind coverage. For many major airports this may not be acceptable. In practice, these airports usually provide runways in a sufficient number of directions, when needed, to ensure usability factors higher than 95 percent. National civil aviation authorities, in fact, may impose far more stringent crosswind coverage requirements than 95 percent at the principal airports of their countries.

Airport designers use historical wind statistics from an airport's site to determine the orientation of the runways that should be provided to achieve adequate crosswind coverage. A *wind rose* conveniently summarizes these statistics graphically. It consists of a series of concentric circles, representing wind speed groupings from 0 to 10 knots, 10 to 16 knots, 16 to 21 knots, etc., and a set of radial lines, usually drawn at intervals of 10°, that cut through the circles (Fig. 9.2). The figure within each resulting "box" indicates the percentage of time during which winds observed at the site are within the corresponding orientation limits and speed limits.

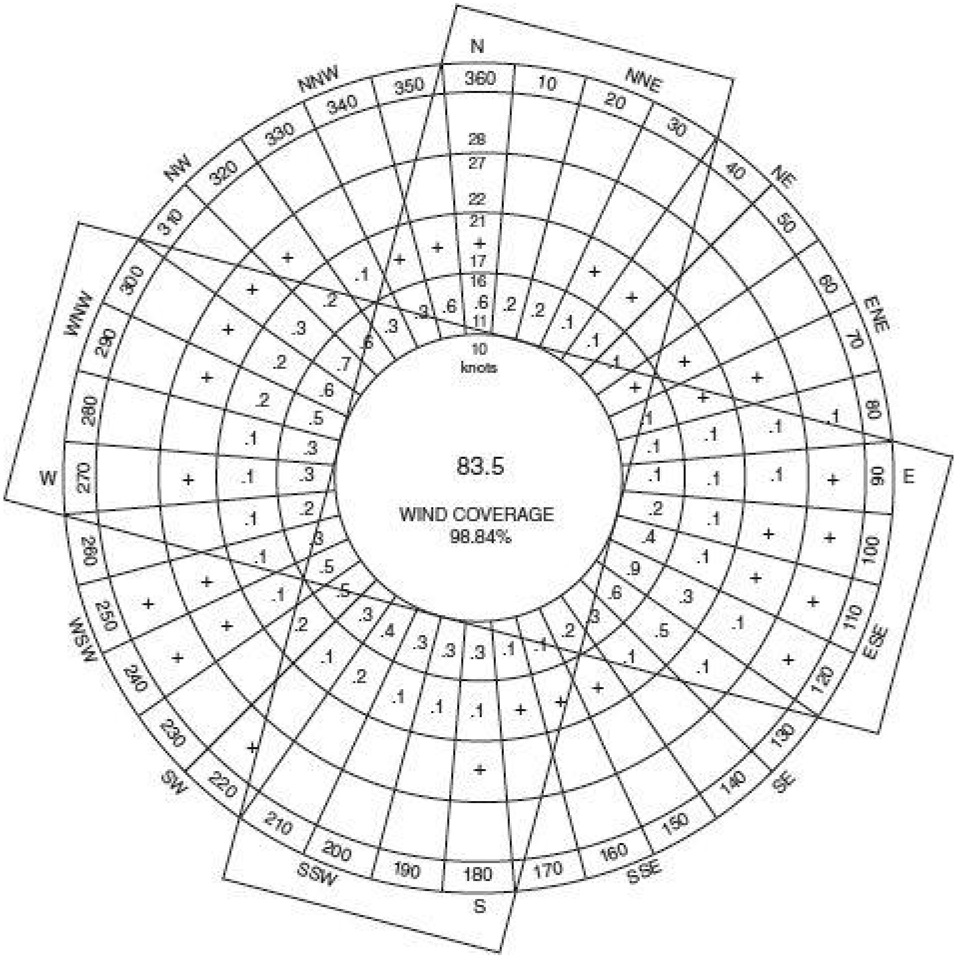


FIGURE 9.2 A wind rose and the wind coverage of two runways. The numbers indicate percent of time and a "+" indicates "nonzero, but less than 0.1 percent." The total coverage achieved by the two runways in this example is 98.84 percent. (*Source*: FAA, 1989.)

In a manual wind analysis, two parallel lines are drawn for each runway orientation examined. These lines are tangent to the circle corresponding to the allowable crosswind limit. For example, one of the runway orientations of interest in Fig. 9.2 is, in opposite directions, 105° and 285° (true) and the allowable crosswind limit is 10 knots. All winds within the rectangle drawn in this way are "covered" for crosswinds of 10 knots or less by a runway with that orientation; those outside the rectangle are not. In Fig. 9.2, two bidirectional runways, one with a 105° and 285° orientation and the other with 15° and 195°, together provide 98.84 percent crosswind coverage when the crosswind limit is 10 knots. In practice, a simple computer program performs this procedure.

Wind analyses are important in the design of new airports and in any study of the effects of weather on existing airports. For this reason, the FAA and the ICAO recommend that extensive historical wind records be compiled at each actual or potential airport site, preferably covering up to 10 consecutive years.

9.4 Airport Layouts

This section provides a brief descriptive review of airport layouts, with emphasis on the configuration of the airfield and of systems of runways. It makes several general observations concerning the range of land areas that airports occupy and the principal factors that influence this parameter. It concludes with a survey of commonly used generic types of airport layouts and a summary of some of their properties and characteristics.

Land Area Requirements and Related Observations

Airfields largely determine the total land area occupied by an airport and play a critical role in determining the airport's functionality and capacity. How much land area is actually occupied by an airfield depends on many factors. Principal among them are the following:

- The number, layout, and geometry of the runways, including runway length, separations between parallel runways, angles between nonparallel runways, ARCs selected for the purposes of airfield design, etc.
- The location of the landside facilities relative to the airside facilities
- The additional land area held in reserve for future expansion or to provide a "buffer" area for mitigation of noise and other environmental effects

The land areas that airports occupy span an enormous range, while the correlation between land area and the amount of traffic *processed* is rather weak. For example, the land area of New York/LaGuardia, which covers 2.6 million m^2 (less than 1 sq. mile) is equal to less than 2 percent of Denver/International's, which covers 136 million m^2 (54 sq. miles). However, in 2011, LaGuardia processed almost half as many passengers as Denver (24 vs. 52.7 million). Atlanta at 89.3 million passengers had nearly 80 percent more passengers than Denver, while occupying only 25 percent as much space. There is, however, a strong correlation between an airport's land area and its *potential capacity* to handle passengers and aircraft.

The airfield takes up most of the land area occupied by an airport. All the landside facilities (passenger buildings, cargo areas, on-site access roads, car parking, etc.) typically occupy only between 5 and 20 percent of the total land area, with the larger percentages applying to airports with small land areas, such as New York/LaGuardia and Washington/ Reagan (3.8 million m^2). The other 80 to 95 percent is dedicated to the complex of runways, taxiways, and aprons.

The number of runways needed to serve air traffic demand is a critical factor in determining land area requirements for airfields. In this respect, there exists an interesting systemic difference among regions of the world. This is suggested by Table 9.4, which lists the 30 busiest airports in the world in 2010, ranked by number of passengers. Of those, 13 were in North America, 10 in Asia[7] and Oceania (Sydney), and 7 in Europe.[8] The table also shows the number of aircraft movements at these airports and the number of passengers per movement. A distinct characteristic of North American (the United States and Canada) airports is that they process very large numbers of aircraft movements. Note, for example, that Chicago/O'Hare, London/Heathrow, and Tokyo/Haneda had similar numbers of passengers in 2010, but Tokyo/ Haneda handled only 75 percent as many movements as London/Heathrow and 39 percent as many as Chicago/O'Hare. This also meant a relatively small number of passengers per movement at North American airports, compared to European and especially to Asian airports. These observations are further borne out by Table 9.5, which provides relevant statistics for the 15 busiest airports in Asia, Europe, and North America in 2010. It shows that the average number of passengers per airport was about 9 percent smaller in the 15 busiest airports in Asia/Oceania than in North America, while the average number of movements per airport was 45 percent smaller, resulting in 67 percent more passengers per movement (145 vs. 87) in Asian airports. The 15 busiest European airports averaged 110 passengers per movement or 26 percent more than the ones in North America.

	Passengers (millions)	Movements (thousands)	Passengers/ Movement
Atlanta	89.3	950	94
Beijing	73.9	518	143
Chicago/O'Hare	66.8	883	76
London/Heathrow	65.9	455	145
Tokyo/Haneda	64.2	343	187
Los Angeles/International	59.1	667	89
Paris/de Gaulle	58.2	500	116
Dallas/Ft. Worth	56.9	652	87
Frankfurt/International	53.0	464	114
Denver	52.2	630	83
Hong Kong/Chek Lap Kok	50.3	307	164
Madrid	49.8	434	115
Dubai	47.2	293	161
New York/Kennedy	46.5	400	116
Amsterdam	45.2	402	112
Jakarta	44.4	339	131
Bangkok/Suvarnabhumi	42.8	267	160
Singapore	42.0	264	159
Guangzhou	41.0	329	125
Shanghai/Pudong	40.6	332	122
Houston/Intercontinental	40.5	531	76
Las Vegas	39.8	506	79
San Francisco/International	39.3	387	102
Phoenix	38.6	449	86
Charlotte	38.3	529	72
Rome/Fiumicino	36.2	329	110
Sydney	36.0	303	119
Miami/International	35.7	376	95
Orlando/International	34.9	308	113
Munich	34.7	390	89

Source: ACI, 2011.

TABLE 9.4 Traffic at the World's 30 Busiest Airports in 2010

15 Busiest Airports in...	Average Annual Passengers (millions)	Average Annual Aircraft Movements (thousands)	Average Passengers per Movement
North America	46.9	541	87
Europe	35.7 (–24%)	325 (–40%)	110 (+26%)
Asia	42.9 (–9%)	296 (–45%)	145 (+67%)

Source: ACI, 2011.

TABLE 9.5 Comparison of 15 Busiest Airports in North America, Asia/Oceania, and Europe

The explanation for these large regional differences lies in the fact that the aircraft mix at the busiest airports in Asia includes a large percentage, often in excess of 40 percent of the total, of large, wide-body jets, performing long-range—often intercontinental—flights. By contrast, the aircraft mix at North American airports is dominated (roughly 90 percent of the total movements at the 15 busiest airports) by a combination of narrow-body jets, regional jets, and nonjets flown by regional airlines. A large fraction of the passengers on the latter two types of aircraft connect to flights on the larger airplanes. Europe is in the middle: following the 1993 deregulation of the airline industry in the European Union, there has been a sharp increase in regional flights at European airports, including "feeder" flights with smaller aircraft. Thus, the busiest North American and, to a lesser extent, European airports need more runways than their Asian counterparts to serve the same number of passengers. The 34 busiest airports in the United States had an average of 4.1 runways in 2010 and the 34 busiest in Europe 2.5, whereas only one major airport in Asia had four runways (Tokyo/Haneda) and only a handful had three.

However, a major change has been taking place at the busiest Asian airports, as a result of the very rapid growth in: domestic air travel in China, India, Indonesia, and several other countries; regional services, especially in Southeast Asia; and low-cost carriers (see Chap. 1). Because narrow-body jets primarily provide these domestic, regional, and low-cost services, the large increase in the numbers of passengers at Asian airports has had to be absorbed through a roughly equally large increase in the number of aircraft movements. As of 2012, Asian airports were experiencing a runway capacity crisis: practically every major airport in the region had either constructed recently one or more new runways or was in the process of planning for such expansion.

The need for additional runway capacity, either by adding new runways to existing airports or by building new airports, is therefore global. The need will persist for as long as air traffic growth outpaces gains in runway capacity obtained through improvements in the ATM system. Building new runways is, however, an extremely difficult proposition politically in many countries. In addition, when expansion is feasible, it is typically a time-consuming and expensive process. Recent runway addition projects at Atlanta, Seattle/Tacoma, and St. Louis/Lambert all cost upward of $1 billion, whereas the Chicago/O'Hare modernization plan that emphasizes runway modifications and additions has an estimated cost in the order of $10 billion. The new second runway on a landfill at Osaka/Kansai that opened in 2007 cost around $10 billion, while in early 2012 Hong Kong/Chek Lap Kok announced plans to build a third runway on landfill (and expand landside facilities) at an anticipated cost of approximately $17 billion.

Airport Layouts

The geometric characteristics of the runway system and of the airfield are very important in determining land area requirements. The length of the runways is a simple example. At major airports they can be as short as 1500 m (roughly 5000 ft) and as long as 4000 m (13,000 ft)—or even longer at high elevations. Obviously, the longer the runways, the more land area is needed, especially if the airport property is to have a regular and compact shape, such as a rectangle. When parallel runways are present, the separation between their centerlines is another parameter critical to determining how much space the airport will occupy and where the landside facilities will be placed. Finally, the physical dimensions of runways, taxiways, and aprons (width, separations between runways and neighboring taxiways, taxilanes, aprons and buildings, separations between parallel taxiways, etc.) depend on the ARC selected for design purposes. The following brief survey of some important types of airport layouts illustrates these points.

Several major and many secondary airports have only a single runway. Because of limited land availability at their sites, it is also unlikely that many of these airports will ever add second runways. Their geometric layout is quite simple, as illustrated by Fig. 9.3, which sketches London/Gatwick. Although the airport seems to have a pair of close parallel runways, it actually operates with a single runway at all times because of the proximity of 08L/26R and 08R/26L: the former is normally used as a taxiway for the latter, except at times when 08R/26L is closed down for maintenance and repairs. At single-runway airports, the landside facilities are to the side of the runway, with passenger buildings and cargo buildings possibly on opposite sides. However, because of site limitations, the buildings may sometimes be inconveniently located relative to the runway. Note that at London/Gatwick, the location of the main passenger building necessitates taxiing distances of about

3500 m (2.2 miles) for aircraft taking off from 08R, which is the case much of the time. The nature of the traffic that single-runway airports handle is dictated in large part by the length of the runway. Milan/Linate, with a 2700-m (8860-ft) runway, is limited to short- and medium-range flights. London/Gatwick and Geneva both have long runways and can accommodate intercontinental, long-range flights. Single-runway airports may be able to handle surprisingly large numbers of passengers, especially if the mix of aircraft includes a high fraction of wide-body aircraft. London/Gatwick handled 33.7 million passengers in 2011 and ranked 36th in the world. San Diego, heavily utilized by medium-size and smaller aircraft, processed 200,000 aircraft movements and 16.8 million passengers that year.

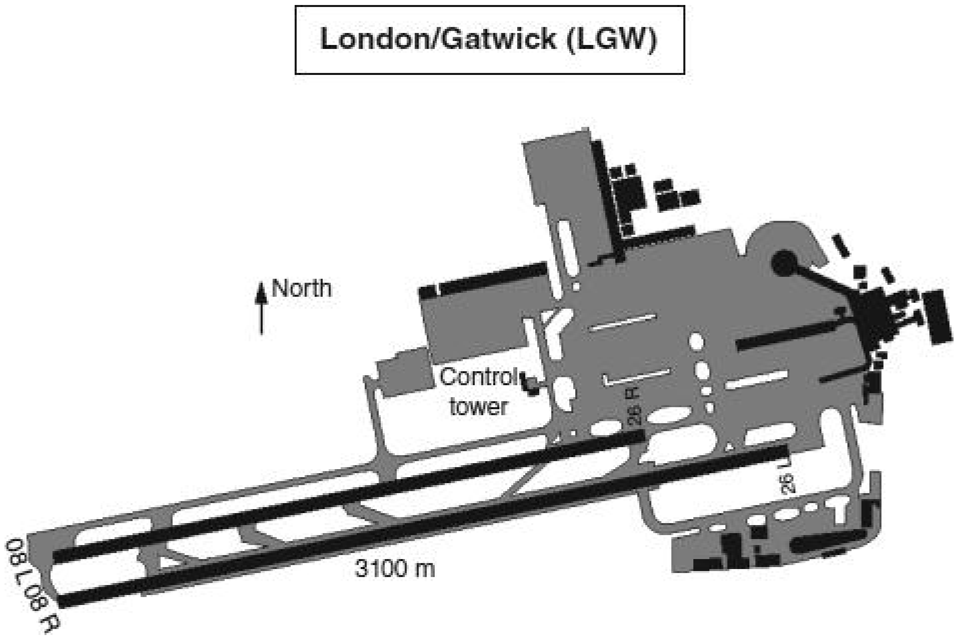


FIGURE 9.3 Layout sketch of London/Gatwick.

The runway systems of many major airports consist of two parallel runways. Depending on the separation between the runway centerlines, these airports are said to have "close," "medium-spaced," or "independent" parallel runways. Although standards differ somewhat from country to country, *close parallel runways* are generally those with centerline separations of less than 2500 ft (762 m). Under instrument flight rules (IFR), movements of aircraft on the two close parallel runways[9] must be coordinated (see Chap. 10). Independent parallel runways, on the other hand, are those whose centerlines are separated by distances greater than 4300 ft (1310 m) or 5000 ft (1525 m), depending on runway instrumentation and the ATM rules in different countries (see Chap. 10). As the name suggests, any pair of aircraft movements on the two independent parallel runways need not be coordinated, as long as a number of ATM and terminal airspace conditions are satisfied. Independence allows simultaneous parallel approaches to the two runways.

Medium-spaced parallel runways lie between the two extremes of close parallel and independent parallel. They permit independent departures from the two runways or independent "segregated" parallel operations, meaning that one runway can be used for arrivals and the other, independently, for departures. However, arrivals on two medium-spaced runways are not independent (see also Chap. 10).

The capacity of the runway system obviously generally increases as one moves from a close to a medium-spaced to an independent pair of parallel runways. However, close and medium-spaced parallel runways may be able to generate capacities close to those of an independent pair when operated under visual flight rules (VFR) in good weather, as they often are in the United States. Parallel runway operations under VFR can be conducted on pairs separated by as little as 700 ft (214 m) according to the FAA and the ICAO, although the FAA recommends 1200 ft (366 m) for runways serving Group V and VI aircraft. Chapter 10 provides details on all these points.

Close parallel and medium-spaced parallel runways do not provide sufficient space for the development of a landside complex between them. Thus, landside facilities at these airports are generally located to one or both sides of the runway pair. Layouts based on two close parallel runways or on two medium-spaced parallel runways (rather than two independent ones) are necessitated by such reasons as limited land availability, environmental restrictions and irregular shape of the airport site. Important examples include Philadelphia, New York/Newark, Frankfurt/International, and Milan/Malpensa.[10] Figure 9.4 shows a sketch of Frankfurt/International, where a third runway supplements the close pair of parallel runways. This third runway usually serves only takeoffs (and only to the south), primarily because of environmental restrictions.[11] Note that all four of the airports mentioned above are on older sites. Because of their local and regional importance, they have all undergone (or are undergoing) major infrastructure improvements and, in some cases, limited expansion of the available land area. All these airports share an important disadvantage: aircraft operating on the runway farther from the passenger building and main apron area must usually cross an active runway or its extension. This increases surface traffic delays and taxi times, as well as air traffic controller workload (see Chap. 10).

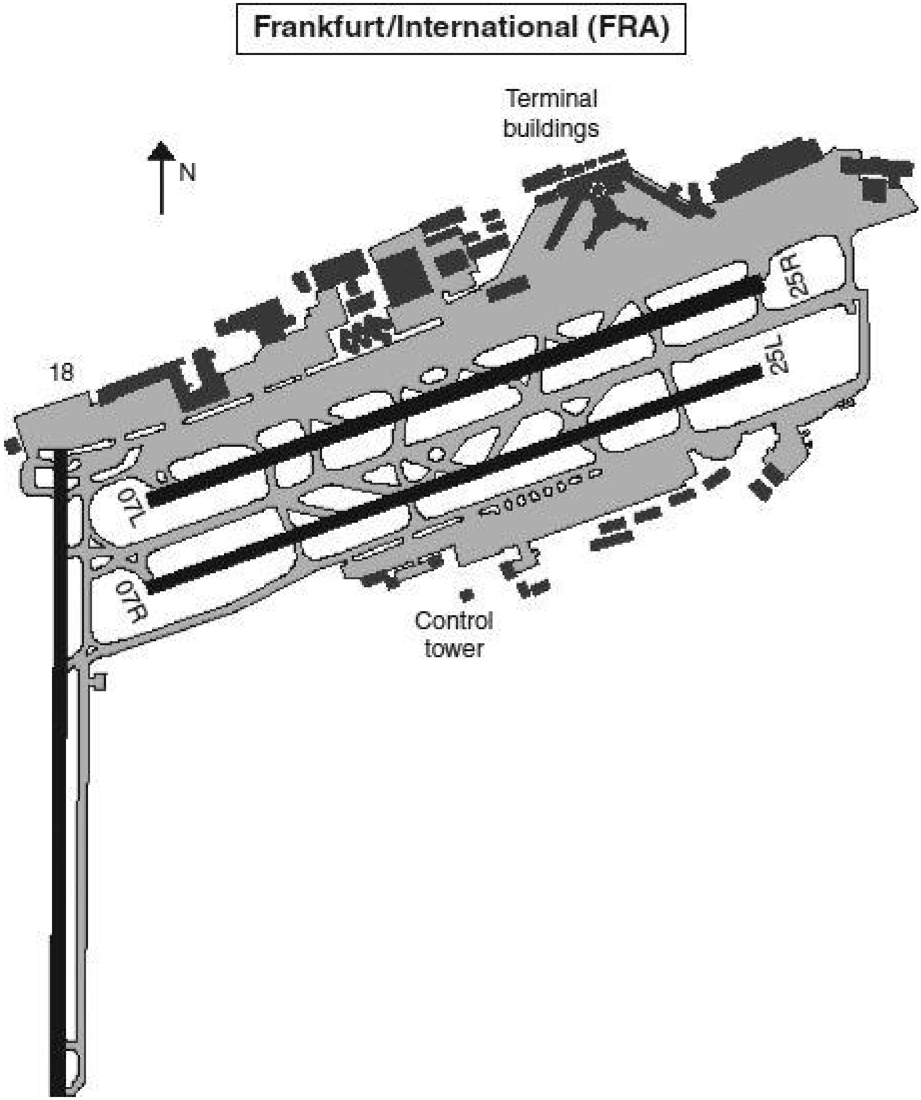


FIGURE 9.4 Layout sketch of Frankfurt/International.

When the runway system occupies the central part of an airport's site, as in the cases of single and close or medium-spaced parallel runways, it is important to have all passenger buildings on the same side of the runways. A number of airports, such as Manila/Aquino, Moscow/Sheremetyevo, Sydney (and other major Australian airports), and (in the future) Frankfurt/International, have passenger buildings on both sides of the runway system. Transfers of passengers and bags between buildings are then difficult, expensive, and time-consuming. There is also wasteful duplication of services and facilities on the two sides of the airport and limited opportunity to achieve economies of scale through sharing of common areas (see Chap. 15). Airlines and connecting passengers generally dislike using airports with these split landside arrangements.

Independent parallel runways provide sufficient space between them to accommodate the bulk of an airport's landside facilities, especially when spaced by 5000 ft (1525 m) or more, as is usually the case. The landside facilities are built mostly along the central axis of the airport. Some of the busiest airports in the world and, especially, many new airports that have started operations since 1990 belong to this category. London/Heathrow, Singapore, Hong Kong/Chek Lap Kok, Guangzhou, Jakarta, Kuala Lumpur, Athens, Oslo/Gardermoen, and Munich (Fig. 9.5) are examples. Some of the main advantages of this popular family of airport designs include the following:

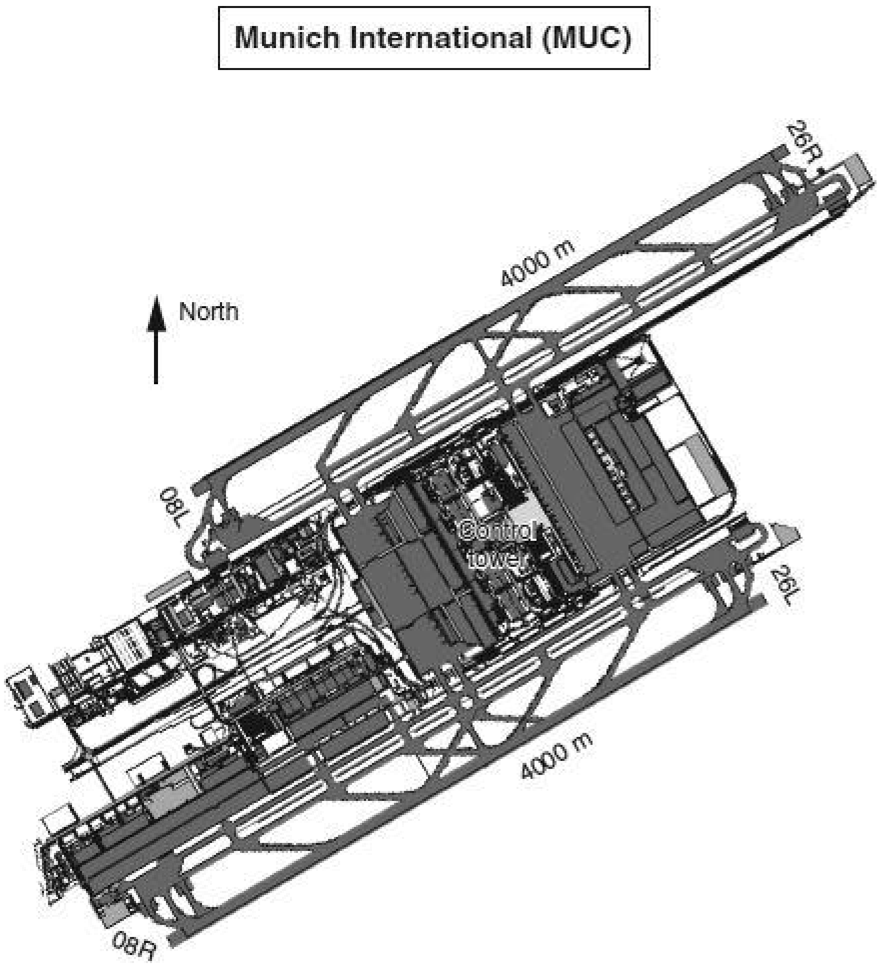


FIGURE 9.5 Layout sketch of Munich Airport.

- *Efficient utilization* of the vast land area between the independent runways, which would otherwise be greatly underutilized
- *Reasonable proximity of passenger and cargo buildings* to both runways, assuming that the landside configuration and apron and taxiway systems are well designed (see Chap. 15)
- *Better airfield traffic circulation*, as aircraft can reach either runway without having to cross another active runway
- *Ability to isolate the airport's landside from the airport's surroundings* and thus better control the landside's development, as well as ground access to the airport

There are disadvantages, as well. One stems from the fact that these layouts typically feature multilane access roads that provide ground connection to the local highway system—and may also have rail access to the passenger buildings. Such transportation links may cut across the entire length or at least a major part of the airport. To ensure good circulation of aircraft on the airport's surface, this necessitates the construction of an extensive taxiway system that includes expensive taxiway bridges passing over the access roads. Munich (see Fig. 9.5) has eight such bridges. A second disadvantage is that the placement of landside facilities along the central axis of the airport restricts somewhat the flexibility for expansion of these facilities when traffic grows. London/Heathrow is an extreme case in point: the space between its two parallel runways is completely saturated by now with various kinds of terminal buildings and other landside facilities.

A site of at least 11 million m^2, or roughly 5 km by 2.2 km, is needed to accommodate an airport with two independent long parallel runways of 3.5 to 4 km length, separated by at least 1525 m between their centerlines.

At several of the airports that operate with two independent parallel runways, the runways are "staggered." This is shown in Fig. 9.5 for Munich: the threshold of runway 08L is displaced relative to the threshold of runway 08R, and the same is true for the thresholds of runways 26R and 26L. One of the benefits of staggered runways is that they provide additional vertical separation between aircraft operating on the two runways. For example, when two aircraft are performing simultaneous parallel approaches to the two runways (e.g., 08R and 08L at Munich, see Fig. 9.5), the aircraft landing on the "farthest" runway (08L in this case) is at a higher altitude than the aircraft aiming for the "closer" one (08R).

Another advantage of staggered runways is the reduced taxiing distances when one runway is used for arrivals only and the other for departures only, as often done in practice.

For instance, when Munich is operating in an eastern orientation, the use of runway 08R for arrivals and 08L for departures reduces the taxiing distances for both landing and departing aircraft between their runway and the terminal area. Conversely, when operations are to the west, assigning arrivals to 26R and departures to 26L accomplishes the same objective. Munich indeed obtains tangible taxiing distance benefits in this way. However, to increase runway capacity as traffic increases, airports are often forced to mix arrivals and departures on both runways (see Chap. 10). In such cases, the reduced taxiing advantage of staggered runways may be largely lost.

A possible disadvantage of staggered runways is that they increase the land area required if the airport property is to retain a rectangular or nearly rectangular shape. Land acquisition is almost always a problem in airport development.

Airports often have runways whose orientations differ. The runways intersect in such cases, either physically or along their projected centerlines.[12] New York/LaGuardia is an example where the runways intersect physically (Fig. 9.6). Airfield geometries with intersecting runways may be necessary at sites that often experience strong winds from several different directions. The different orientations of the runways make it possible to operate the airport under most weather conditions and provide the 95 percent or greater crosswind coverage recommended by the ICAO, the FAA, or other civil aviation authorities (see Sec. 9.3).

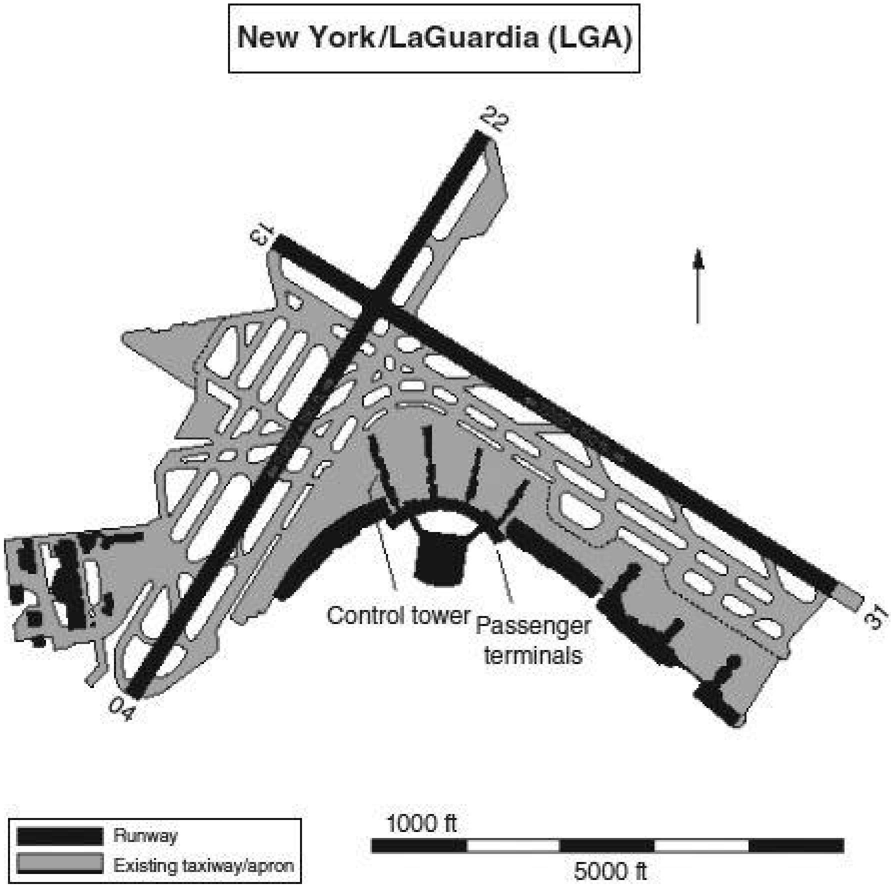


FIGURE 9.6 Layout sketch of New York/LaGuardia.

Airports with intersecting runways are more difficult to operate from the ATM point of view. From Fig. 9.6 it is clear that when both runways are active, aircraft movements on each must be carefully coordinated with those on the other runway. Moreover, the capacity of the runway pair will vary depending on the direction in which the operations take place and the location of the intersection point (see Chap. 10). When strong winds in one direction force one of the two runways to close down, the airfield capacity is also affected in a major way. Thus, airports with intersecting runways are often challenging operationally, especially when local weather conditions are highly variable.

The simplest way to increase the capacity of airports with two independent parallel runways is by adding a close parallel runway next to one of the two existing ones. Examples include Seoul/Incheon, Shanghai/Pudong, and, in the near future, Bangkok/Suvarnabhumi. The airport thus ends up with a pair of close parallel runways and a third independent runway. In such situations, one of the close parallels is used solely for arrivals, the other solely for departures, and the third independent runway for mixed operations (arrivals and departures).

Airports with two independent pairs of close parallel runways, one pair on each side of the landside complex, represent the next stage of capacity evolution. Examples include Paris/Charles de Gaulle, Los Angeles/International, Delhi/Indira Gandhi (in the near future), and, until 2006, Atlanta (Fig. 9.7). These airports typically operate each close pair of runways by having departures on the inner runways (08R/26L and 09L/27R in Atlanta) and arrivals on the outer (08L/26R and 09R/27L). The distance between the two runways used for arrivals is sufficient to operate these two runways independently. This runway layout is well suited for airports processing 50 million or more passengers per year. The complex of the four runways can provide a total capacity of 140 or more movements per hour, even under IFR. Airports with a land area of 15 million m^2—and preferably more—may be able to accommodate this type of layout.

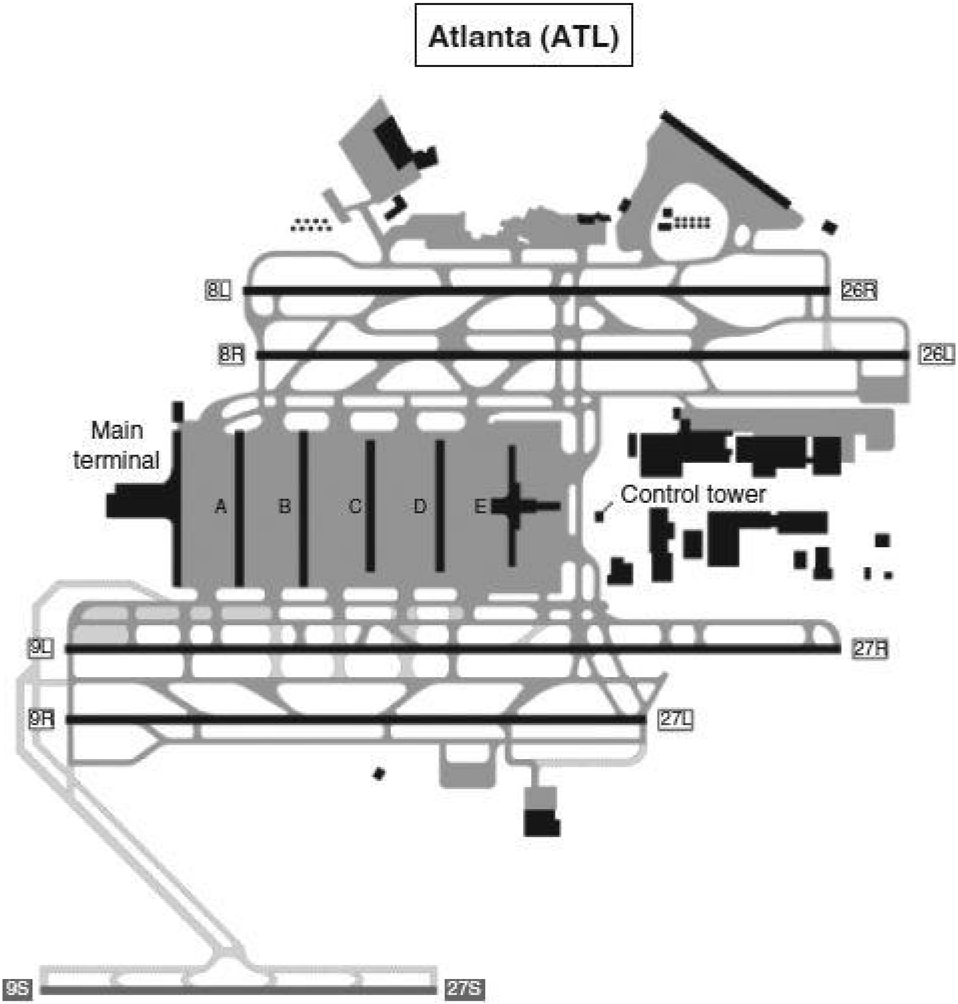


FIGURE 9.7 Layout sketch of Atlanta Airport.

Airports that occupy 30 million m^2 or more have enormous capacity potential, as they can accommodate six or more parallel (or nearly parallel) runways in the primary orientation of operations, if necessary. Moreover, at least three of these runways can be separated by more than approximately 5000 ft (1525 m) from one another, as required by the FAA in order for all three to be used for approaches simultaneously. This means a capacity of 100 or more arrivals per hour or 200 or more movements per hour when all six runways are active (see Chap. 10). The United States has several such potential "mega-airports" at Denver/International, Dallas/Ft. Worth (Fig. 9.8), Chicago/O'Hare, Atlanta, and Orlando/International.

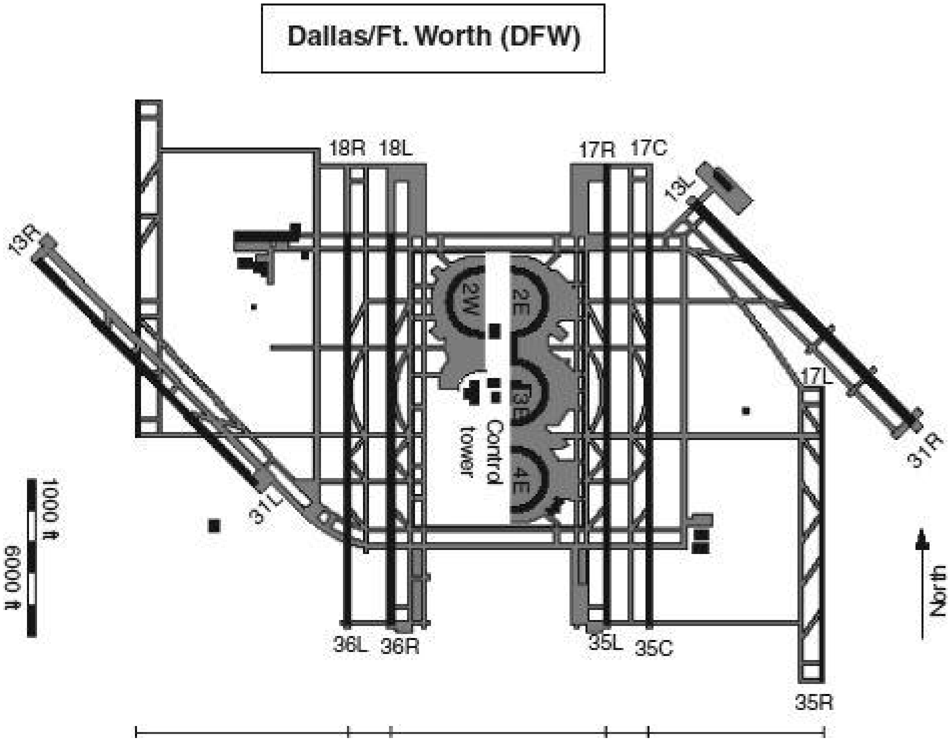


FIGURE 9.8 Layout sketch of Dallas/Ft. Worth; the north–south runway at the upper left (without a runway number) has not been constructed.

Allowing for local variations, the airfield layouts discussed previously are typical of those encountered at most major airports around the world. However, they do not exhaust the range of possibilities. Several multirunway airports have layouts that do not fit any of the models described so far. San Francisco/International, Toronto/Pearson, Amsterdam/Schiphol, and Zürich are examples of airports with four, five, six, and three runways, respectively, that fall in this category. Boston/Logan has six runways with three different orientations. Chicago/O'Hare has had seven runways with three different orientations, but its modernization program is transforming it to a configuration of six parallel and two crosswind runways.

9.5 Runway Length

Many factors affect the runway length required for the landing or takeoff of any aircraft on any given day. The most important among those are the following:

- Weight of the aircraft on takeoff or on landing and the settings of its lift- or drag-increasing devices (e.g., wing flaps)
- Stage length (or nonstop distance) to be flown
- Weather, particularly temperature and surface wind
- Airport elevation
- Runway characteristics, such as runway gradient (see Sec. 9.6) and surface condition (wet or dry pavement, surface texture)

The qualitative relationships between runway length required and these factors are quite obvious. The greater the total actual weight of the aircraft (operating weight empty plus payload plus fuel), the longer are the takeoff or landing distances. Longer stage lengths mean more fuel and thus higher likelihood that the aircraft will be operating at its maximum takeoff weight, and thus require longer takeoff distances. The stronger the headwind, the shorter the required runway length; conversely, a tailwind increases the length of runway required. High temperatures create lower air densities, resulting in lower output of thrust and reduced lift, thus increasing runway length required. An airplane taking off on an uphill gradient requires more distance than one on a level or downhill gradient. Similarly, a wet runway will increase the runway length required, especially on landing. The higher the elevation of an airport, the longer will be the runway required, everything else being equal.

For runway design purposes, airport performance manuals (APM) capture many of these relationships quantitatively, in graphical or tabular form. Manufacturers prepare APM for each type of aircraft they produce. Appendix 1 of FAA (2005) provides the links to the manufacturer web sites for their APM. The APM assist in determining the design runway length, given (1) any set of local conditions, such as the airport's elevation and the mean maximum daily temperature during the hottest month of the year, and (2) the most demanding aircraft that will be using the runway. The relevant charts provided refer to standardized sets of conditions for landing and for takeoff. For instance, for the 767-200ER, Boeing's APM (Boeing, current) provides two charts for takeoff field length requirements: one (Fig. 9.9) for takeoff at zero wind, zero runway slope, and standard atmosphere, which includes 15°C as one of its attributes; and a second chart for the same conditions, but for a "standard atmosphere plus 17°C," that is, a temperature of 32°C. Similarly, the APM provides two charts for the required length of the landing runway: one for zero wind, zero runway slope, no reverse thrust, anti-skid on, automatic speed brakes and flaps set at 25°; and the other for the same conditions, but with flaps at 30°. Each of the landing runway charts shows the requirements for both dry runway and wet runway conditions.

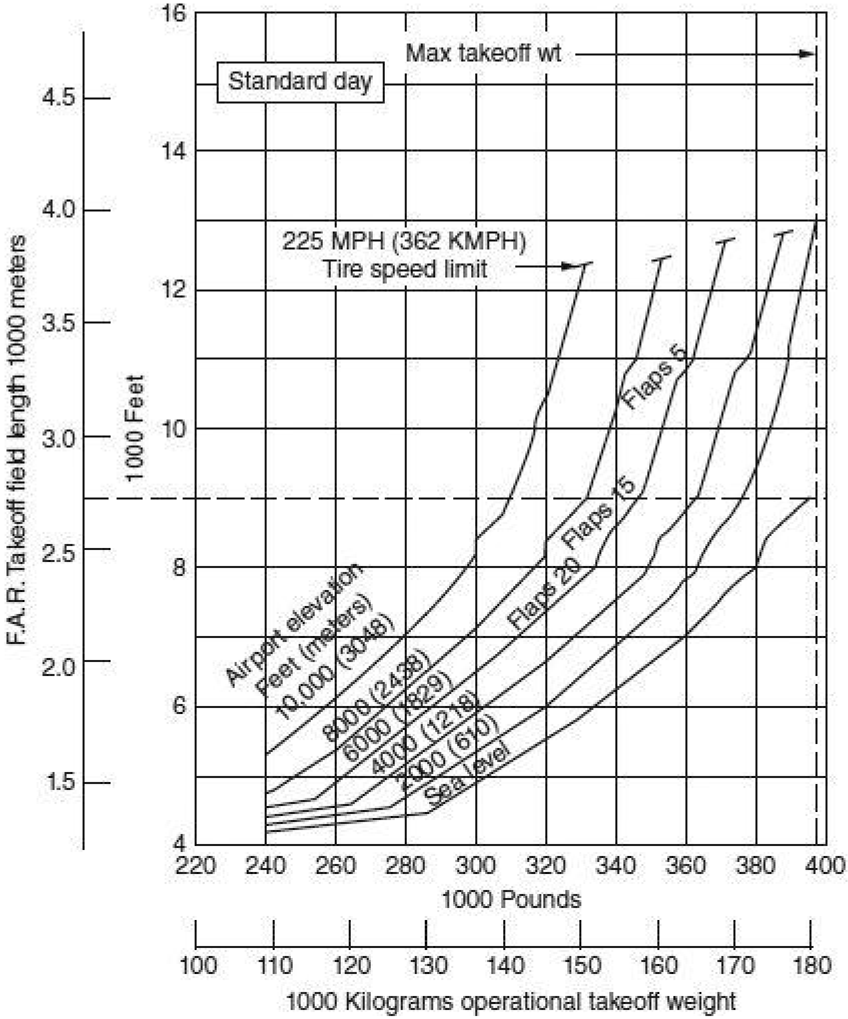


FIGURE 9.9 Take-off distances for the Boeing 767 at zero wind, zero runway slope, and standard atmospheric conditions. (*Source*: Boeing Company, current.)

It is easy to determine approximately the design length requirements using the APM charts. Consider, for example, Fig. 9.9. For a Boeing 767-200ER aircraft taking off at sea level under the conditions specified in the chart and at its maximum takeoff weight (179.2 tons, see Table 9.3*a*), the required takeoff runway length can be seen to be approximately 9000 ft or 2750 m (cf. Table 9.3*a*). Interpolation between the curves shown in Fig. 9.9 is permissible, both regarding elevation and operational takeoff weight. For example, the reader can verify by inspection that the required design takeoff runway length for this aircraft is approximately 8400 ft or 2560 m, if taking off under the same conditions assumed in Fig. 9.9, but with a planned operational weight of 170 tons and from a runway at an elevation of 1000 ft. Interpolation across charts prepared for different mean temperatures (e.g., 15°C and 32°C) is also permissible.

Adjustments must subsequently be made to APM-derived estimates of design length to account for local conditions (ICAO, 2006; FAA, 2005). For instance, an approximate guideline is that the design length should be increased by 10 ft (3 m) per foot (0.3 m) of difference between the high and low points of the runway centerline elevations (this is the "gradient adjustment). For example, 15 ft of difference between the high and low points of the runway requires 150 ft of additional runway length. Other important examples of approximate guidelines include that (a) takeoff runway length should be increased by 7 percent for each 300 m (~ 1000 ft) of airport elevation, and (b) the required landing runway length with a wet pavement is about 15 percent greater than with a dry pavement. Such approximations, however, are typically valid only for relatively limited deviations from the base length estimates.

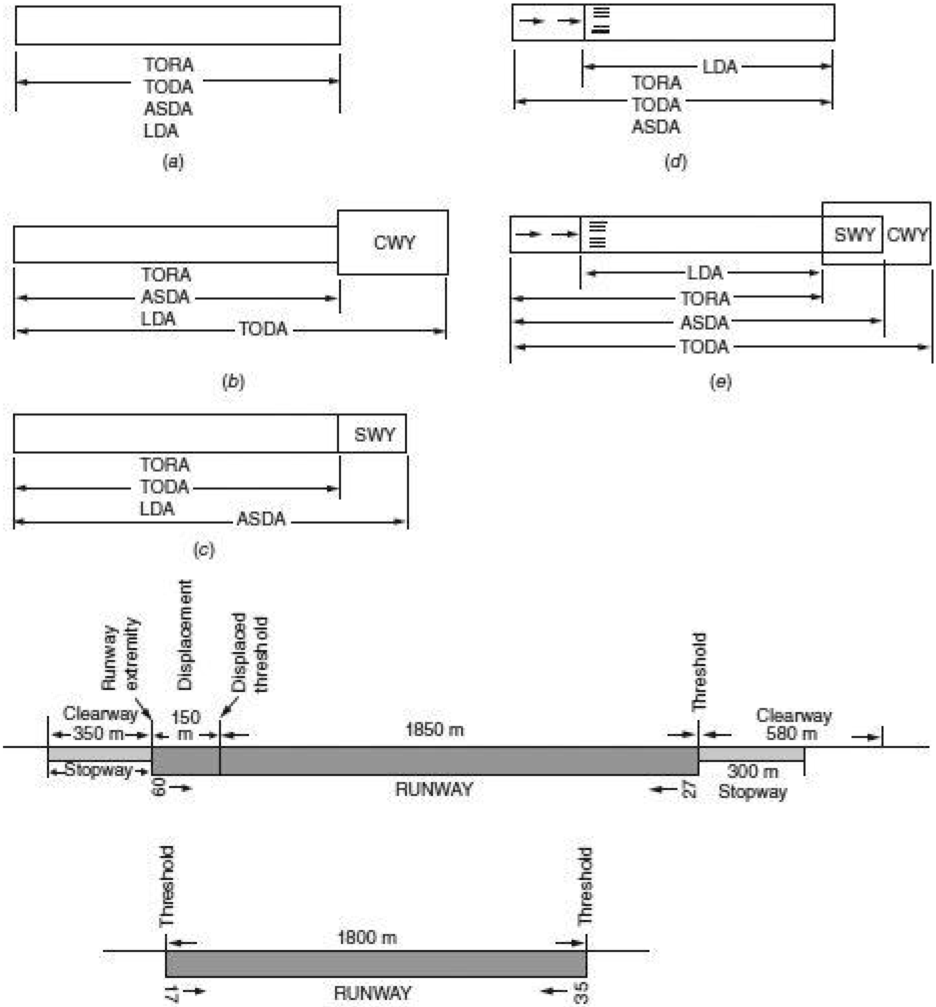
Readers should consult ICAO (2006) and FAA (2005) for additional details and guidance on the subject of design length. The next three subsections describe first the fundamental approach to determining whether a runway of a given length and characteristics can be used for the landing or takeoff of a specific aircraft ("usability") and then summarize some practical points concerning runway design length.

Declared Distances

The concept of *declared distances* is central to understanding the usability of a runway for any specific aircraft movement. It also helps explain how the design runway lengths at an airport are determined in the first place. For any given runway, four declared distances are defined (ICAO, 2006; FAA 2012):

- TORA, the *takeoff run available*: The length of runway declared available and suitable for the ground run of an airplane taking off
- TODA, the *takeoff distance available*: The length of the takeoff run available (TORA) plus the length of the clearway (see as follows), if one is provided
- ASDA, the *accelerate–stop distance available*: The length of the TORA plus the length of the stopway (see as follows), if one is provided
- LDA, the *landing distance available*: The length of the runway declared available and suitable for the ground run of an airplane landing

These definitions are now explained through Fig. 9.10. When there is no clearway, stopway, or displaced threshold, all four declared distances are equal to the length of the runway (see Fig. 9.10*a*). However, some runways may have one, two, or all three of these features.



RUNWAY	TORA	ASDA	TODA	LDA
	m	m	m	m
09	2000	2300	2580	1850
27	2000	2350	2350	2000
17	NU	NU	NU	1800
35	1800	1800	1800	NU

(*f*)

FIGURE 9.10 Declared distances for a runway. (*Source*: ICAO, 2009.)

A *clearway* (CWY), when available, is a rectangular area, beginning at the end of the runway and centered on the runway's extended centerline, over which an airplane can make the initial portion of its flight on takeoff (CWY in Fig. 9.10). It can be on ground or water. It must be clear of any obstacles or terrain at an upward slope of 1.25 percent. Its width must be at least 500 ft (150 m per ICAO recommendations) and its length cannot exceed 1000 ft, according to the FAA's specifications.[13] It must be under the control of the airport operator or other appropriate organization. When a clearway is available (see Fig. 9-10*b*), TODA is equal to the sum of the lengths of the takeoff runway and the clearway.

A *stopway* (SWY), when available, is a rectangular area, beginning at the end of the runway and centered on the runway's extended centerline, which has been prepared as a suitable area where an aircraft can be stopped in the case of an aborted takeoff without suffering structural damage. The stopway must be at least as wide as the runway. When a stopway is available (see Fig. 9.10*c*), ASDA is equal to the sum of the lengths of the takeoff runway and of the stopway.

The *runway threshold* is the beginning of that portion of the runway, which is usable for landing. Because of the presence of obstacles on the approach path or for environmental or other local reasons, the threshold of a runway is sometimes *displaced* and does not coincide with the physical beginning of the runway. When this is the case (see Fig. 9.10*d*), LDA is reduced by the distance the threshold is displaced. Note that this does not affect LDA in the opposite direction of operations.

Figure 9.10*e* shows the four declared distances when a clearway, stopway, and displaced threshold are all present. Note that the stopway and the clearway necessarily overlap partially, if both exist. The lower part of Fig. 9.10 shows a display of declared distances at an airport with two runways, 09/27 and 17/35. The table gives the TORA, TODA, ASDA, and LDA for each of the four runway orientations. Runway 09/27 has asymmetrical clearways and stopways on its two ends, and a displaced threshold for landings on 09. Runway 17/35 is assumed to be unidirectional: it only serves arrivals on runway 17 and departures on 35. The notation "NU" means "not usable."

Usability of a Runway

Consider now the usability of a runway by a particular type of turbine-powered aircraft[14] (turbojet or turbofan). For the runway to be usable for landing, the aircraft *must be able to come to a full stop within a distance of at most 60 percent of the landing distance available*, LDA, assuming the aircraft makes a normal approach to the runway and flies over the threshold of the runway at a height of 50 ft (15 m). Note that this leaves a large margin of safety to account for deviations, such as coming over the threshold at a higher altitude or landing at a higher than normal speed.

The requirements are more complex when it comes to takeoffs. Both the takeoff distance available, TODA, and the accelerate–stop distance available, ASDA, must be considered. Two cases must now be analyzed, a normal takeoff and a takeoff during which the failure of an engine occurs. In the first case, a distance, TOD1, is computed that is equal to 115 percent of the distance needed by the aircraft to reach a height of 35 ft (10.7 m) with all engines assumed available throughout. This provides for a margin of 15 percent to allow for variability in performance and pilot technique, much greater than the variability normally expected.

The second case assumes that the failure of one engine occurs during takeoff and the pilot uses the following strategy: A decision speed V_1 (also known as critical engine-failure speed) is defined. If the failure occurs before the aircraft has reached V_1, the takeoff is aborted and the aircraft is brought to a stop. If, on the other hand, the failure occurs at a speed greater than V_1, the takeoff must continue as there is not enough distance left on the runway and, possibly, the stopway to brake to a stop. Two distances are then computed for the situation when the engine fails at exactly the decision speed V_1. One is the distance, TOD2, which is equal to the total distance (from start of the takeoff run) needed for the aircraft to attain an altitude of 35 ft (10.7 m) if it continues the takeoff under the conditions described. (Note that no margin, like the earlier 15 percent, is applied, as this is a low-probability case, given the reliability of turbine engines.) The second quantity, ASD, is the total distance (from start of the takeoff run) needed to bring the aircraft to a full stop if the takeoff is aborted.

For the runway to qualify for takeoffs by the aircraft in question, both of the following two conditions must be satisfied:

- TODA must be greater than or equal to the greater of TOD1 and TOD2
- ASDA must be greater than or equal to ASD

In words, this requires that TODA should be sufficient to accommodate a normal takeoff to 35 ft with a 15 percent margin, as well as an engine failure at a speed of V_1 or greater; and, at the same time, ASDA should provide sufficient distance for the aircraft to stop if an engine failure occurs at a speed of V_1 or less.

Obviously, both TODA (through TOD2) and ASDA depend on the decision speed V_1. What should this speed be? Note that as V_1 increases, the distance needed to come to a stop will also increase, but the distance needed to attain the height of 35 ft, if the takeoff continues with one failed engine, will decrease. The converse is also true.

It follows that, *when there is no clearway and no stopway,* V_1 should be such that the distance between the point where V_1 occurs and the point where the aircraft comes to a stop should be exactly equal to the distance between the same point and the point where the aircraft reaches the required altitude of 35 ft with one engine out. This means that TODA must be equal to ASDA. Moreover, both of these quantities should be equal to the length of the TORA; otherwise, the airport could not take advantage of the full length of the available runway. This simple relationship among these three quantities is referred to as the concept of *balanced field length*.

In the presence of a clearway and/or stopway, a somewhat different value of V_1 may be adopted, depending on the relative cost of preparing a clearway and or constructing a stopway. However, at major airports, where the TORA is typically much greater than that of any stopway or clearway that can be provided, the balanced field length approach provides optimal or near-optimal values for TODA and ASDA.

Practical Considerations

The approach for deciding the appropriate length of the runway (or runways) to be provided at an airport follows directly from the preceding discussion. The airport designer must select the most demanding aircraft type (*critical aircraft*) to be accommodated by a runway and the *conditions of use* by that critical aircraft. The runway length that will accommodate that aircraft under these conditions is then computed. "Conditions of use" essentially mean

- The longest nonstop distance (stage length) to be flown by the critical aircraft from/to the runway
- The most demanding environmental conditions during runway use, such as the mean daily temperature for the hottest month of the year at the airport

The FAA recommends (FAA, 2005) that the critical airplane/ flight combination should be a service operating for at least 250 days in a year (= 500 landings and takeoffs), such as a flight scheduled on a Monday-through-Friday basis every week. A better approach is to identify target markets that an airport should be designed to serve as part of its long-term strategy and then design, but not necessarily build, accordingly. It should be emphasized, however, that such target markets should be chosen realistically. For example, several secondary airports in Europe built in the 1980s and 1990s runways long enough to serve scheduled intercontinental flights to the United States with Boeing 747 aircraft, but such flights never materialized.

Another common mistake in selecting runway lengths is building too many very long runways. Consider an airport with two parallel runways in its primary orientation of operations. At least in the early phases of the airport's development, the airport operator may not wish to construct both runways to equal length, if the number of flights requiring a long runway is not large. For example, building one of the runways to a length of 3600 m (11,800 ft) and the other to 2700 m (8900 ft) may be perfectly adequate for an airport at sea level in an area with nonextreme summer temperatures. The long runway will suffice for the takeoff of practically any long-range flight, whereas the shorter one will be adequate for nearly any short- and medium-range flight. The two runways can then share all operations, except for long-range flights, which must be assigned to the long runway. If long-range traffic increases to the point where two long runways are necessary, the shorter runway can be lengthened—assuming that adequate plans have been made for this eventuality.

Finally, the following should be interpreted as *only rough* (and somewhat conservative) *indications* of the typical length of runway needed to serve various types of airline flights with jet aircraft at sea level and nonextreme climates. They should *not* be used as substitutes for detailed computation of runway length requirements:

- 2000 m (6600 ft) will accommodate regional jets and many short-range flights (up to roughly 2000 km or 1200 miles) by narrow-body conventional transport jets.
- 2300 m (7500 ft)—practically all short- to medium-range flights (3000 km or 1900 miles).
- 2700 m (8900 ft)—practically all medium-range flights (4500 km or 2500 miles).
- 3200 m (10,500 ft)—most longer-range flights (9000 km or 5600 miles).
- 3500–4000 m (11,500–13,100 ft)—all feasible stage lengths in other than extremely high temperatures.

9.6 Runway Geometry

The ICAO and the FAA specify the design standards for all the geometric characteristics of runways, other than runway length, on the basis of the applicable ARC. The required or recommended dimensions and separations ensure the safe operation of the most demanding aircraft type for which the runway is designed. The FAA provides dimensional standards for the runway itself and for runway-associated elements that include the following:

- Runway shoulders
- Runway blast pad
- Runway safety area (RSA)
- Obstacle-free zone (OFZ)
- Runway object-free area (ROFA)

- Clearway
- Stopway

Section 9.5 discussed clearways and stopways. The following text briefly describes the other elements. Table 9.6 summarizes the associated FAA dimensional standards. Additional details can be found in FAA (2012), its amendments, appendices, and supporting documents. The ICAO design standards are similar, although the terminology may be somewhat different.

	Airplane Design Group					
	I	II	III	IV	V	VI
Runway width	100 ft 30 m	100 ft 30 m	150 ft 46 m	150 ft 46 m	150 ft 46 m	200 ft 61 m
Runway shoulder width	10 ft 3 m	10 ft 3 m	25 ft 8 m	25 ft 7.5 m	35 ft 10.5 m	40 ft 12 m
Runway blast pad width	120 ft 37 m	120 ft 37 m	200 ft 61 m	200 ft 61 m	220 ft 67 m	280 ft 85 m
Runway blast pad length	100 ft 30 m	150 ft 46 m	200 ft 61 m	200 ft 61 m	400 ft 122 m	400 ft 122 m
Runway safety area (RSA) width length beyond runway end	 500 ft (152 m) 1000 ft (305 m)					
Runway object-free area (ROFA) width length beyond runway end	 800 ft (244 m) 1000 ft (305 m)					
Runway obstacle-free zone (ROFZ) width length beyond runway end	 400 ft (122 m) 200 ft (61 m)					
Precision obstacle-free zone (POFZ) width length	 800 ft (244 m) 200 ft (61 m)					

*Dimensions shown are for sea-level runways; for adjustments for runway elevation, see FAA (2012).

Source: FAA, 2012.

TABLE 9.6 Dimensional Standards* for Runways

The *runway shoulders* (Fig. 9.11) are adjacent to the structural pavement of the runway. They provide resistance to jet blast erosion and accommodate maintenance and emergency equipment. A natural surface with dense, well-rooted turf may suffice for the runway shoulders at secondary airports, but paved shoulders are recommended for runways (as well as taxiways and aprons) that accommodate Group III or higher aircraft. Groups V and VI normally require paved shoulder surfaces.

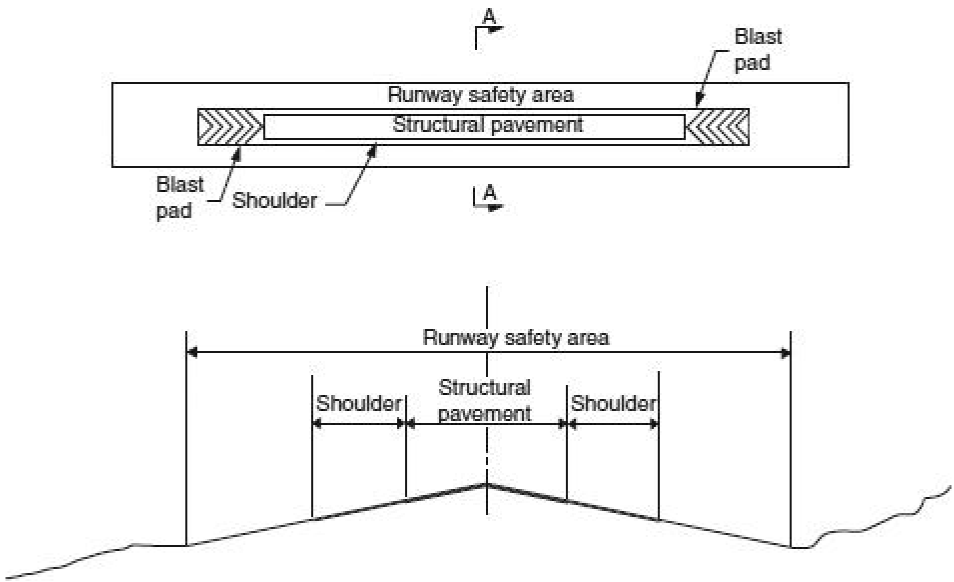


FIGURE 9.11 The runway safety area. (*Source*: FAA, 2012.)

Runway blast pads (Fig. 9.11, top) provide blast erosion protection beyond runway ends. They should extend across the full width of the runway plus its shoulders.

The *runway safety area* (RSA) includes, but is not limited to, the structural pavement, runway shoulders, runway blast pads, and stopways (Fig. 9.11). It was formerly known as the "landing strip." It must be cleared, graded, and free of hazardous surface variations; drained through sufficient grading or storm sewers; capable of supporting snow-removal equipment, rescue, and firefighting equipment and the occasional passage of aircraft without causing structural damage to the aircraft; and free of objects, except those that functionally need to be located in the RSA (such as runway lights). Objects higher than 3 in (7.6 cm) above grade should be constructed on frangible mounted structures with the frangible point no higher than 3 in above grade.[15]

The *obstacle-free zone* (OFZ) defines a volume of protected airspace below 150 ft (45 m) above the established airport elevation (Fig. 9.12). It is centered above the runway and the extended runway centerline and is intended to provide clearance protection for aircraft landing or taking off from the runway and for missed approaches ("balked landings" in ICAO terms). It must be clear of all objects, except for frangible visual navigational aids that need to be located within the OFZ. The OFZ is subdivided into the following:

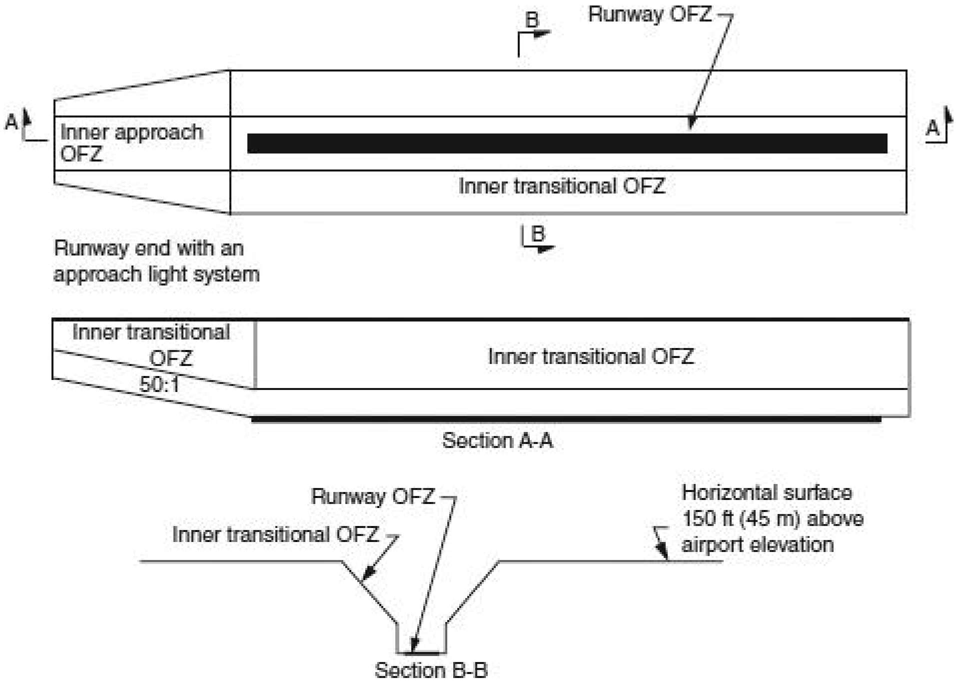


FIGURE 9.12 Obstacle-free zone for runways serving large aircraft with lower than 0.75-statute-mile (1200-m) approach visibility minimums. (*Source*: FAA, 2012.)

- *Runway OFZ*, the volume of airspace above a surface centered on the runway centerline and extending 200 ft (61 m) beyond each end of the runway
- *Inner-approach OFZ*, centered on the extended runway center-line and applicable only to runways with an approach lighting system; it is intended to protect the approach to the runway
- *Inner-transitional OFZ*, the airspace above the surfaces located on the outer edges of the runway OFZ and the inner-approach OFZ and applicable only to runways with an approach visibility minimum lower than 0.75 mile (1.2 km)

The *runway object-free area* (ROFA) occupies ground centered on the centerline of a runway. It is an area kept free of all objects, except those needed for air navigation or aircraft maneuvering purposes. It is acceptable to taxi or temporarily hold aircraft within the ROFA, when necessary. Note (Fig. 9.13) that the ROFA surrounds a runway and extends beyond its ends.

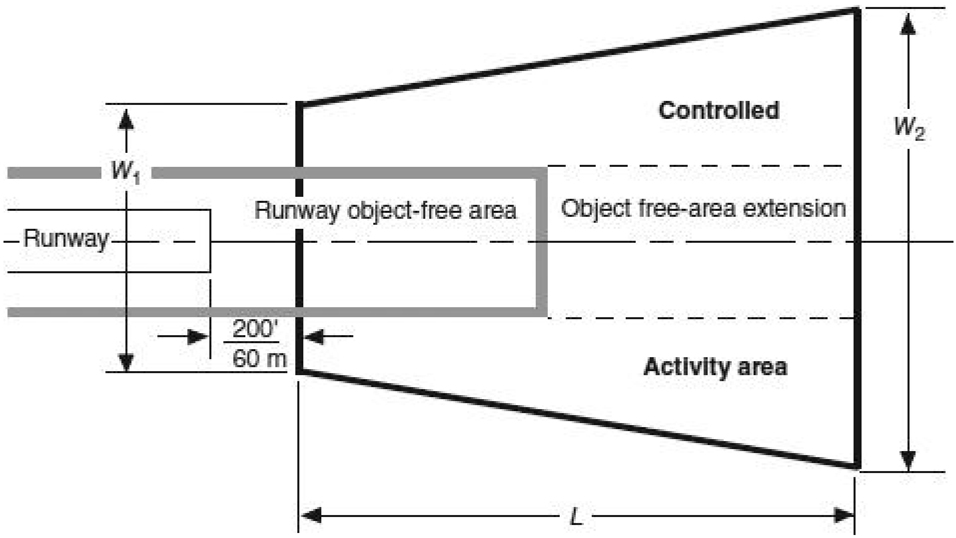


FIGURE 9.13 Runway protection zone. (*Source*: FAA, 2012.)

Beginning more recently, the FAA requires that any new authorized precision approaches with less than 0.75-mile visibility also provide a *precision object-free area* (POFA). The POFA is centered on the runway centerline, begins at the runway threshold, extends to 200 ft along the runway's centerline, and is 800 ft wide (FAA, 2012). In Fig. 9.13, the POFA would occupy the gap between the runway threshold and the runway protection zone and extend 400 ft to each side of the extended runway centerline.

In addition to the elements identified above, the *runway protection zone* (RPZ) is an area off the end of the runway intended to enhance the protection of people and property on the ground.[16] As Fig. 9.13 shows, the RPZ contains a part of the ROFA. The portion of the RPZ beyond and to the sides of the ROFA, the *controlled activity area*, should be under the control of the airport operator, as much as possible. The controlled activity area should be reserved for uses and activities that do not interfere with airport operations and with navigational aids. Some agricultural activities that do not attract wildlife are expressly permitted, but residences, places of public assembly, and fuel storage facilities are prohibited. For precision approach runways, the length, L, of the RPZ (see Fig. 9.13) should be 2500 ft (750 m) and the widths, W_1 and W_2, 1000 and 1750 ft, respectively. This results in a total land area for the RPZ of about 80 acres (32 ha) at each end of the runway.

Separations from Other Parts of the Airfield

The required distances between runways and other parts of the airfield are an important aspect of geometric design. Table 9.7 summarizes some of the FAA standards for separations between runway centerlines and holdlines, taxiway/taxilane centerlines, and locations where aircraft park. Holdlines keep aircraft that are waiting to use a runway at a distance sufficient to ensure that no part of the aircraft penetrates any obstacle limitation surfaces. Holding aircraft and other vehicles behind holdlines should also not interfere with the operation of navigation aids.

	Airplane Design Group (ADG)					
	I	II	III	IV	V	VI
Runway centerline to:						
Parallel runway centerline (VFR operations)	700 ft 213 m	700 ft 213 m	700 ft 213 m	700 ft 213 m	1200 ft 213 m	1200 ft 213 m
Parallel runway centerline[a] (IFR operations)	4300 ft 1311 m	4300 ft 1311 m	4300 ft 1311 m	4300 ft 1311 m	4300 ft 1311 m	4300 ft 1311 m
Hold line[b]	250 ft 76 m	250 ft 76 m	250 ft[c] 76 m	250 ft[d] 76 m	280 ft[d] 85 m	280 ft[d] 85 m
Taxiway/taxilane centerline[e]	400 ft 122 m	400 ft 122 m	400 ft 122 m	400 ft 122 m	400/450/500 ft[f,g] 122/137/152 m	500 ft[g] 152 m
Aircraft parking area	500 ft 152 m					

[a]Applies to dual simultaneous approaches; the FAA will consider proposals with separations down to a minimum of 3000 ft (914 m) in cases where a 4300-ft (1311-m) separation is impracticable.

[b]For all aircraft with FAA reference letter D or E, this distance is increased 1 ft for each 100 ft above sea level.

[c]For ADG III, this distance is increased 1 ft for each 100 ft above 5100 ft above sea level.

[d]For ADG IV–VI, this distance is increased 1 ft for each 100 ft above sea level.

[e]Distances shown are at sea level, unless otherwise noted.

[f]The distance increases to 450 ft (135 m) for airports with elevations between 1345 and 6560 ft (410–2000 m) and to 500 ft (150 m) for airports with elevations above 6560 ft (2000 m).

[g]For ADV V and VI, approaches with visibility less than 0.5 mile (0.8 km), the separation distance increases to 500 ft (152 m) and 550 ft (168 m), respectively, plus any required adjustment for elevation.

Source: FAA, 2012.

TABLE 9.7 Separation Requirements between Runways and Other Facilities or Parts of the Airfield for Approach Categories C, D, and E

The most significant differences between Table 9.7 and the corresponding ICAO specifications concern the distances between runway centerlines and taxiway centerlines for instrument runways with approach minima lower than 0.75-mile (1.2-km) visibility.[17] For runways handling aircraft with ICAO code letters C, D, E, and F (FAA Groups III–VI), the ICAO standards call for 168, 176, 182.5, and 190 m of separation, respectively—compared to the 122, 122, 122, and 182 m shown for Groups III through VI in Table 9.7. For the other separation distances listed in Table 9.7, the ICAO standards are also generally more conservative.[18] For example, for the distance between runway centerlines and holdlines (as well as holding bays, see Sec. 9.7) at airports designed to ICAO code number 4, the ICAO calls for distances of 90 m for precision approach runways serving ICAO code letter C, D, and E aircraft and of 107.5 m for code letter F. The FAA counterparts of these three values in Table 9.7 are 76 m for ICAO code letters C and D (FAA Groups III and IV), and 85 m for E and F (FAA Groups V and VI).[19]

Vertical Profile

The FAA and the ICAO also have strict standards for runway and taxiway surface gradients and lines of sight. The brief review here is limited to FAA standards for runways designed to serve aircraft in approach categories C, D, and E of the FAA's ARC. The relevant longitudinal and transverse gradient standards are shown in Figs. 9.14 and 9.15, respectively. Analogous but different standards apply to approach categories A and B.

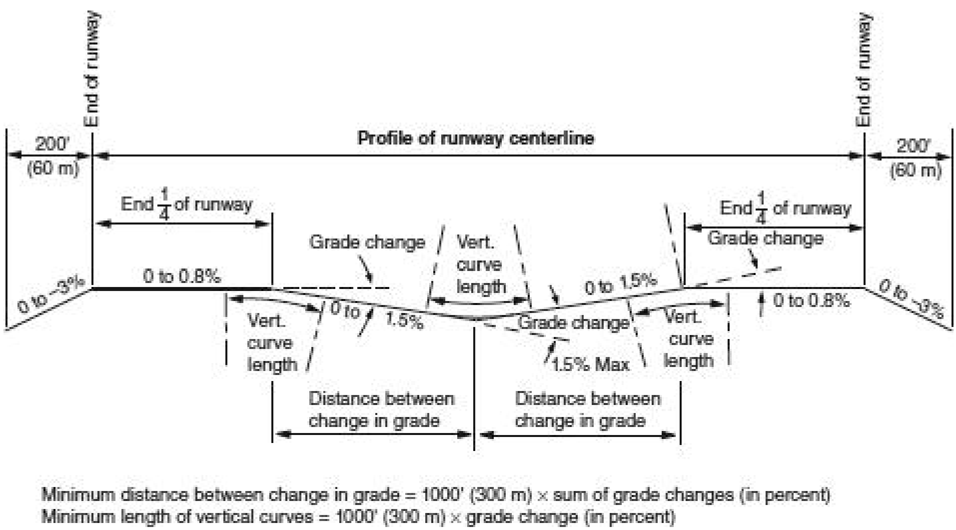


FIGURE 9.14 Longitudinal grade limitations for FAA aircraft approach categories C, D, and E. (*Source*: FAA, 2012.)

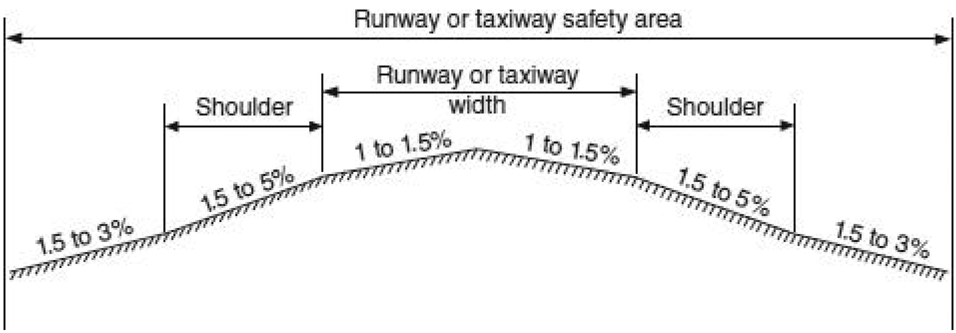


FIGURE 9.15 Transverse grade limitations for FAA aircraft approach categories C, D, and E. (*Source*: FAA, 2012.)

The *longitudinal* specifications apply restrictions on the runway grades allowed, on the changes in grades, and on the distances between changes. In general, it is desirable to keep the longitudinal grades, as well as the number and size of changes in grades, to a minimum. The maximum longitudinal grade allowed is ±1.5 percent. However, in the first and last quarter of the runway, the grade may not exceed ±0.8 percent. At the same time, the maximum allowable grade change is ±1.5 percent. Parabolic vertical curves should be used to effect longitudinal changes in grade. To ensure a smooth transition between grades, the length of the vertical curve should be at least 1000 ft (300 m) for each 1 percent of change in grade. Moreover, to avoid frequent changes in grade, the minimum allowable distance between successive changes in grade is 1000 ft (300 m) multiplied by the sum, in percent, of the grade changes, associated with the two vertical curves. Thus, the distance between a 1 percent grade change and a 0.5 percent grade change must be at least 1500 ft. The longitudinal grades applied to a runway should also be applied to the entire RSA. Figure 9.14 also shows the longitudinal grade standards for the parts of the RSA that extend beyond the runway ends.

Figure 9.15 shows the maximum and minimum transverse grades for runways, taxiways, and stopways, including the runway shoulders and the RSA. In general, the smallest transverse grades that satisfy local drainage requirements should be used.

To enhance safety, line-of-sight standards have also been set for runways. An acceptable runway profile permits any two points 5 ft (1.5 m) above the runway centerline to be mutually visible for the entire runway length. However, if the runway has a full-length parallel taxiway, this requirement is reduced to the two points being mutually visible for *one-half* the runway length. A clear line of sight between the ends of intersecting runways is recommended, as well. For detailed guidance on this point, see FAA (2012). The FAA has no line-of-sight requirements for taxiways.[20] However, the sight distance along a runway from an intersecting taxiway must be sufficient to allow a taxiing aircraft to enter safely or cross the runway.

9.7 Taxiways

Figures 9.3 through 9.8 show that taxiway systems at major airports can be extensive, complex in configuration, and thus costly to construct and maintain. The Munich taxiway system, for example, is approximately 30 km (~19 miles) long, not including taxilanes in apron areas. In contrast, the combined length of the two runways is 8 km (5 miles).

All too often in airfield design, the taxiway system is almost an after-thought. Typically, the positioning and configuration of the runways and of the landside facilities, including the ground access roadways and other guideways, are fixed first. The taxiway system is then designed to provide connections between the runways and the apron areas near and around passenger and cargo buildings, maintenance areas, etc. This can be a costly approach in terms of both fixed and operating costs. It may, for example, lead to the construction of an unnecessarily large number of expensive (segments of) taxiways on bridges, so that automobiles on access roads can reach the passenger buildings. A poorly designed taxiway system may also require aircraft to take circuitous routes between the runways and the apron stands, increasing airline operating costs and wasting time. Airports with landside facilities located at midfield are particularly prone to such problems, if not designed in a manner that integrates the planning of the taxiway system into the overall process.

To appreciate the magnitude of the economic quantities involved, consider the taxiing time that aircraft experience in traveling between aprons and runways at a busy airport handling, for example, 300,000 movements and 25 million passengers in a year. A design improvement that reduces taxiing time by 1 minute, on average, translates to savings of 5000 aircraft-hours and over 400,000 passenger-hours per year. If the average direct operating cost for taxiing aircraft at this airport is $2000 per aircraft-hour, saving a minute of average taxi times saves about $10 million per year in direct airline costs alone—the equivalent of about $100 million in capital investment.[21] The value of this saving in taxi time could more than double if savings in passenger time at $40 per hour saved (typical for such analyses in the United States) were also taken into consideration. In practice, airport planners unfortunately often fail to take a systems view of the design and, in this case, focus on the capital costs incurred by the airport and not the system benefits received by the users, the airlines and passengers. Airports need to recognize, however, their potential to recover costs through user charges, and to strive to maximize the overall productivity of the system. Section 14.3 picks up this discussion.

Another common mistake is overbuilding in the early stages. Runways that are not used intensively during the early years of an airport's operation can usually be adequately supported, depending on the airport's geometry, by a single taxiway running parallel to the entire length of the runway ("full-length taxiway"). Even when a second parallel taxiway may be needed to ensure smooth circulation of airport surface traffic under all circumstances, that second taxiway need not be full-length. It may be sufficient in most cases to build that second parallel taxiway for only part of the length of the runway. Unfortunately, airport designers often tend to emphasize symmetry. Almost reflexively sometimes, they design taxiway systems with two full-length parallel taxiways per runway and then build the entire taxiway system in a single step. The key to economic efficiency is to design carefully a flexible taxiway system and then build it up in phases over the years as traffic grows.

Dimensional Specifications and Separations

National and international design guidelines for taxiways include recommendations for their width; curves; minimum separation distances between taxiways and parallel taxiways, taxiways, and objects; longitudinal slope changes; sight distances; and transverse slopes (FAA, 2012; ICAO, 2005, 2009). Both the FAA and the ICAO make special reference to taxiways located around and within apron areas. These taxiways are divided into two types: *apron taxiways*, which may surround aprons or may provide a route across them; and *taxilanes* (or *aircraft stand taxilanes*), which are corridors in an apron reserved for circulation of aircraft and for providing access to the aircraft stands only. Standards and recommendations for apron taxiways are essentially the same as for regular taxiways. However, some separation requirements for taxilanes are less conservative than for taxiways. This is because aircraft move more slowly in apron areas and thus their paths typically adhere more closely to the centerlines of taxilanes than to those of regular taxiways.

In a break with past practices, the FAA decided in 2012 to base certain taxiway design specifications on a new classification called the *taxiway design group* (TDG) rather than on its Group I through VI classification of aircraft. The TDG of an aircraft is determined by the width of the overall main gear (OMG in Table 9.3) and the distance between the cockpit and the aircraft's main gear (FAA, 2012). These two dimensions provide an indication of the maneuverability and turning radii of the aircraft on the taxiways. The FAA thus now classifies aircraft types into TDG Classes 1 through 7. It uses this classification to specify the width of taxiways and taxiway shoulders, as well as the recommended distance between the centerline of a taxiway or taxilane and the centerline of a parallel taxiway and taxilane (Table 9.8). The *edge safety margin* is defined as the minimum acceptable distance between the outside of the airplane wheels and the pavement edge. Note that the width of the taxiway is determined by adding the taxiway edge safety margins to the OMG dimension assumed for each TDG class of aircraft. For example, TDG 7 assumes that the OMG dimension of its most demanding aircraft is 15 m. The recommended width of 25 m for TDG 7 taxiways is thus the sum of its OMG and 10 m (twice the relevant taxiway edge safety margin) (see Table 9.8).

	Taxiway Design Group (TDG)						
	1	2	3	4	5	6	7
Dimensional Standards							
Width	25 ft 7.5 m	35 ft 10.5 m	50 ft 15 m	50 ft 15 m	75 ft 23 m	75 ft 23 m	82 ft 25 m
Edge safety margin	5 ft 1.5 m	7.5 ft 2 m	10 ft 3 m	10 ft 3 m	15 ft 5 m	15 ft 5 m	15 ft 5 m
Shoulder width	10 ft 3 m	10 ft 3 m	20 ft 6 m	20 ft 6 m	25 ft 7.5 m	35 ft 10.5 m	40 ft 12 m
Separation Distances							
Taxiway centerline to parallel taxiway or taxilane centerline	69 ft 21 m	69 ft 21 m	160 ft 49 m	160 ft 49 m	240 ft 73 m	350 ft 107 m	350 ft 107 m

Note: All the dimensions for TDG 3 and 4 are identical.

Source: FAA, 2012.

TABLE 9.8 Taxiway design standards based on TDG

This new classification may seem confusing. However, in practical terms, there is a close correspondence between the TDG and ADG classification of aircraft types. Specifically, most (but not all) Group III aircraft are in TDG Class 3 or 4, Group IV and a few Group III aircraft in TDG 5, and Groups V and VI aircraft in TDG 6 and 7. Table 9.3 indicates the TDG classification of the most common current types of aircraft.

Dimensional and separation specifications that do not depend on the dimensions of aircraft undercarriage, but on wingspan and tail height are listed in Table 9.9. Note that these dimensions are analogous to the dimensions defined for runways. For instance, the *taxiway safety area*, by analogy to the RSA, includes, but is not limited to, the taxiway and taxiway shoulders and it must be cleared and graded, drained, and free of objects.

	Airplane Design Group					
	I	II	III	IV	V	VI
Dimensional Standards						
Taxiway safety area (TSA) width	49 ft 15 m	79 ft 24 m	118 ft 36 m	171 ft 52 m	214 ft 65 m	262 ft 80 m
Taxiway object-free area width	89 ft 27 m	131 ft 40 m	186 ft 57 m	259 ft 79 m	320 ft 98 m	386 ft 118 m
Taxilane object-free area width	79 ft 24 m	115 ft 35 m	162 ft 49 m	225 ft 69 m	276 ft 84 m	334 ft 102 m
Separation Distances						
Taxiway centerline to:						
Parallel taxiway or taxilane centerline*	69 ft 21 m	105 ft 32 m	158 ft 48 m	215 ft 66 m	267 ft 81 m	324 ft 99 m
Fixed or movable object	44.5 ft 14 m	65.5 ft 20 m	97 ft 30 m	129.5 ft 39 m	160 ft 49 m	193 ft 59 m
Taxilane centerline to:						
Parallel taxilane centerline*	64 ft 20 m	97 ft 30 m	146 ft 45 m	195 ft 59 m	245 ft 75 m	298 ft 91 m
Fixed or movable object	39.5 ft 12 m	57.5 ft 18 m	84 ft 26 m	112.5 ft 34 m	138 ft 42 m	167 ft 51 m
Wingtip clearance						
Taxiway wingtip clearance	20 ft 6 m	26 ft 8 m	35 ft 10.5 m	44 ft 13 m	53 ft 16 m	62 ft 19 m
Taxilane wingtip clearance	15 ft 5 m	18 ft 5 m	23 ft 7 m	27 ft 8 m	31 ft 9 m	36 ft 11 m

*These values are based on wingtip clearances; if 180° turns between parallel taxiways are needed; use this dimension or the dimension specified in Table 9.8, whichever is larger.

Source: FAA, 2012.

TABLE 9.9 Taxiway Design Standards Based on ADG

The ICAO has very similar dimensional standards with the FAA for taxiways [see Chap. 3 of ICAO (2009) and, especially, ICAO (2005)]. For example, for airports serving large aircraft, the principal difference with the FAA standards is that, for its code letters D, E, and F, the ICAO recommends an edge safety margin (*clearance distance*) of 4.5 m, by comparison to the FAA's 5 m for TDG Classes 5, 6, and 7, which comprise all ICAO code D, E, and F aircraft.

Note that the FAA separation distances in the lower half of Table 9.9 are different for taxiways and for taxilanes. These separation distances are based on a set of simple formulas that may be used to take into consideration special local conditions. Thus the separation between a taxiway centerline and a parallel taxiway or taxilane centerline is calculated as 1.2 times the maximum wingspan of the airplane design group plus 10 ft or 3 m. For Group VI, with a 262-ft (80-m) maximum wingspan (see Table 9.3*b*), this gives 324 ft (99 m) of separation, as shown in Table 9.9. Note that this leaves a margin of 62 ft or 19 m (see "wingtip clearance" at the bottom of Table 9.9) for the total possible deviation from the centerlines of the parallel taxiways when two Group VI airplanes are moving (usually in opposite directions) on the taxiways. The margin is, of course, larger when one or both of the aircraft moving on this pair of taxiways are smaller than Group VI.

For the separation between the centerlines of a taxilane and a parallel taxilane, 1.1 times the maximum wingspan (instead of 1.2) plus 10 ft or 3 m is used. Similarly, a taxiway centerline should be separated by 0.7 times the maximum wingspan plus 10 ft (3 m) from a fixed or movable object, whereas 0.6, instead of 0.7, is used in the case of a taxilane. The ICAO uses somewhat different formulas than the FAA to determine a set of separation standards [see Table 3.1 in ICAO (2009)], which are very similar to the ones shown in the lower part of Table 9.9—in most cases within 1.5 m of the corresponding values. ICAO, however, does not draw a distinction between taxiways and taxilanes, when it comes to separation from a parallel taxilane.

Special Cases

A taxiway system includes segments and special-purpose elements that require more detailed design considerations, because of their particular characteristics. Included in this category are the following:

- Curved segments of taxiways
- Taxiway intersections or junctions
- Taxiways on bridges
- Exit taxiways, including high-speed (or rapid or acute-angle) exit taxiways
- Holding bays and bypass taxiways

For *curved segments of taxiways* and for *intersections and junctions* with runways, aprons, and other taxiways, dimensional specifications are driven by the edge safety margins (clearance distance, in ICAO terminology) shown in Table 9.8, as well as by the *wheelbase* of the design aircraft, which is captured by its TDG number. The wheelbase is the distance from the nose gear of the aircraft to the geometric center of the main gear. Additional taxiway width or taxiway fillets[22] are provided to ensure that the applicable edge safety margins are maintained, assuming the aircraft's nose gear stays on centerline markings (Fig. 9.16). Detailed guidance on the design of these curved segments can be found in ICAO (2005) and FAA (2012).

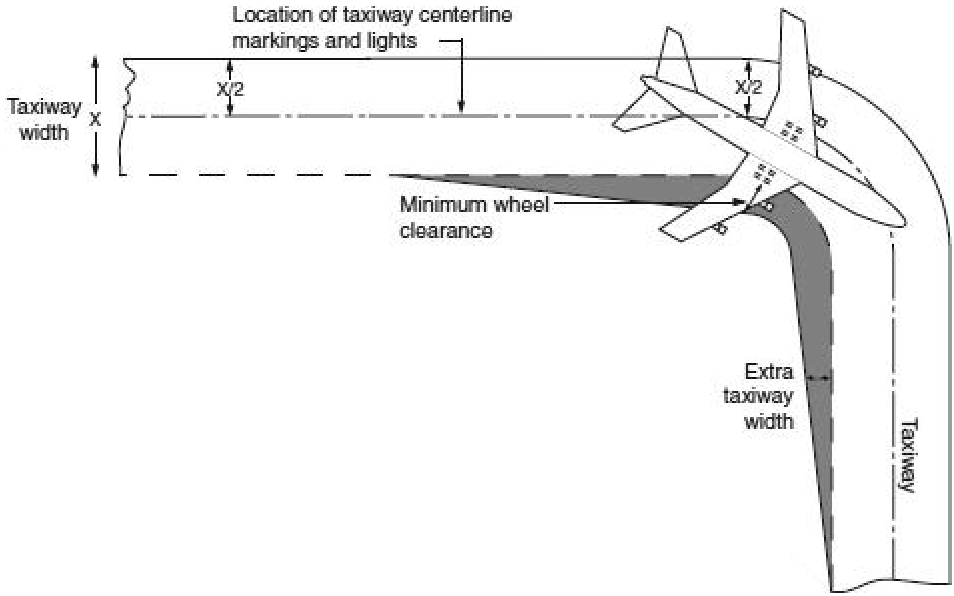


FIGURE 9.16 A curved segment of a taxiway; additional taxiway width may be needed to ensure that taxiway edge safety margins are preserved. (*Source*: ICAO, 2009)

Taxiways on bridges are becoming increasingly common, especially at airports where the landside facilities occupy the central portion of the airport[23] (Sec. 9.3). Because of the high cost of these structures, airport designers try to minimize their number. This often leads to difficult tradeoffs between construction (fixed) costs, on the one hand, and taxiing and operations (variable) costs, on the other, as noted earlier in this section. The ICAO and the FAA guidelines regarding taxiway bridges are very similar. The width of the part of the taxiway that lies on the bridge should be at least as large as the width of the taxiway safety area (FAA) or the graded part of the taxiway strip (ICAO). Where this may not be possible, edge protection should be provided on the bridge, as well as engine blast protection for vehicles or people crossing under the bridge. There should also be adequate space on the bridge for access by firefighting and rescue equipment from both sides of the bridge and for aircraft evacuation slides. The bridges should be on straight segments of taxiways and away from high-speed exit taxiways or other special-purpose parts of the airfield (FAA, 2012; ICAO, 2005, 2009).

Exit taxiways provide egress paths from runways for arriving aircraft. The centerlines of conventional exits form a 90° angle with the centerlines of the runway (Fig. 9.17). *High-speed exits* (or *rapid exit taxiways* or *acute-angle exit taxiways* in ICAO and FAA terms, respectively) are those whose centerlines form an angle significantly less than 90° with the runway centerline. With a 30° high-speed exit, aircraft can theoretically initiate a turn while traveling as fast as 90 km/h (56 mi/h) or more. However, in practice, most pilots are more conservative. Figure 9.18 shows a typical geometric design of a high-speed exit. An angle of 30° between the centerlines of the high-speed exit taxiway and the runway is common. The ICAO specifies that the radius of the turnoff curve of a high-speed exit (see Fig. 9.18) should be at least 550 m to enable exit speeds of 93 km/h (~58 mi/h) under wet runway surface conditions for ICAO code number 3 or 4 aircraft (ICAO, 2009). For a high-speed exit of this type, the FAA requires at least 180 m (600 ft) of separation between the centerline of the runway and that of the parallel taxiway (FAA, 2012). Note that the FAA's runway-to-taxiway separation distances shown in Table 9.7 apply to the case where conventional, right-angle exit taxiways are used. These distances may have to be increased to accommodate high-speed exits.

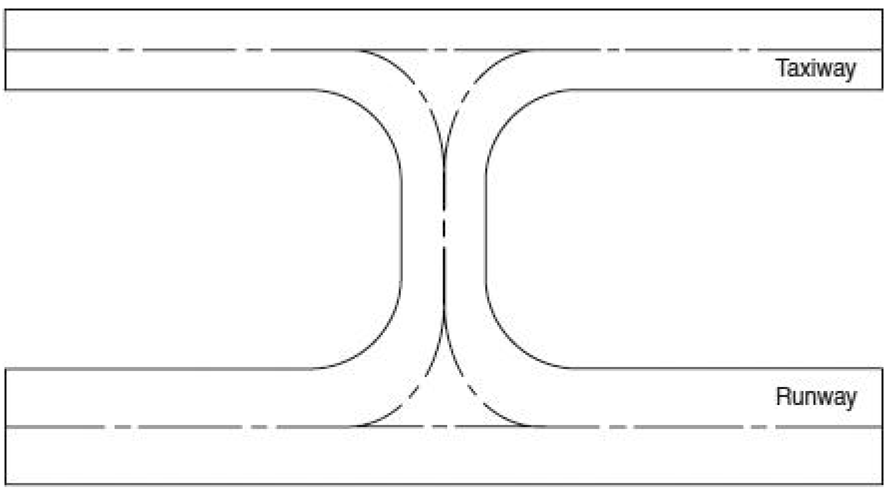


FIGURE 9.17 Conventional, right-angle exit taxiway. (*Source*: FAA, 2012.)

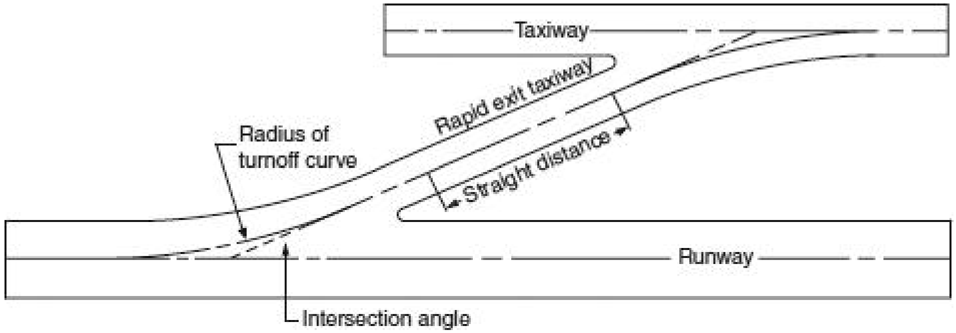


FIGURE 9.18 Acute-angled exit taxiway. (*Source*: FAA, 2012.)

The location of exit taxiways plays a significant role in determining runway occupancy times and, under certain conditions (see Chap. 10), runway capacity. To take an extreme example, average runway occupancy times will obviously be much longer on a long runway that has a single exit at its far end than on an equally long runway that has several well-located exits. There is, however, a point of diminishing returns, after which little is gained by constructing additional exits. High-speed exits do contribute to reducing runway occupancy times. Given, however, their higher cost and diminishing returns as their number increases, it is difficult to make a case for constructing more than, at most, three high-speed exits for each direction of operation of a runway, if other conventional exits at 90° are provided. Moreover, this can be justified only for intensively utilized runways, with more than 30 peak-period movements per hour.

Many studies have been performed on the optimal location of exit taxiways, beginning in the 1950s. However, few general statements can be made in this regard, because of the many local factors that play a role in the exit selection process. Included among those are the mix of aircraft types using the runway, pilot technique, the condition of the runway surface (wet or dry), and the location of aircraft stands relative to the runway. On this last point, it has been observed that pilots often aim for the exit(s) that will most facilitate access to the stands and adjust the deceleration of the aircraft accordingly. As an approximate rule, a long runway that may be used in either direction at a busy airport should have exit taxiways at its two ends and at approximately 450 m (~1500 ft) intervals after its middle in each direction, up to a distance of about 600 m (~2000 ft) from its ends. Thus, a long 3500-m (~11,500-ft) runway should have seven or eight exit taxiways, including the two at its ends. The exit configuration of the two Munich runways, 08L/26R and 08R/26L, provides a good example (see Fig. 9.5).

A common mistake in airport design is the placement of high-speed exits on runways (or in runway directions) that will seldom be used for arrivals. It is often entirely appropriate to equip one direction of operation of a given runway with high-speed exits and the other direction with only conventional exits. An example is Runway 01L/19R at Amsterdam/Schiphol (Fig. 9.19). This asymmetry is because Runway 01L is almost always used for departures only, whereas 19R for arrivals only. In general, it is useful to remember that high-speed exits offer essentially no capacity benefits at runways used only for departures, some capacity benefits at runways used only for arrivals, and significant capacity benefits, under some movement-sequencing strategies, at runways used in a mixed operations mode (see Chap. 10).

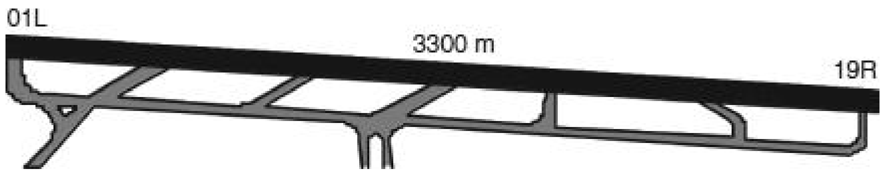


FIGURE 9.19 Runway 01L-19R and exit taxiways at Amsterdam/Schiphol.

Holding bays are areas adjacent to taxiways where aircraft may be held temporarily without impeding the circulation of other taxiing aircraft. These are usually placed close to runways, so they can provide a waiting area for aircraft that are not yet ready for takeoff, and allow air traffic controllers to sequence departing aircraft in a particular way, if desired. Holding bays take several different geometric shapes (ICAO, 2005). Their location should keep aircraft out of obstacle-free zones and RSAs, and prevent interference with instrument landing systems and other navigational aids. *Dual taxiways* provide a second taxiway segment that makes it possible to bypass aircraft near critical points of the airfield, typically runway ends. Figure 9.5 shows the holding bays and the dual taxiway pairs next to the four runway ends at Munich.

9.8 Aprons

Aprons provide the interface between airside and landside facilities at airports. Depending on the type of aircraft stands they contain, they can be classified as passenger building aprons, cargo building aprons, long-term parking aprons, service and hangar aprons, or general aviation aprons. Long-term parking aprons, if any, are located at remote areas of the airfield and can be used by aircraft parked for periods ranging from overnight to months (due to temporary grounding). General aviation aprons may be separated into areas for "itinerant" aircraft passing through and areas for aircraft based at the airport in question. These latter aprons are also known as "tie-downs."

Passenger building stands, on which this section focuses, are either *contact* or *remote*, depending on their location relative to the buildings. Figure 9.20 shows schematically generic configurations of passenger building aprons. The configuration of the aprons clearly depends on the configuration of the passenger buildings. Chapters 14 and 15 provide a more detailed discussion of the advantages and disadvantages of the different configurations. Only brief related observations are made here.

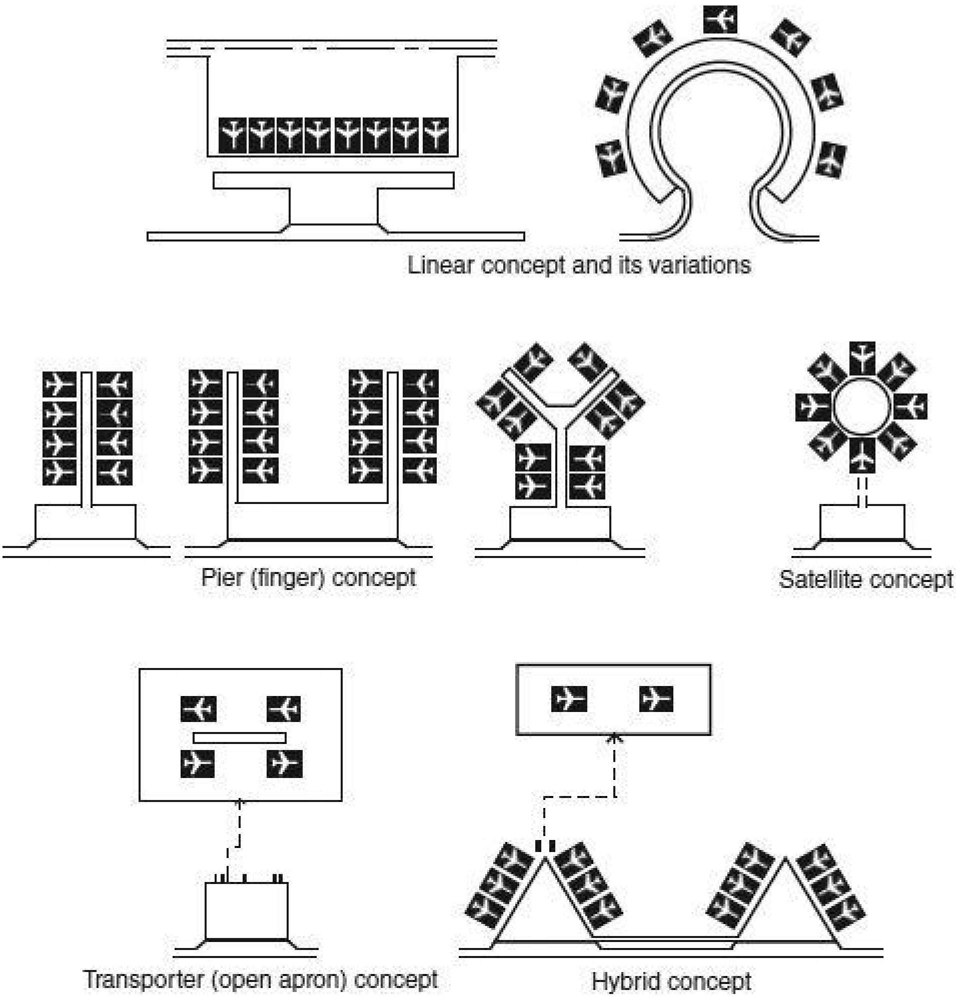


FIGURE 9.20 Standard and hybrid configurations of passenger building and aprons. (*Source*: ICAO, 2005.)

The objective of apron design is to develop a configuration that respects all safety-related requirements, while maximizing efficiency for aircraft moving in and out of the apron and providing flexibility. On the safety side, ICAO requires (ICAO, 2005, 2009) that the following minimum clearances be provided at an aircraft stand between any part of the aircraft and any adjacent building, aircraft on another stand, or other object, except for vehicles and equipment servicing the aircraft:

- 3 m for code letters A and B
- 4.5 m for code letter C
- 7.5 m for code letters D, E, and F

However, these may be reduced in the specific case of the clearance between the nose of the aircraft and the passenger building (including passenger loading bridges).

Of special importance are the expandability of the apron area and its ability to accommodate the full range of aircraft using the airport. The latter poses a particularly difficult challenge. At one extreme, the dimensions of the stands in an apron could be selected so all are large enough to accommodate all potential aircraft sizes (e.g., Group V or smaller, for airports not expected to serve Group VI aircraft). This, however, would be extremely inefficient in most cases. At the opposite end, the mix of stand sizes could be identical to the mix of aircraft that currently park at the airport during peak demand periods. The obvious disadvantage of this second approach is that it offers little flexibility if the mix changes in the direction of a greater presence of larger aircraft. The proper compromise is to provide a mix of stand sizes that is biased toward having a higher fraction of larger stands than is warranted by the current aircraft mix. However, developing the specifics of this approach is a complex task that depends on much more than just apron-related considerations. For example, if the airport relies heavily on contact stands, the choice of the mix of stand sizes will clearly affect directly the dimensions of the passenger building. In this case, any bias toward a high percentage of larger stands must be tempered by the associated very large capital costs. Chapters 14 and 15 address this question further.

The efficiency of operations associated with different apron designs involves another set of complex questions. For example, tradeoffs must be made between ease of movement of aircraft versus passenger convenience and passenger building operating costs. To take an obvious example, the transporter (or "open apron") concept shown at the lower left of Fig. 9.20 greatly facilitates the movement of aircraft—as they can park in less space-constrained areas and do not usually have to be pushed back from their stands. However, this concept also requires transporting passengers to/from the aircraft, with buses or special-purpose vehicles, implying higher variable costs and necessitating "closing" of acceptance of passengers for departing flights 30 minutes or more prior to departure time. As a second example, the linear concept (top of Fig. 9.20) also facilitates aircraft movements, but it makes inefficient use of the frontage of the passenger building, as aircraft can park on only one side of the building. It also requires considerable duplication of landside services and, thus, leads to higher operating costs (see Chaps. 14 and 15). For airports with a high volume of connecting passengers, a concept that has emerged as a good solution to this difficult design optimization problem is the midfield satellite terminal (see Chap. 14). The best-known—and the first—example of this particular design is Atlanta (see Fig. 9.7).

An important problem that occurs frequently at major airports is caused by the blocking of aircraft movements in the taxilanes that serve sets of stands. This problem occurs primarily around pier/finger-shaped buildings and may reduce significantly the capacity of the apron. A typical example is the one involving parallel piers (see Fig. 9.20, left middle row). Ideally, the distance between the piers should be sufficient to allow for two parallel taxilanes, one serving incoming and the other outgoing traffic (Fig. 9.21, bottom row). Note that the distances shown in Fig. 9.21 are based on the FAA's formulas for separations between parallel taxilane centerlines and between a taxilane centerline and an object described in Sec. 9.7. If the distance between the piers is not sufficient for a dual taxilane, a single taxilane must serve all the stands in the apron.

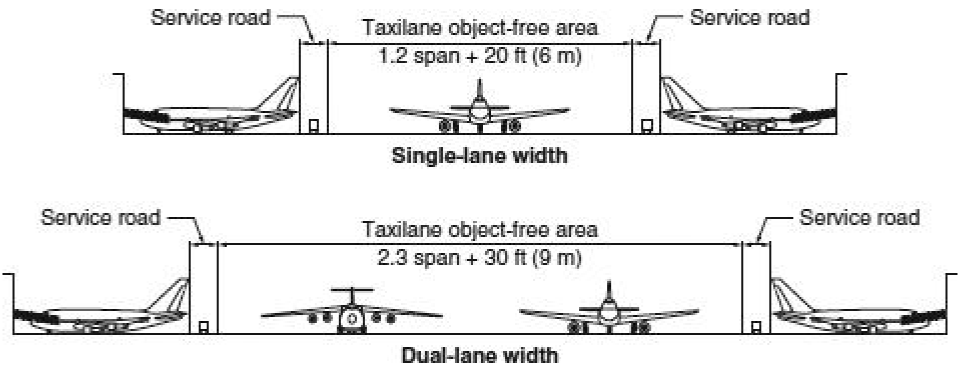


FIGURE 9.21 Single versus dual taxilanes and wingtip clearances at an apron. (*Source*: FAA, 1989.)

With a parallel pier arrangement, movement into and out of many of the stands may be blocked whenever an aircraft is either pushing back from a stand or moving on the taxilane in either direction.[24] Serious delays to traffic may result. A simple rule of thumb is that such delays will indeed occur if there are *more than four to six stands* on each side of the single taxilane. The same problem arises with numerous variants of the single-taxilane geometric configuration. As an example, Fig. 9.22 shows the narrow entrance and exit to an apron area between passenger buildings B and C at Boston/Logan. This apron area is shaped like a horseshoe. Any aircraft occupying the narrow space between the two passenger buildings effectively blocks circulation into or out of the "horseshoe."

FIGURE 9.22 The "horseshoe" apron between passenger buildings B and C at Boston/ Logan and the entrance to that area. (*Source*: Massport.)

9.9 Physical Obstacles

To ensure the safety of operations in the airspace in the immediate vicinity of airports, both the ICAO and the FAA have established a set of obstacle limitation surfaces that define the limits to which objects may project into airspace. These surfaces protect approaches to runways, takeoffs, and missed approaches ("balked landings") from obstructions. Objects that penetrate these surfaces are considered obstacles to air navigation and are removed, when possible. An obstacle can be any fixed or mobile object, including terrain, natural objects such as trees, and man-made ones such as antennas or buildings. Several types of airport charts display these obstacles for each airport. The International Air Transport Association, among others, maintains an online, current *Airport and Obstacle Database* with information on several thousand airports worldwide (IATA, current).

Obstacle limitation surfaces are very important to aviation safety. They provide guidance for zoning restrictions regarding the height of buildings, antennas, and other structures near airports. They may also play a major role in determining the construction costs of new airports when these are located in difficult terrain. More than 50 million m^3 of soil, rocks, and other materials had to be removed and transported from hills around the site of the new airport at Athens in order to clear obstacles to air navigation per Annex 14 of ICAO.

The obstacle limitation surfaces defined by the ICAO are the following:

- Conical surface
- Inner horizontal surface
- Approach surface
- Transitional surfaces
- Inner approach surface and inner transitional surfaces
- Balked landing surface
- Takeoff climb surface

These surfaces are shown schematically in planar view at the top part of Fig. 9.23 and in side view for two different sections at the lower part (ICAO, 2009). Note in the planar view what Secs. A-A and B-B correspond to. The surfaces are described briefly in the next paragraphs. The relevant dimensions are given in Table 9.10 for nonprecision and precision approach runways and in Table 9.11 for departure runways. The FAA has established a very similar set of surfaces, called *imaginary surfaces*, which are defined in Part 77, paragraph 25, of the Federal Aviation Regulations (FAR) (FAA, current). Some differences between the ICAO and FAA specifications are identified at the end of this section. FAR Part 77 also describes the requirements for adequate advance notice to the FAA of any proposed construction that may affect navigable airspace (Subpart B), as well as the procedure for evaluating, reviewing, and determining the need for remedial action (Subparts D and E).

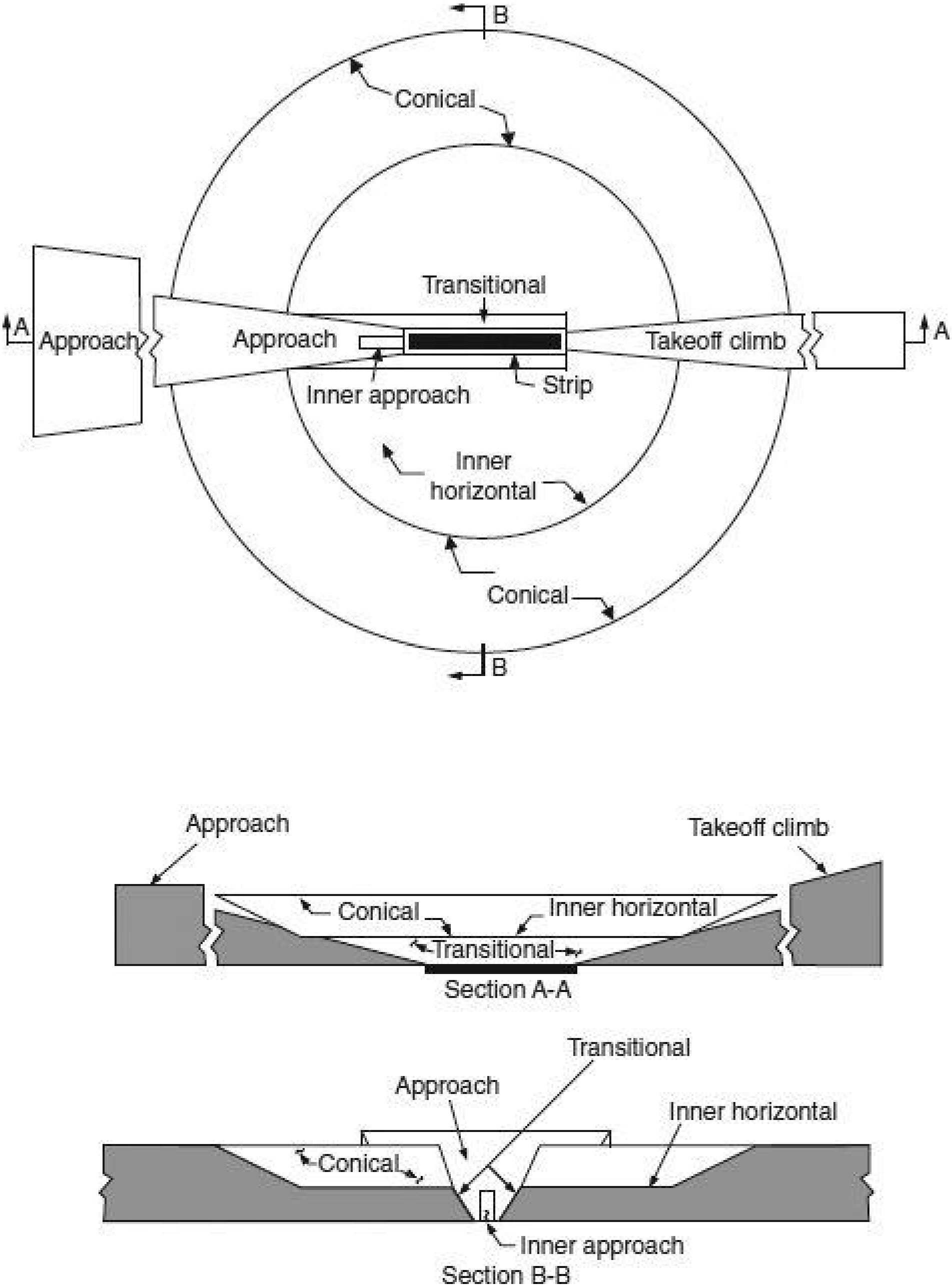


FIGURE 9.23 Obstacle limitation surfaces as defined by the ICAO. (*Source*: ICAO, 2009.)

Surface and Dimensions	Nonprecision Approach Runway			Precision Approach Runway		
				Cat I		Cat II or III
	Code Number			Code Number		
	1, 2	3	4	1, 2	3, 4	3, 4
Conical						
Slope	5%	5%	5%	5%	5%	5%
Height	60 m	75 m	100 m	60 m	100 m	100 m
Inner Horizontal						
Height	45 m	45 m	45 m	45 m	45 m	45 m
Radius	3.5 km	4 km	4 km	3.5 km	4 km	4 km
Inner Approach						
Width	—	—	—	90 m	120 m	120 m
Distance from threshold	—	—	—	60 m	60 m	60 m
Length	—	—	—	900 m	900 m	900 m
Slope	—	—	—	2.5%	2%	2%
Approach						
Length of inner edge	150 m	300 m	300 m	150 m	300 m	300 m
Distance from threshold	60 m	60 m	60 m	60 m	60 m	60 m
Divergence (each side)	15%	15%	15%	15%	15%	15%
First section of approach						
Length	2.5 km	3 km	3 km	3 km	3 km	3 km
Slope	3.33%	2%	2%	2.5%	2%	2%
Second section of approach						
Length	—	3.6 km	3.6 km	12 km	3.6 km	3.6 km
Slope	—	2.5%	2.5%	3%	2.5%	2.5%
Horizontal section of approach						
Length	—	8.4 km	8.4 km	—	8.4 km	8.4 km
Total length	—	15 km	15 km	15 km	15 km	15 km
Transitional						
Slope	20%	14.3%	14.3%	14.3%	14.3%	14.3%
Inner Transitional						
Slope	—	—	—	40%	33.3%	33.3%
Balked Landing Surface						
Length of inner edge	—	—	—	90 m	120 m	120 m
Distance from threshold	—	—	—		1.8 km	1.8 km
Divergence (each side)	—	—	—	10%	10%	10%
Slope	—	—	—	4%	3.33%	3.33%

Source: ICAO, 2009.

TABLE 9.10 Obstacle Limitation Surfaces, Approaches

	Code Number		
	1	2	3 or 4
Length of inner edge	60 m	80 m	180 m
Distance from runway end	30 m	60 m	60 m
Divergence (each side)	10%	10%	12.5%
Final width	380 m	580 m	1,200 m 1,800 m
Length	1,600 m	2,500 m	15,000 m
Slope	5%	4%	2%

Source: ICAO, 2009.

TABLE 9.11 Obstacle Limitation Surfaces, Takeoff Climb

The *inner horizontal surface*, as defined by the ICAO, is a horizontal plane above an airport and its environs. It should normally be a circle whose radius depends on the type of runway(s) available (noninstrument approach, nonprecision approach, or precision approach). The circle is centered at the airport's reference point, the designated geographic location of the airport, located "near the initial or planned geometric center of the airport" and reported in degrees, minutes, and seconds (ICAO, 1999). The height of the inner horizontal surface is 45 m (150 ft in the FAA's FAR part 77) above the established elevation of the airport.

The *conical surface* projects upward at a slope of 5 percent from the periphery of the inner horizontal surface to a specified height above that surface. That height depends on the type of runway(s) available (see Table 9.10).

The *approach surface*, as the name suggests, protects the approach to the runway from obstructions. In planar view, it resembles a trapezoid inclined relative to the horizontal plane and may have a first section and a second section each with a different slope. For nonprecision or precision approaches by aircraft with code number 3 or 4, the approach surface has a horizontal section, as well (see far left side of Sec. A-A in Fig. 9.22). In these cases, a 3-km first section begins 60 m from the runway threshold with a slope of 2 percent. Then follows a 3.6-km second section with an upward slope of 2.5 percent and finally a horizontal section (at an elevation of 150 m) with a length of 8.4 km. The total length of these three sections is 15 km (~9 miles), although in some cases it may be even longer (ICAO, 2009).

The *transitional surfaces* on either side of the runway and of the approach surface incline upward and outward to the height of the inner horizontal surface (see planar and side views in Fig. 9.24). The elevation of any point at the lower edge of the approach surfaces is given by either the elevation of the approach surface at that point (when the point is along the side of the approach surface) or the elevation of the runway strip at that point (when the point is along the strip).

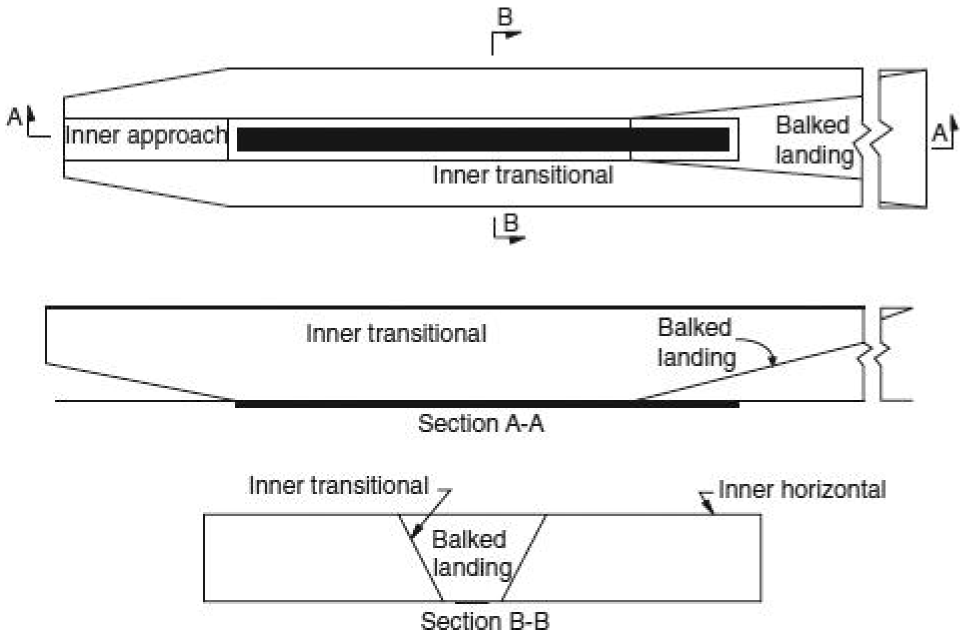


FIGURE 9.24 Inner approach, inner transitional, and balked landing surfaces. (*Source*: ICAO, 2009.)

For precision approach runways, an *inner approach surface* and associated *inner transitional surfaces* are also defined. The inner approach surface protects the part of the approach closest to the runway threshold, whereas the inner transitional surfaces are the controlling obstacle limitation surfaces for navigation aids, aircraft, and other vehicles that must be near the runway. The inner transitional surfaces are not to be penetrated except for frangible objects.

The *balked landing* (missed approach) *surface* is also defined only for precision approach runways. It provides an obstacle-free volume of airspace at the back end of the approach runway. Note in Table 9.10 that the balked landing surface begins at the end of the runway strip or at 1800 m from the approach threshold of the runway, whichever is closer. When a runway can be used for precision approaches from both directions, the approach surface provided at the opposite end of the runway, with a first-section slope of 2.5 or 2 percent, is more restrictive than the balked landing surface that requires a 4 or 3.33 percent slope. This means that the approach surface also provides, by default, a balked landing surface at both ends of a bidirectional precision approach runway.

Finally, the *takeoff climb surface* is an inclined plane intended to prevent obstructions to the paths of departing aircraft near a runway. Table 9.11 lists its specifications. For the ICAO code numbers 3 and 4, the final width of the surface is shown as 1200 or 1800 m. The larger number applies to cases where the flight track includes changes of heading greater than 15° for operations conducted in instrument meteorological conditions (IMC) or in visual meteorological conditions (VMC) at night. The 2 percent slope for code letters 3 and 4 may be reduced, if local conditions make this desirable. The ICAO recommends that, if no object at a runway currently reaches the 2 percent takeoff climb surface, new objects should be limited to a slope of 1.6 percent to protect future options.

As already mentioned, the standards for the imaginary surfaces established under FAR Part 77 for the FAA (Fig. 9.25) are similar in concept and in specified parameters to those of the ICAO. A detailed review of the FAA specifications and rationale can be found in ACRP (2010). One notable difference between FAA and ICAO standards is that the inner horizontal surface under FAR Part 77 is not a circle, but an oval (FAA, current). The oval consists of semicircles at the two ends of the runway connected with straight lines. Each semicircle is centered at one end of the runway and has a radius of 10,000 ft. Another significant difference is that the approach surface defined in FAR Part 77 is the same for landings and takeoffs. The inner edge of the approach surface for precision instrument runways is 1000 ft wide and its outer edge 16,000 ft wide. Its total length is 50,000 ft (roughly the same as the ICAO-specified 15 km), with a slope of 2 percent for the first 10,000 ft and 2.5 percent for the next 40,000 ft.

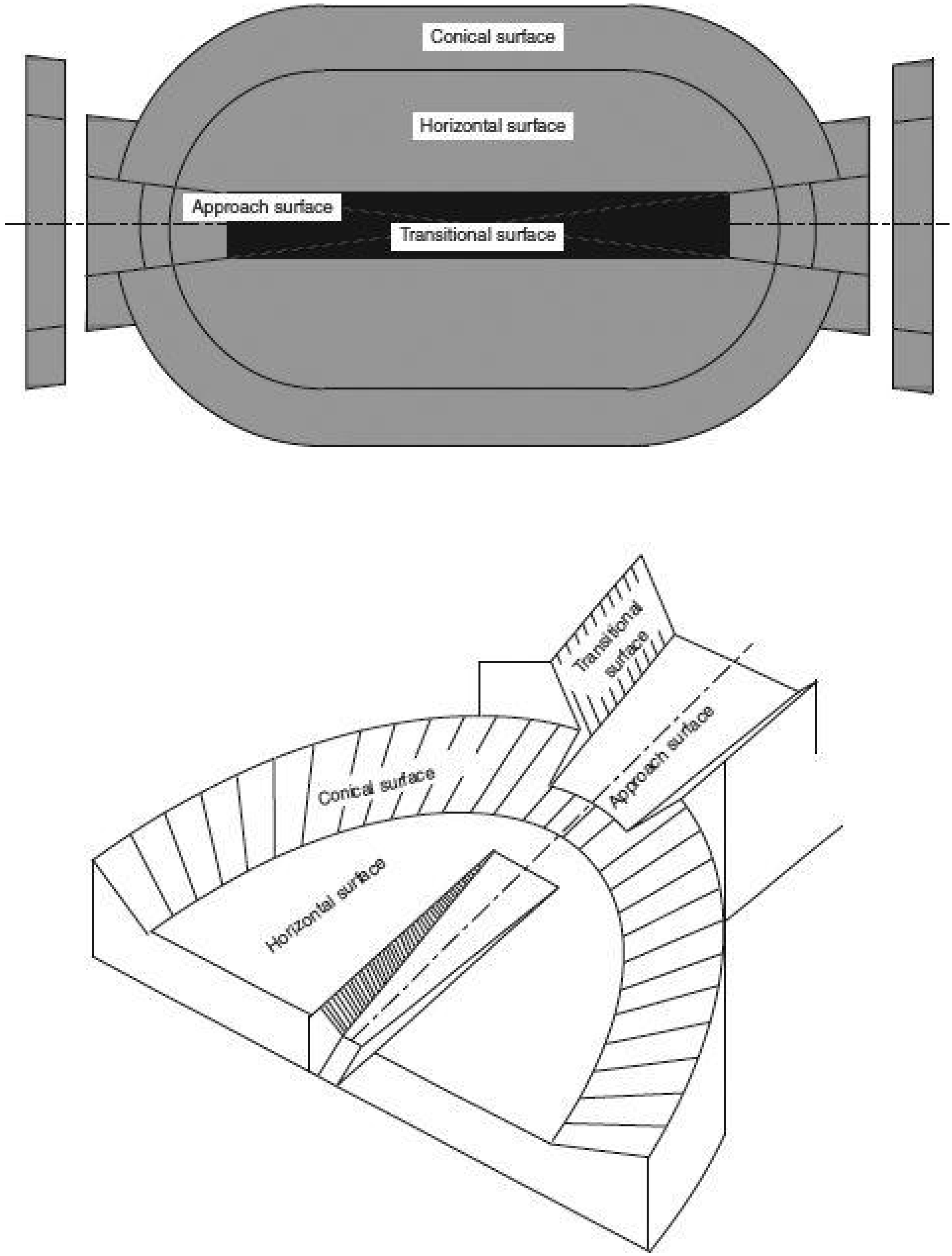


FIGURE 9.25 Imaginary surfaces as defined by the FAA. (*Source*: FAA, current.)

Exercises

9.1. Consider a single runway airport whose most demanding aircraft are in FAA approach category D and airplane design Group V. It has a long linear passenger building running parallel to the runway, with contact stands on the side of the building facing the runway. There are 14 contact stands: 5 can accommodate the Boeing 747-400, 4 the Boeing 767-200ER and other Group IV aircraft, and 6 the Boeing 737-800 and other Group III aircraft. The nine stands for the larger aircraft are at the central part of the building and the other six at the two end parts (three on each end). Arriving aircraft park nose-in and are pushed back on departure. A vehicle road that is 13 ft (4 m) wide lies behind the aircraft stands followed by a taxilane that runs in parallel to the full length of the face of the passenger building. The taxilane provides sufficient space for the aircraft pushback maneuver. Beyond this taxilane, two full-length taxiways run parallel to the entire length of the runway. They are located between the apron and the runway. Assume the runway is 11,200 ft (or 3400 m) long and is used in either direction, depending on prevailing winds, for both arrivals and departures. Aircraft not served at the contact stands are parked at remote stands in an apron area that does not affect the operation of the main apron next to the passenger building or the required distance between the main apron and the taxiways. Provide an approximate layout plan for this airfield showing its key dimensions. Make sure to indicate the minimum linear frontage of the passenger terminal, the approximate dimensions of the main apron, the width of runways and taxiways, and the separations between the centerlines of the taxiways and adjacent runways, taxiways, taxilanes and fixed objects. Do not go into details such as the design of curved segments and fillets but indicate high-speed versus conventional exit taxiways and their approximate location. Feel free to work in the set of units you are most comfortable with.

9.2. Consider again Exercise 1, but now assume that the runway system consists of two medium-spaced parallel runways, with a distance of 2500 ft (760 m) between their centerlines. The runway farthest from the passenger building is used for arrivals and the other for departures. Modify the layout plan of Exercise 1 to account for the new runway. Would you construct a new parallel taxiway between the two runways? What are the advantages and disadvantages of doing so?

9.3. Return to Exercise 1 and assume now that the most demanding aircraft is in FAA approach category D and airplane design Group VI. For simplicity assume that, as far as contact stands are concerned, the airport operator will want to convert the five stands that can now accommodate the Boeing 747-400 into stands that can accommodate Group VI aircraft, while still maintaining the four stands for Group IV aircraft, and the six for the Group

III aircraft. Indicate what other changes will have to be made to the approximate airport design that you developed in Exercise 1.

References

Airbus Industrie (2001) *A380 Airport Compatibility*, http://www.airbus.com/A380/.

ACI, Airports Council International (2011 and annual), *Airport Traffic Statistics*, http:www.airports.org.

ACRP, Airport Cooperative Research Program (2010), *Understanding Airspace, Objects, and Their Effects on Airports*, ACRP Report 38, prepared by Leigh Fisher, ASRC Research and Technology Solutions, Landrum and Brown, and the Ohio State University Dept. of Aviation, Washington, DC. www.trb.org

Boeing Company (current) *Boeing 767 Airport Performance Manual*, http://www.boeing.com/commercial/airports/acaps/767sec3.pdf.

FAA, Federal Aviation Administration (1983) *Airport Design*, Advisory Circular 150/5300-12, U.S. Government Printing Office, Washington, DC.

FAA, Federal Aviation Administration (2005) *Runway Length Requirements for Airport Design*, Advisory Circular 150/5325-4B, incorporates Change 1, U.S. Government Printing Office, Washington, DC.

FAA, Federal Aviation Administration (current) *FAR Part 77—Objects Affecting Navigable Airspace*, in Code of Federal Regulations (CFR) 14, Office of Federal Register, National Archives and Records Administration, Washington, DC.

FAA, Federal Aviation Administration (2012) *Airport Design*, Advisory Circular 150/5300-12, U.S. Government Printing Office, Washington, DC.

IATA, International Air Transport Association (current) *Airport and Obstacle Database*, http://www.aodb.iata.org/.

ICAO, International Civil Aviation Organization (1987) *Airport Planning Manual, Part 1, Master Planning*, 2d ed., reprinted April 1999, Doc. 9184, Montreal, Canada.

ICAO, International Civil Aviation Organization (2005) *Aerodrome Design Manual, Part 2—Taxiways, Aprons and Holding Bays*, 4th ed., Doc. 9157, ICAO, Montreal, Canada.

ICAO, International Civil Aviation Organization (2006) *Aerodrome Design Manual, Part 1—Runways*, 3d ed., Doc. 9157, ICAO, Montreal, Canada.

ICAO, International Civil Aviation Organization (2009) *Aerodromes, Annex 14 to the Convention on International Civil Aviation, Volume I: Aerodrome Design and Operations*, 5th ed., ICAO, Montreal, Canada.

ICAO, International Civil Aviation Organization (current-a) *Aerodrome Design Manual*, Doc. 9157, Montreal, Canada.

ICAO, International Civil Aviation Organization (current-b) *Airport Planning Manual*, Doc. 9184, Montreal, Canada.
ICAO, International Civil Aviation Organization (current-c) *Airport Services Manual*, Doc. 9137, Montreal, Canada.

---

[1]The site http://www.icao.int/publications/Pages/catalogue.aspx provides the list of ICAO publications; they can be purchased online.

[2]The FAA provides advisory circulars and FARs at: http://www.faa.gov/regulations_policies/advisory_circulars, and http://www.faa.gov/regulations_policies/faa_regulations, respectively.

[3]The standard atmosphere is defined as: temperature of 15°C, pressure of 76 cmHg at sea level, and a temperature gradient of 0.0065°C/m from sea level to an altitude of 11,000 m.

[4]The magnetic azimuth and the true azimuth of a runway may differ, depending on the airport's location. In the United States this difference may range from 0° to a little more than 20°.

[5]Dallas/Ft. Worth has two more runways, 13L/31R and 13R/31L, for a total of seven, as of 2012 the largest number (along with Chicago/O'Hare) of any existing airport.

[6]Operations may sometimes be conducted with a slight tailwind. The maximum tailwind permitted is typically 5 or 6 knots (about 9–11 km/h).

[7]Dubai is included in the number of Asian airports.

[8]The corresponding statistics for 2000, 19 in North America, 5 in Asia/Oceania, and 6 in Europe, indicate the rapid ascendancy of Asian airports during the decade 2001–2010.

[9]The ICAO specifies this distance as 760 m (ICAO, 2009), whereas the FAA converts the distance of 2500 ft to 762 m (FAA, 2012). Many minor discrepancies of this type exist, because of different practices in converting between systems of units.

[10]The first three of these airports actually have additional runways with a different orientation from the close or medium-spaced parallel pair. However, the great majority of movements take place on the parallel pair.

[11]In 2011 Frankfurt/International opened a fourth runway for *arrivals only*. It is to the northwest of the terminal complex and quite far from it. It permits arrivals independent from those taking place on the two main close parallel runways. Its utilization will be low during its initial years of operation.

[12]Runways that intersect only at a point along their projected centerlines are called "converging" or "diverging," depending on the direction of operations.

[13]For details on design standards for clearways, see FAA (2012) and ICAO (2009).

[14]Clearways and stopways, even if present, are not taken into consideration when computing the declared distances in the case of piston aircraft. Thus in this case TORA = TODA = ASDA. Balanced field length (see later) always applies to takeoffs of piston aircraft.

[15]Frangible objects are those with low mass designed to break, distort, or yield on impact, so as to present the minimum hazard to aircraft (ICAO, 2009).

[16]The FAA requires an approach RPZ and a departure RPZ. Because the dimensions and requirements of the approach RPZ exceed those of the departure RPZ, the approach RPZ requirements must be satisfied at both ends of the runway in the case of bidirectional runways,.

[17]For detailed descriptions and special cases, see Tables 3.1 and 3.2 in ICAO, 2009.

[18]Comparisons cannot be exact. The ICAO and FAA specify these standards somewhat differently in terms of their airport reference codes and runway instrumentation classes.

[19]The FAA provides for the possibility of increased distances to holdlines, if aircraft intrude on areas critical to the operation of instrument landing systems.

[20]The ICAO does have a sight distance requirement for taxiways. For example, where the ICAO reference code letter is C, D, E, or F, an observer 3 m above the surface of the taxiway should be able to see the entire surface of the taxiway to a distance of at least 300 m (ICAO, 2009).

[21]Any translation between capital and annual costs is imprecise. With depreciation and maintenance, the annual cost of a building is about 10 percent of its capital cost. This is the basis for using a factor of 10 to translate annual savings into an approximation of justifiable additional capital expenditure.

[22]The term *extra taxiway width* is often used to refer to any added width of curved segments of taxiways, whereas *fillet* usually refers to additional width provided at junctions and intersections; in either case, the strength of the additional pavement provided must be the same as that of the taxiway.

[23]A few airports also have parts of runways on bridges that pass over highways.

[24]An important factor in apron operations is the turning radius of aircraft. This is the distance between the pivot around which the aircraft turns and the part of the aircraft farthest from the pivot. The pivot is located along the centerline of the main undercarriage of the aircraft, at some distance from the center of the aircraft's fuselage, typically under the inner part of the wing. The point farthest from the pivot is usually a wing tip, but, for some aircraft, it can be the nose or the tips of the horizontal stabilizers at the tail of the aircraft. The turning radius of the common types of commercial jets typically equals between 65 and 100 percent of the wingspan of the aircraft.

CHAPTER 10

Airfield Capacity

This chapter reviews the subject of airfield capacity, a topic fundamental to modern airport planning and design. The capacity of the airfield and especially of runway systems typically determines the ultimate capacity of an airport.

Maximum throughput capacity is the principal and most fundamental measure of the capacity of a runway system. It indicates the average number of movements that can be performed on the runway system in 1 hour in the presence of continuous demand, while adhering to all the separation requirements imposed by the air traffic management (ATM) system. Practical hourly capacity (PHCAP), declared capacity, and sustained capacity are measures designed to estimate the number of hourly movements at which operations can be performed over an extended period of time at acceptable levels of delay. All three are "derivative" measures, in the sense that they are based on and can be derived from the fundamental measure of maximum throughput capacity. These three capacities are typically equal to 80 to 90 percent of the maximum throughput capacity.

The (maximum throughput) capacity of a runway system depends on many parameters and factors. The most important of these are the number and geometric layout of the runways, the ATM separation requirements, weather conditions (visibility, precipitation, wind direction and strength), mix of aircraft types, mix and sequencing of runway movements, type and location of runway exits, performance of the ATM system, and environmental restrictions on operations. The runway system capacities that one encounters at major airports in various parts of the world span a wide range. In developed countries, capacity per runway at major airports ranges from 25 to 60 movements per hour. A few mega-airports in the United States operate with as many as four or more simultaneously active runways and can serve more than 200 movements in 1 hour.

The range of capacities available at an airport over a long period of time, such as a year, and the frequency with which these capacities are available can be summarized through the capacity coverage chart (CCC). The CCC makes the assumptions that (1) the operations mix consists of 50 percent arrivals and 50 percent departures and (2) the runway configuration in use at any given time is the one that provides the highest capacity under the prevailing conditions. An "uneven" CCC indicates an airport where the supply of runway capacity is not reliable. In practice, this may create serious operational problems.

Simple mathematical models can be used to obtain good approximations to the capacity of runway systems with simple geometric configurations. Moreover, these models provide insight into the sensitivity of capacity to changes in such parameters as separation requirements, traffic mix and characteristics, etc. Runway capacity envelopes of simple runway systems can also be computed approximately from such models. These envelopes indicate the capacity that can be achieved for all possible mixes of arrivals and departures. A number of computer-based simulation and mathematical models are available for estimating airfield capacity and delay at airports with more complex airfield layouts.

The capacity of taxiway systems depends greatly on local conditions and the geometric configuration at hand at each airport. The capacity of the taxiway system of major airports almost always exceeds that of the runway system by a considerable margin. Delays sustained at specific "choke points" are typically much smaller than those experienced due to the capacity limitations of the runway system. However, some exceptions may exist at older, space-constrained airports. Taxiway capacity problems are airport-specific and must be resolved in the context of local conditions.

It is important to distinguish between the static capacity of an apron, that is, the number of aircraft that can be stationed there at any particular instant, and the dynamic capacity, which indicates the number of aircraft that can be served at the apron per unit of time. The dynamic capacity depends strongly on the stand blocking time (SBT). Its determination is often difficult, due to the differences in the sizes of stands and the large number of constraints and conditions regarding stand assignments. These constraints and conditions vary greatly among airports. It is also difficult sometimes to compare apron capacity with runway system capacity. In the long run, all but the most space-constrained airports should be able to increase their apron capacity to a level greater than the capacity of the runway system.

10.1 Introduction

This chapter reviews the subject of airfield capacity. The emphasis is on the capacity of runway systems of major commercial airports. This is a topic fundamental to modern airport planning and design, because it is the capacity of the airfield and especially of runway systems that typically determines the ultimate capacity of an airport. The runway complex is usually the principal "bottleneck" of the ATM system because, quite simply, air traffic transitions from three-dimensional flows in airspace to a "single-file" regime at the runway and the final approach airspace in its immediate vicinity. Moreover, it is usually extremely difficult and time-consuming to increase substantially the capacity of the runway system of a major airport. New runways, along with associated protection zones, noise buffer space, etc., typically require acquisition of a large amount of additional land area. Equally important, runways have significant environmental and economic impacts that necessitate long and complicated review-and-approval processes with uncertain outcomes. By contrast, the capacity of landside facilities (passenger and cargo terminals, road access, etc.) and of other airfield facilities (taxiways, apron stands) can usually be increased, in one way or another, to equal or exceed the capacity of the runway system.

The subject of airport capacity and delay has received a great deal of attention, not only by airport professionals but also by the public at large, as air traffic delays have increased and spread geographically. The problem is particularly acute in North America, Western Europe, and the Pacific Rim. Many airline executives and aviation officials believe that the principal threat to the long-term future of the global air transportation system is the apparent inability of available runway capacity to keep up with growing air traffic demand at many of the world's most important airports.

Chapter 10 begins with a review and explanation of the several definitions of airfield capacity that are currently in use—and have, unfortunately, been the cause of much confusion among aviation professionals (see Sec. 10.2). The various factors that determine the capacity of a runway system are then discussed in some detail in Sec. 10.3. The objective is to help the reader appreciate the complex relationships that play a role in determining runway capacity, as well as the reasons why it may be very difficult to increase capacity beyond a certain level at any particular location. It is also noted that the capacity of a runway system is not a constant, but a highly variable quantity, as it depends on several parameters that vary probabilistically and dynamically. Section 10.4 presents a brief survey of the range of runway system capacities at major airports around the world. It also introduces the notion of the CCC as a means of displaying the range of capacity values associated with a runway system over time and the relative frequency with which these values occur. Section 10.5 turns to the issue of computing the capacity of an airport. The standard approach for computing the capacity of a single runway is outlined first and one of the best-known mathematical models available for this purpose is presented in some detail. Section 10.6 then discusses generalizations of the single-runway model of Sec. 10.5 and describes briefly more detailed versions of the model. It also introduces the important concept of the capacity envelope and suggests ways to extend the single-runway methodology to more complex runway systems. Finally, Sec. 10.7 is concerned with the capacity of the taxiways and aircraft stands of an airport, concentrating primarily on the estimation of the capacity of the apron area, a quite complex problem.

10.2 Measures of Runway Capacity

Several alternative measures of runway capacity are currently in use, all of them intended to provide an estimate of how many aircraft movements (arrivals and/or departures) can be performed on the runway system of an airport during some specified unit of time—typically one hour. To utilize them properly and to avoid confusion, one must understand clearly the definitions of these alternative measures.

It is essential to realize at the outset that, from a long-term perspective, runway capacity is a probabilistic quantity—a “random variable”—which can take on different values at different times, depending on the circumstances involved. For a simple example, the number of arrivals and departures that can be performed on a runway during any particular hour at a busy airport depends on the “mix” of aircraft that will be using the runway during that hour. If, for instance, the mix happens to include a high percentage of wide-body aircraft (B747, B777, A340, etc.), the capacity will generally be lower than at times when the mix consists, for the most part, of smaller aircraft (regional jets, turboprops, B737, A320, etc.). The reason is that bigger airplanes generate wake vortices that may pose a threat to aircraft flying immediately behind them. To ensure safety, providers of ATM services [e.g., the Federal Aviation Administration (FAA), in the United States] require longer separations (in terms of time or distance) between pairs of consecutive aircraft whenever the first aircraft in the pair is a heavy one. And even with identical mixes, the number of movements that can be performed may vary depending on winds, visibility, proficiency of air traffic controllers, and many other factors. One should not forget that the figures that are cited for the runway system capacity of any airport (e.g., 90 movements per hour) typically only refer to the “average number” or, more formally, the expected number of movements that can be performed per unit of time. The actual capacity during any given hour may deviate significantly from that average value.

The first and, as will be seen later, the most fundamental measure of runway capacity can now be introduced. The *maximum throughput capacity* (or saturation capacity) is defined as the expected number of movements that can be performed in 1 hour on a runway system without violating ATM rules, assuming continuous aircraft demand.

Two points should be noted right away. First, to compute the maximum throughput capacity, one needs to know the specific conditions under which runway operations are conducted. This means specifying the ATM separation requirements in force, the mix of aircraft, the mix of movements (arrivals and departures), the allocation of movements among the runways (if the runway system consists of more than one runway), and several other factors that will be described in Sec. 10.3.

Second, the definition of maximum throughput capacity makes no reference to any level-of-service (LOS) requirements. In other words, all one cares to know is how many aircraft movements can be processed on average per hour, if the runway system is utilized to its maximum potential in the presence of “continuous aircraft demand.” Whether this means a delay per movement of a few minutes or of several hours is immaterial, as far as this measure of capacity is concerned.

It is the absence of any reference to LOS that has motivated the use in practice of three other measures of hourly capacity. The common characteristic of all three is that they define capacity indirectly, through the explicit or implicit specification of an acceptable threshold of LOS or of air traffic controller workload. The three measures include the PHCAP, the sustained capacity, and the declared capacity.

The *practical hourly capacity* (PHCAP) is the oldest of these measures, having originally been proposed by the FAA in the early 1960s. It is defined as the expected number of movements that can be performed in 1 hour on a runway system with an average delay per movement of 4 minutes.

Note that this definition specifies a threshold value for acceptable LOS ("average delay of 4 minutes per movement") and states that the runway system "reaches its capacity" when that threshold is exceeded. As a rule of thumb, the PHCAP of a runway system turns out to be approximately equal to 80 to 90 percent of its maximum throughput capacity, depending on the specific conditions at hand. Note that today the average delay per movement is considerably higher than 4 minutes at practically every major airport, especially during peak traffic hours, so that all these airports are operating "above capacity" by this definition. This, however, does not invalidate the notion of a "practical" hourly capacity tied to a threshold of acceptable LOS.[1] Instead, it simply indicates that the failure of runway capacity to keep up with demand has forced many airports to operate routinely at a LOS much lower than what was considered acceptable in 1960 and that a threshold higher than 4 minutes might be more appropriate today. In fact, the selection of the particular threshold value of 4 minutes is not unreasonable, in any case, and is supported, as will be seen in Chap. 11, by observations from queuing theory, the mathematical theory of waiting lines.

The *sustained capacity* of a runway system is a measure defined, rather fuzzily, as the number of movements per hour that can be reasonably sustained over a period of several hours. "Reasonably sustained" refers primarily to the workload of the ATM system and of air traffic controllers. To achieve maximum throughput capacity, the ATM system must operate to its full potential. This may not be feasible in practice for more than 1 or 2 consecutive hours, at most. It is thus argued that one should specify a more realistic target than maximum throughput, when it comes to operations over periods of several hours or of entire days of air traffic activity.

A good example of the application of the notion of sustained capacity is the setting of performance targets[2] at many major airports in the United States. These targets are determined after discussions between FAA specialists and the local air traffic controller teams and specify desirable levels of runway system capacity to be achieved at each participating airport over periods of several hours. For example, the sustained capacity for Boston/Logan in good weather conditions and operations to the northeast was set in 2000 to approximately 110 movements per hour. This capacity is usually further subdivided into a sustained arrival capacity, the airport acceptance rate (AAR), and sustained departure capacity, the airport departure rate (ADR)—see also Chap. 13. Typically, the sustained capacity is set to approximately 90 percent of maximum throughput capacity when runway configurations with high maximum throughput capacity are in use (e.g., in good weather conditions) and to almost 100 percent of maximum throughput capacity with configurations with low maximum throughput capacity.[3] The reasoning is that low-capacity conditions, usually associated with poor weather, prevail for only a few consecutive hours at a time and it is critical to operate as close as possible to the maximum available capacity during those periods.

Declared capacity is another measure based on the same general notion as sustained capacity. It is defined, again somewhat ambiguously, as the number of aircraft movements per hour that an airport can accommodate at a reasonable LOS. Delay is used as the principal indicator of LOS. While practical hourly capacity and sustained capacity are measures used (with decreasing frequency) in the United States, the notion of declared capacity is one that has been widely adopted in the rest of the world and provides the basis for the worldwide practice of "schedule coordination" and "slot allocation" that will be described in detail in Chap. 12. Under this practice, each airport that experiences congestion "declares" a capacity, which is then used to set a limit on the number of movements per hour that can be scheduled there. For example, in the summer of 2011, the declared capacity of Frankfurt/International was set to 81 to 84 movements per hour, with the exact number varying with the time of day, depending on the mix of arrivals and departures and the percent of scheduled wide-body movements in each hour.

Unfortunately, there is no generally accepted definition of declared capacity and no standard methodology for setting it. It is essentially left up to local or national airport and civil aviation organizations, in cooperation with other interested parties, to compute and set the declared capacity (see Chap. 12). The approaches used for this purpose vary from country to country and, often, even from airport to airport within the same country. In fact, there are instances of airports where the declared capacity is dictated by the capacities of the passenger terminal or of the apron area, which are believed to be more constraining than the capacity of the runway system. In most instances, however, the declared capacity of major airports outside the United States is set close to approximately 85 to 90 percent of the maximum throughput capacity of the runway system. As in the case of sustained capacity, the reasoning is that this choice will ensure reliability of airport operations, as well as a reasonable LOS.

The advantages and disadvantages of the measures of runway capacity introduced in this section can now be summarized. Among these measures, the maximum throughput capacity is clearly the most fundamental and least subjective. It provides an estimate of capacity in its truest sense: how many operations can be performed per unit of time, on average, when the runway system is pushed to its limits. Indeed, it is possible to obtain rough estimates of maximum throughput capacity by collecting data in the field. All one needs to do is to observe the runways and count the number of movements taking place during a continuously busy period, that is, when all movements experience some delay. Note this means that it is much easier to measure maximum throughput capacity at a very congested airport than at a less busy one. Moreover, the data should be collected during peak traffic hours rather than at off-peak. Equally important, maximum throughput capacity can be computed quite accurately through a number of existing analytical and simulation models—see Secs. 10.5 and 10.6. These models make it possible to estimate capacity under hypothetical future conditions, in addition to existing ones.

The principal disadvantage of maximum throughput capacity is that it does not consider LOS in any way. In fact, as will be seen in Chap. 11, extremely long delays will be experienced whenever the average number of movements scheduled at an airport per hour is very close to the runway system's maximum throughput capacity for several hours in a row. By contrast, when demand remains close, on average, to the practical hourly capacity or to sustained capacity,[4] the LOS, as measured by the amount of delay per flight, will usually remain at acceptable levels. Thus, measures such as PHCAP, sustained capacity, and, in most cases, declared capacity are good indicators of how much demand can be accommodated at a reasonable LOS. PHCAP, sustained capacity, and declared capacity are also useful measures for planning purposes. When the average demand per hour, over a period of several hours of a day, grows over the years to a level close to the PHCAP or the sustained capacity or the declared capacity, this is a clear signal that an increase in the airport's capacity is highly desirable. Even relatively small increases in demand beyond that critical level will probably lead to unacceptable delays and airfield congestion.

In conclusion, PHCAP, sustained capacity, and declared capacity are somewhat subjective measures of capacity that can, however, be very useful if applied properly. They are also "derivative" measures, in the sense that one needs to compute the maximum throughput capacity before one can estimate these other capacity measures. This will be discussed further in Chap 11.

Henceforth in this and subsequent chapters, the term *runway capacity* is used to refer to the maximum throughput capacity of a runway system. Whenever reference is made to some other measure of capacity (e.g., the "declared capacity"), this will be stated explicitly.

The reader should also bear in mind the following convention: the often-heard statement, "the capacity of Airport A is X movements per hour," typically makes the implicit assumption that X consists of approximately 50 percent arrivals and 50 percent departures. When this is not the case, the statement is usually more detailed (e.g., "the arrival capacity of the runway is Y," or "when two runways are used for arrivals and one for departures, the arrival capacity is Z and the departure capacity is W").

Finally, note that all the measures of capacity mentioned so far use the hour as their unit of time. Another natural measure of great practical interest is the annual capacity of an airfield. This is a number that can be compared readily with airport demand forecasts that are typically given in terms of annual estimates ("500,000 aircraft movements expected by

2020”). In fact, the FAA has been using for some years the measure of practical annual capacity (PANCAP) for this purpose. As is discussed in Chap. 11, PANCAP and other similar estimates of annual capacity can be derived from the fundamental measure of maximum throughput capacity per hour. Annual measures of capacity must necessarily be tied to a LOS and should take into consideration the daily and seasonal patterns of demand at an airport.

10.3 Factors Affecting the Capacity of a Runway System

The dependence of the capacity of any runway system on many different factors is emphasized in the previous section. This section provides an overview of the following important factors and of the ways in which each affects runway capacity:

- Number and geometric layout of the runways
- Separation requirements between aircraft imposed by the ATM system
- Visibility, cloud ceiling, and precipitation
- Wind direction and strength
- Mix of aircraft using the airport
- Mix of movements on each runway (arrivals only, departures only, or mixed) and sequencing of movements
- Type and location of taxiway exits from the runway(s)
- State and performance of the ATM system
- Noise-related and other environmental considerations and constraints

One of the objectives of the discussion below is to make the reader aware of the complex relationships that are often at play.

Number and Geometric Layout of the Runways

The most obvious, and usually single most important, factor influencing a runway system’s capacity is the number of runways at the airport and their geometric layout. From a practical point of view, the surest way to achieve a “quantum increase” in the capacity of an airport is by constructing a well-located (relative to the other existing runways) and well-designed runway.[5] Unfortunately, as noted at the beginning of this chapter, adding a new runway is a task that today ranges from “very difficult” to “impossible” at most of the world’s busiest and most congested airports. The following are some introductory observations on how the number and geometric layout of runways affect capacity.

First, it is important to distinguish between the number of runways at an airport and the number that are active at any given time. For example, Boston/Logan and Amsterdam/Schiphol have five main runways each, but no more than three of them are ever active simultaneously, due to the geometric layouts of the runway systems and to noise restrictions. By contrast, Atlanta has five runways and uses all five simultaneously during most of the busy hours of the day. Similarly, Dallas/Fort Worth has seven runways and typically uses six or sometimes all seven during busy hours. It is the number of simultaneously active runways that is a primary factor in determining airfield capacity.

Second, the number and identity of runways in use at any given time, as well as the allocation of types of aircraft and movements to them, may change several times a day at many airports. The selection of the specific set of runways to be operated at any one time depends on many of the factors to be discussed further in this section: demand (e.g., during periods of low demand an airport may accommodate all its traffic on a single runway, even though more than one runway may actually be available); weather conditions, including visibility, precipitation, and wind speed and direction; mix of movements (e.g., during peak periods for flight arrivals, one or more runways may be dedicated to serving arrivals exclusively—and conversely for peak departure periods); and noise restrictions, which, for instance, may prohibit or discourage the use of certain runways during the night or during certain parts of the year. For an airport with several runways, there can be a large number of combinations of simultaneously active runways, weather conditions, and assignments of aircraft types and movements (arrivals and/or departures) to the active runways. Each of these combinations is called a *runway configuration*. For example, Boston/Logan, with five runways, can operate in about 20 different configurations!

Third, the details of the geometric layout of any set of runways are extremely important, as they determine the degree of dependence among the runways. Section 9.4 has already provided an overview of this topic, and more will be said below in connection with the description of the relevant ATM separation requirements.

ATM Separation Requirements

Every ATM system, no matter how advanced or primitive, specifies a set of required minimum separations between aircraft flying under instrument flight rules (IFR). Obviously, the purpose of these requirements is to ensure safety. In turn, the separation requirements determine the maximum number of aircraft that can traverse each part of the airspace or can use a runway system per unit of time.

Required separation distances between aircraft operating under IFR at major airports in the United States are typically among the smallest (or "least conservative") anywhere, reflecting in part the need to maximize airport capacity, as well as the outstanding proficiency and training of the air traffic controllers. Several major European airports, such as London/Heathrow, London/Gatwick, Copenhagen, Frankfurt/International, Munich, and Amsterdam/Schiphol, have also come to be operated in recent years with separation requirements that are essentially identical to those used at the busiest airports in the United States. Such "tight" separation requirements recognize the need for more capacity at these airports and have been made possible by the outstanding ATM capabilities that have been developed there.

Separation Requirements for Aircraft Operating to/from the Same Runway

The longitudinal separation requirements for aircraft landing on or departing from the same runway are of special importance in determining runway capacity. Typically, each type of aircraft is assigned to one of a small number (usually, three or four) of classes according to the aircraft's size and/or weight. The separation requirements are then specified in units of distance or of time. Each set of requirements gives the minimum separation that must be maintained at all times between two aircraft operating consecutively on the runway. The requirements are specified for every possible pair of classes and every possible sequence of movements: "arrival followed by arrival," A-A; "departure followed by departure," D-D; "arrival followed by departure," A-D; and "departure followed by arrival," D-A (see Example 10.1).

Example 10.1 In the United States, the FAA assigns all aircraft to three classes, according to their maximum certified takeoff weight (MTOW): heavy (H), large (L), and small (S). Aircraft with

MTOW greater than 255,000 lb (~ 116 tons) are in Class H.

MTOW between 41,000 lb (~ 19 tons) and 255,000 lb (~ 116 tons) are in L.

MTOW less than 41,000 lb (~ 19 tons) are in S.

In addition, the FAA also identifies the Boeing 757, whose MTOW places it at the borderline between the L and H classes, as an aircraft class by itself for some terminal airspace separation purposes, because of its strong wake-vortex effects.

Wide-body commercial jets generally belong to the H class (see also Table 9.3). [However, the "super heavy" aircraft in FAA Group VI (or ICAO Reference Code F)—see Sec. 9.2—are being treated as special cases, as is explained later.] Class L includes practically all types of narrow-body commercial jets and regional jets and some types of turboprops used by regional (or "commuter") air carriers. Finally, Class S includes most general aviation aircraft, including many general aviation jets, as well as many types of nonjet aircraft used by regional air carriers and by air taxi operators.

Table 10.1 summarizes the FAA separation requirements for movements on the same runway under IFR (FAA, 2011). Note that requirements are specified for all four possible combinations of movements (A-A, A-D, D-D, D-A) and for all possible pairs of aircraft classes. For example, it can be seen (A-A separations) that when the landing of a large (L) aircraft is followed immediately by the landing of a small (S) aircraft, the minimum separation allowed between the two when the leading aircraft is at the threshold of the runway is 4 nautical miles (nmi). If a departure is to be followed immediately by an arrival (and regardless of the classes of aircraft involved), the arriving aircraft must be at least 2 nmi away from the runway at the time when the departure run begins[6] and cannot touch down on the runway before the preceding departing aircraft has lifted off (D-A separations). In the reverse situation, that is, when an arrival is followed immediately by a departure, the arriving aircraft must be safely out of the runway before the takeoff run can begin, again regardless of the classes of the two aircraft involved. Note that some separations are specified in terms of time or of occurrence of an event ("clear of runway"), whereas others are specified in units of distance.

Arrival followed by arrival (A-A)

A. Throughout final approach, successive aircraft must be separated by at least the distance (in nautical miles) indicated by the table below. (Asterisked separations are required when leading aircraft is at the threshold of the runway.)

		Trailing Aircraft		
		H	**L + B757**	**S**
Leading aircraft	H	4	5	5/6*
	B757	4	4	5
	L	2.5 (or 3)	2.5 (or 3)	3/4*
	S	2.5 (or 3)	2.5 (or 3)	2.5 (or 3)

B. The trailing aircraft cannot touch down on runway before the leading aircraft is clear of it.

Arrival followed by departure (A-D)

Clearance for takeoff run of the trailing departure is granted after the preceding landing is clear of the runway.

Departure followed by departure (D-D) (separations are approximate—see text)

Clearances for takeoff run of successive aircraft must be separated by at least the amount of time (in seconds) indicated by the table below.

		Trailing Aircraft		
		H	**L + B757**	**S**
Leading aircraft	H	90	120	120
	B757	90	90	120
	L	60	60	60
	S	45	45	45

Departure followed by arrival (D-A)

The trailing arrival on final approach must be at least 2 nmi from runway when departing aircraft begins its takeoff run, and cannot touch down until departing aircraft is clear of the runway.

TABLE 10.1 Single-Runway IFR Separation Requirements in the United States in 2010

Aircraft pairs in which the first aircraft is in Class H or B757 generally require greater separations than other pairs, both in the A-A and in the D-D cases. The reason is that these aircraft classes generate severe wake turbulence (wake vortices) behind them. A wake vortex poses the threat of destabilizing a trailing aircraft that runs into it, especially if the trailing aircraft belongs to Class S.

As Table 10.1 shows, the A-A case requires that two conditions be satisfied: (1) The two landing aircraft should not be on the runway at the same time;[7] and (2) while airborne on final approach, the two aircraft must be separated by a minimum distance specified in units of nautical miles. The separations of 4, 5, and 6 nmi in the A-A case are all related to the potential presence of wake turbulence due to the leading aircraft, whereas the 2.5-nmi separations apply in pairings where it is believed that wake vortices are not a factor. The two A-A separation requirements denoted with an asterisk in Table 10.1 apply only at the time when the leading aircraft (H or L, as the case may be) is at the runway threshold, that is, about to touch down on the runway. All the other A-A separation requirements apply at all points of the final approach path to the runway. In the specific case of an H-S pair, the trailing Class S aircraft is required to be at least 5 nmi behind the leading Class H aircraft at all points on final approach and at least 6 nmi behind at the instant when H is at the runway threshold—with the more restrictive of the two requirements dictating the actual separation. The 2.5-nmi separation between the indicated pairs of aircraft classes is used only at the busiest airports in the United States. At other airports 3 nmi is used, as indicated in parentheses in Table 10.1.

Turning next to the D-D case, the requirements shown in Table 10.1 give approximate (and somewhat conservative) estimates of the time separations that result in practice[8] from the following more complicated set of rules:

1. In the case where the leading departing aircraft belongs to Class L (Class S) and thus wake turbulence is not a factor, the takeoff run of the trailing aircraft can start after the leading aircraft is airborne and (a) is at a distance of more than 6000 ft (4500 ft) from the trailing aircraft or (b) has either cleared the runway end or has turned out of conflict.

2. In the case where the leading departing aircraft belongs to Class H or to Class B757 (and thus wake turbulence is a factor), the takeoff run of the trailing aircraft can start as soon as one of the following two conditions has been satisfied: (a) 2 minutes have elapsed since the start of the takeoff run of the leading aircraft or (b) the following separations, in nautical miles, have been assured when the trailing aircraft becomes airborne.

		Trailing Aircraft		
		H	L or B757	S
Leading aircraft	H	4	5	5
	B757	4	4	5

The separations of 90 and 120 seconds behind H-class and B757 aircraft shown in Table 10.1 are approximate estimates of the earliest time it takes to satisfy, in each case, the less constraining of the above two conditions, (a) and (b).

Finally, a special note about the FAA Group VI or ICAO Reference Code F aircraft: Despite the fact that the A380 has been operating (in small numbers) since 2008, the ICAO and Air Navigation Service Providers (ANSP) around the world have not yet settled on a definite set of separation requirements for these aircraft. Because of concerns about their vortex effects, Group VI/Code F aircraft are now classified as “Super Heavy” (SH) aircraft which, when leading an A-A or D-D pair, require additional separation from the trailing aircraft. In the case of A-A, this additional separation is currently 2 nmi, while in the case of D-D it is about 60 seconds. Therefore, when the leading aircraft is in Group VI/Code F, the first row of the A-A separations in Table 10.1 becomes 6 nmi if the trailing aircraft is in Class H (or is a Group VI/Code F aircraft), 7 nmi if it is in Class L or B757, and 7 (or 8*) nmi if it is in Class S. The corresponding requirements for D-D are 150, 180, and 180 seconds, respectively. These requirements should be viewed as temporary and may be revised, possibly downward, in the future as more experience is acquired concerning the wake effects of these very large aircraft.

Obviously, the larger (or more “conservative”) the separations required by the ATM system, the lower the capacity of a runway. To emphasize this point, Table 10.2 lists the separation requirements used at some smaller regional airports in Europe (and until about the late 1990s at many major airports, as well). Clearly, these separations are considerably more conservative than those of Table 10.1 that apply to major airports in the United States. It stands to reason that the runway system capacity of an airport in the United States is typically higher than the capacity of regional airports in Europe with similar runway system layouts, even for the same mix of aircraft and of movements. For instance, the declared capacity of the single-runway airport at Milan/Linate in 1998 was 32 movements per hour. By comparison, the single-runway San Diego airport often handles as many as 60 movements per hour.

Arrival followed by arrival (A-A)				
A. Throughout final approach, successive aircraft must be separated by at least the distance indicated by the table below (in nautical miles):				
		Trailing Aircraft		
		H	**L**	**S**
Leading Aircraft	**H**	5	5	7
	L	5	5	5
	S	5	5	5
B. The trailing aircraft cannot touch down before the leading aircraft is clear of the runway.				
Arrival followed by departure (A-D)				
Clearance for takeoff run of the trailing departure is granted after the preceding landing is clear of the runway.				
Departure followed by departure (D-D)				
Clearances for the takeoff run of successive departures must be separated by at least 120 s.				
Departure followed by arrival (D-A)				
At the start of the takeoff run of the leading (departing) aircraft, the trailing landing aircraft must be at least 5 nmi from the end of the runway. (Typically this translates to about 2–2.5 min between beginning of the takeoff run of the leading departure and the subsequent touchdown of the trailing arrival.)				

TABLE 10.2 Simplified Single-Runway IFR Separation Requirements in Effect at Rome and Milan Airports until 1998

Separation Requirements for Aircraft Operating to/from Parallel Runways

The separation requirements for aircraft landing on or departing from a pair of parallel runways play a critical role at many airports that often operate with more than one active runway. Most of these multirunway airports rely largely on operations to/from parallel runways.

Table 10.3 summarizes the FAA separation requirements for operations on parallel runways under IFR. The "arrival/arrival" column refers to the required separation between a pair of arriving aircraft, the first of which is landing on one of the parallel runways and the second on the other. Similarly, "departure/arrival" refers to the situation in which the first aircraft in the pair will depart from one of the parallel runways and the second will land on the other. The "departure/departure" and "arrival/departure" columns should be interpreted in a similar way.

Separation between Runway Centerlines	Arrival/ Arrival	Departure/ Departure	Arrival/ Departure	Departure/ Arrival
Up to 2500 ft (up to 762 m)	As in single runway	As in single runway	Arrival must be over runway and committed to land	As in single runway
2500–4300 ft (762–1310 m)	1.5 nmi	Independent	Independent	Independent
4300 ft or more (1310 m or more)	Independent	Independent	Independent	Independent

Source: FAA, 1989

TABLE 10.3 IFR Separation Requirements between Aircraft Movements on Parallel Runways in the United States

The critical parameter is now the distance between the centerlines of the runways. For runway centerlines that are separated by less than 2500 ft (762 m), the separation requirements in the "arrival/arrival" case are the same as those in Table 10.1, when the two aircraft are landing on the same runway. In other words, the second aircraft should follow the first by 2.5 (or 3), 4, 5, or 6 nmi, depending on the classes of the two aircraft. Similarly, the separations in Table 10.1 also apply to the "departure/arrival" case (the landing aircraft must be at least 2 nmi from the parallel runway when the departure roll on the parallel runway begins and should not touch down on its runway before the departure on the parallel runway has lifted off), as well as to the "departure/departure" case. The only change from the separation requirements of Table 10.1 occurs in the "arrival/departure" case: the departing aircraft does not have to wait for the landing aircraft to exit the parallel (arrival) runway—as was the case with a single runway—but can begin its takeoff roll as soon as the landing aircraft touches down on the parallel runway (or, in a less conservative interpretation of the rules, crosses the threshold of that runway).

The situation changes considerably when the separation between the centerlines of the two parallel runways exceeds 2500 ft (762 m). Now, the two parallel runways may operate independently when both are used for departures or when one is used for arrivals and the other for departures. “Independently” means that, absent airspace constraints, an aircraft movement on one runway does not have an impact on a movement on the other. When both runways are used for arrivals, the trailing aircraft has to be at least 1.5 nmi behind the leading one when the two centerlines are between 2500 ft (762 m) and 4300 ft (1310 m). The 1.5-nmi distance is measured diagonally; that is, it represents the direct distance between the two aircraft (Fig. 10.1). Finally, when runway centerlines are more than 4300 ft (1310 m) apart, the two parallel runways may be operated independently,[9] even if both are used for arrivals. At some airports, the FAA has also authorized simultaneous approaches to parallel runways separated by as little as 3400 ft (1035 m) when a precision runway monitor system (PRM) is available. The FAA will, in fact, consider (FAA, 2012) on a case-by-case basis, authorizing simultaneous approaches to parallel runways with centerline separations down to 3000 ft (915 m).

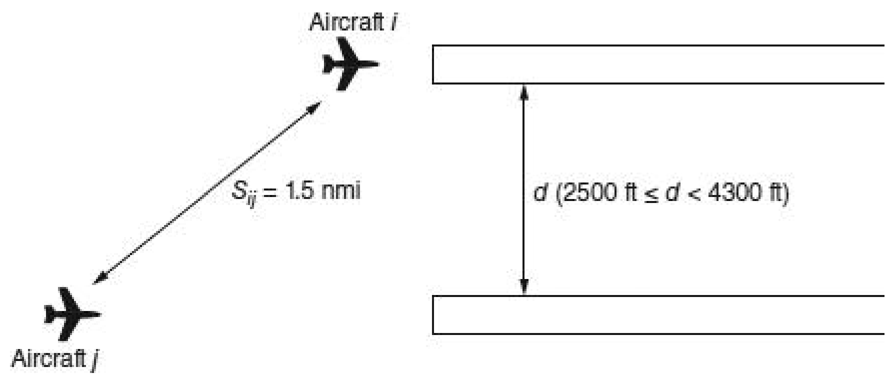


FIGURE 10.1 The diagonal separation between two aircraft approaching medium-spaced parallel runways.

Two additional points should be mentioned. First, when both runways are used for departures, independent movements are allowed only if the aircraft departing from each of the parallel runways will follow diverging climb paths after takeoff—as is most often the case in such circumstances. If not, one must apply the same separation requirements as for departures from a single runway just as in the case of close-spaced (under 2500 ft) parallel runways. Second, Table 10.3 assumes that the parallel runways are not “staggered”—i.e., their thresholds are not offset (Fig. 10.2). If they are staggered, an "effective separation distance" between the centerlines of the two runways should be computed. Specifically, when arrivals are to the "near end" in the direction of operations (see Fig. 10.2), the 2500-ft (762-m) separation requirement between runway centerlines is reduced by 100 ft (30 m) for each 500 ft (150 m) of threshold offset, down to a minimum of 1200 ft (366 m). For example, when the offset is 1000 ft (300 m), a separation of 2300 ft (690 m) between the runway centerlines is equivalent to 2500 ft (762 m) when there is no offset. In other words, arrivals on one runway and departures on the other can be performed independently on a pair of parallel runways whose centerlines are 2300 ft apart and whose thresholds are staggered by 1000 ft, as long as the arrivals are assigned to the "near-end" runway, as shown in Fig. 10.2. The reverse applies when arrivals are to the far threshold: the 2500 ft separation between runway centerlines must then be increased using the same method.

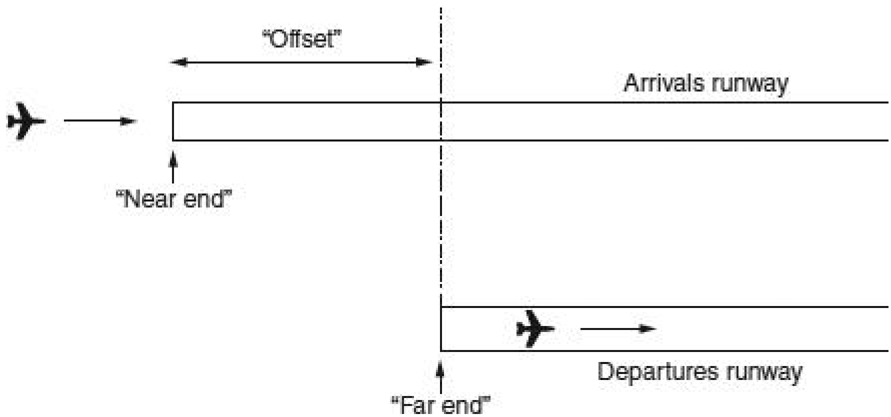


FIGURE 10.2 Staggered parallel runways; the "near" runway is used for arrivals and the other for departures.

While Table 10.3 shows the separation requirements that apply to operations on parallel runways in the United States, the pattern it presents is also typical of separation requirements for parallel runways elsewhere. Generally speaking, the greater the distance between runway centerlines (and the greater the offset of the runway thresholds), the "less dependent" operations on the two runways are. However, differences with the specific values used in the United States abound. Following the FAA's lead, the International Civil Aviation Organization (ICAO) recommends that a distance between centerlines of at least 1035 m (3400 ft) be required for simultaneous instrument approaches, provided appropriate instrumentation and procedures are in place (ICAO, 2009). Most countries, however, still require at least 5000 ft (1525 m) of separation between runway centerlines for independent simultaneous approaches to a pair of parallel runways.

For independent approaches to three parallel runways under IFR, the FAA requires a 5000-ft (1525-m) separation between the centerline of the middle runway and the centerlines of each of the outer runways (FAA, 1989). An approved FAA aeronautical study is required before authorizing triple approaches at airports located above 1000 ft (305 m) mean sea level.

Separation Requirements for Aircraft on Intersecting, Converging, or Diverging Runways

When it comes to runways that either intersect physically or converge/diverge (so that the projections of their centerlines intersect), the applicable operating procedures and separation requirements vary from airport to airport and from country to country. Examples of the considerations involved include the location of the intersection of the runways, the angle between their centerlines, the mix of aircraft and of movements on each runway, and the local missed-approach procedures. It is therefore impossible to provide a general summary, analogous to Table 10.3, for the separation requirements that apply to such cases. However, these requirements can be specified for any set of local conditions of use.

Clearly, the combined capacity of runways that intersect or converge/diverge will vary significantly, depending on all the factors mentioned. The highest capacities for pairs of runways that intersect physically are usually achieved when the intersection is at the very beginning of both runways in the direction of operations.[10] A pair of intersecting runways can then provide as much capacity as a pair of close-spaced parallels or even medium-spaced parallels, under some modes of use. At the opposite extreme, when crosswinds preclude operations on one of the two runways, their capacity will be the same as that of a single runway.

Visibility, Ceiling, and Precipitation

Airport capacity is affected in critical ways by weather conditions. Cloud ceiling and visibility are the two parameters that determine the weather category in which an airport operates at any given time. Figure 10.3 shows a classification of weather conditions according to these two parameters at a typical airport in the United States. The designations “VFR,” “MVFR,” “IFR,” and “LIFR” for the various regions shown are informal, but widely used in practice. MVFR stands for “marginal VFR” and LIFR for “low IFR.”

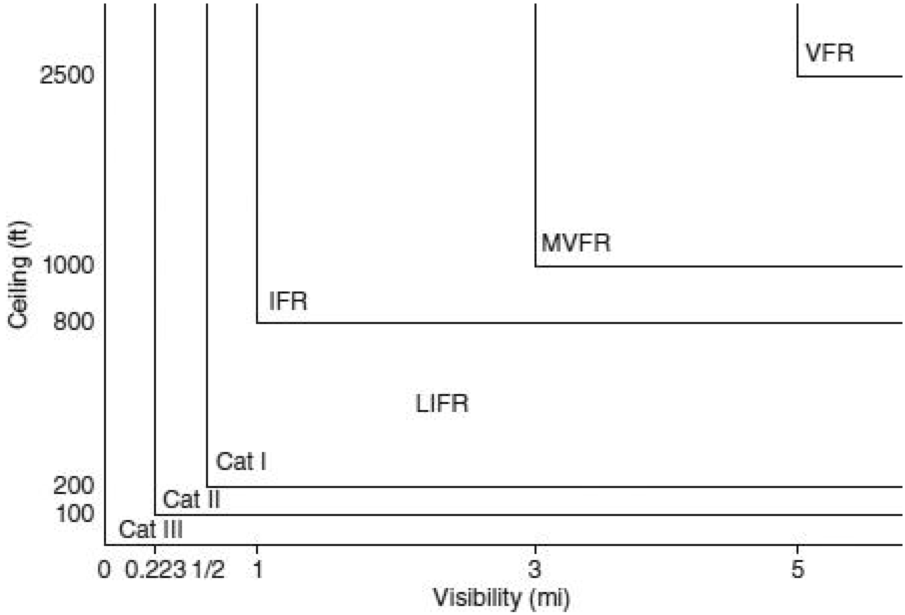


FIGURE 10.3 A typical classification of weather conditions (ceiling and visibility) at an airport in the United States. For CAT II operations the minimum visibility is 1200-ft runway visual range (RVR) or approximately 0.223 statute miles.

The regions denoted as VFR, with a cloud ceiling of 2500 ft (762 m) or higher and visibility of 5 miles or more, and as MVFR are associated with visual meteorological conditions (VMC). The other two correspond to instrument meteorological conditions (IMC) of increasing severity. Note that Category I, II, and III conditions (see Chap. 13) are all part of LIFR. Depending on the instrumentation of the runways and on local topography, different approach, spacing, and sequencing procedures may be used under the ceiling/visibility combinations associated with MVFR, IFR, and LIFR. This means that airport capacity may also change considerably.

An important example of these effects on runway capacity is the (frequent) use of visual separations on final approach at major airports in the United States. At commercial airports outside the United States, IFR separations, such as those in Tables 10.1 and 10.3, are always maintained—at least officially—between landing and/or departing aircraft, regardless of prevailing weather conditions. However, in the United States, under VMC, pilots are often requested by air traffic controllers to maintain visually a safe separation from preceding aircraft during the final spacing and final approach phases of flight. This practice results in higher capacities per runway than can be achieved with strict adherence to IFR.

Equally important, it allows for more efficient use of parallel runways than suggested by Table 10.3.

Boston/Logan illustrates this last point well (right side of Fig. 10.4). Procedures have been established that allow simultaneous, parallel landings in VFR weather on runways 04L and 04R, which are separated by only approximately 1600 ft (490 m). Typically, non-jets land on 04L and practically all jets on 04R. These procedures have been extended for use in MVFR weather, as well.

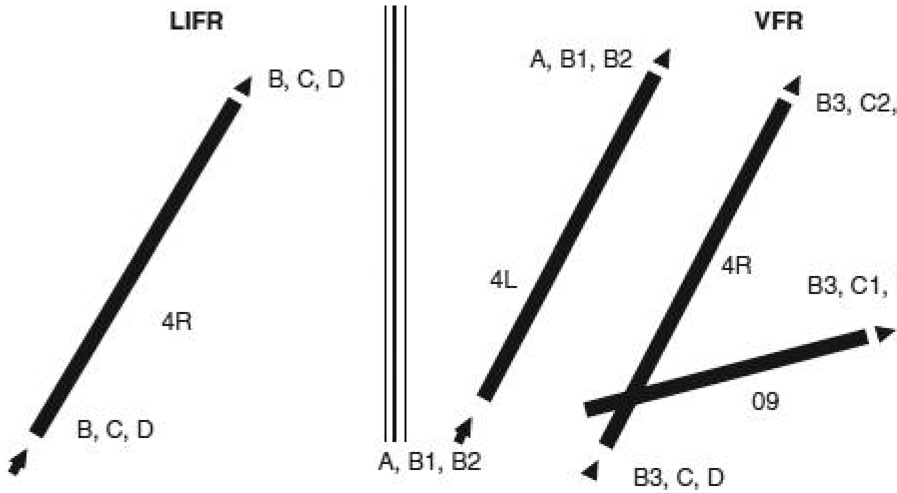


FIGURE 10.4 Two configurations at Boston/Logan with the same orientation but under different weather conditions. The notations A, B, B1, etc., indicate classes of aircraft assigned to each runway. For example, C1 are narrow-body jets on short- and medium-range flights.

More generally, FAA procedures in VMC allow for parallel landings and takeoffs on pairs of parallel runways whose centerlines are separated by as little as 700 ft (214 m) when the runways are used by aircraft in airplane design Groups I through IV, and by 1200 ft (366 m) when aircraft in airplane design groups V and VI are involved.[11] It is for such reasons that the capacity of major airports in the United States under VMC is, in practice, considerably higher than would have been predicted if one applied strictly the IFR separation requirements of Tables 10.1 and 10.3. Some busy European airports (and, more recently, a few in Asia and Oceania, as well) have moved toward applying similar procedures in VMC, under certain conditions, and have consequently gained capacity.

A second example of the effect of weather is provided by operations under low-ceiling and low-visibility conditions (LIFR). A first and obvious effect is that only certain runways, those equipped with a qualified instrument landing system (ILS), can be used under such conditions. For example, Runway 04L at Boston/Logan cannot be used for arrivals in LIFR conditions, as it is not equipped with an ILS, in part because of its proximity to 04R. This means that, when operations are to the northeast in LIFR, Boston/Logan may operate with only one arrival runway, 04R, which must accommodate all aircraft (left side of Fig. 10.4). A second effect is that, to minimize interference with the ILS signal that each aircraft receives (see Chap. 13), separations between aircraft landing consecutively on the same runway in Categories II and III weather are typically increased from those shown in Table 10.1—and can become as large as 9 nmi or several minutes between successive movements. This, of course, reduces dramatically the arrival capacity under these conditions.

Finally, precipitation and icing may affect severely the capacity of runways because of poor visibility, poor braking action, and the need for aircraft deicing. For example, when braking action is poor, the crosswind limits (see Sec. 9.3) for approaches to a runway may be reduced. More extreme weather events, such as snowstorms and thunderstorms, often lead to the temporary closing of an airport's runways.

Wind Direction and Strength

Winds may also affect airport capacity in crucial ways. As explained in Sec. 9.3, a runway can be used only when crosswinds are within prescribed limits and tailwinds do not exceed 5 or 6 knots (9–11 km/h). This means that the orientation of runway operations and, more generally, the availability of runways largely depend on the direction and strength of the prevailing winds at any given time. At locations that may experience strong winds from several different directions at different times, this can be the cause of considerable variability in the available capacity of the runway system. Boston/Logan again provides a good example. With strong winds from the northeast or the southwest, the airport usually operates in VMC with two arrival runways—04L and 04R when operations are to the northeast, 22L and 27 when operations are to the southwest (Fig. 10.5). However, with strong winds from the northwest, only one runway, 33L, was truly available for arrivals until 2007, because runway 33R is very short (2300 ft, ~ 700 m) and can be used by only some nonjet aircraft (Fig. 10.6). This meant that with strong northwest winds, aircraft arrivals at Boston/Logan, even in VMC, experienced severe delays that exceeded 2 hours on some days. As a result, a new runway, 14/32 was opened in 2007 at Boston/Logan, with its centerline approximately 4300 ft away from the centerline of 15R/33L. In the 32 direction, this new runway (which is only 5000 ft, but still more than twice as long as 33R) provides a nearly parallel and quasi-independent approach to the approach to Runway 33L, thus increasing greatly the capacity of Boston/Logan under the conditions shown in Fig. 10.6. This has resulted in a significant reduction of delays at the airport when strong northwesterly winds prevail.

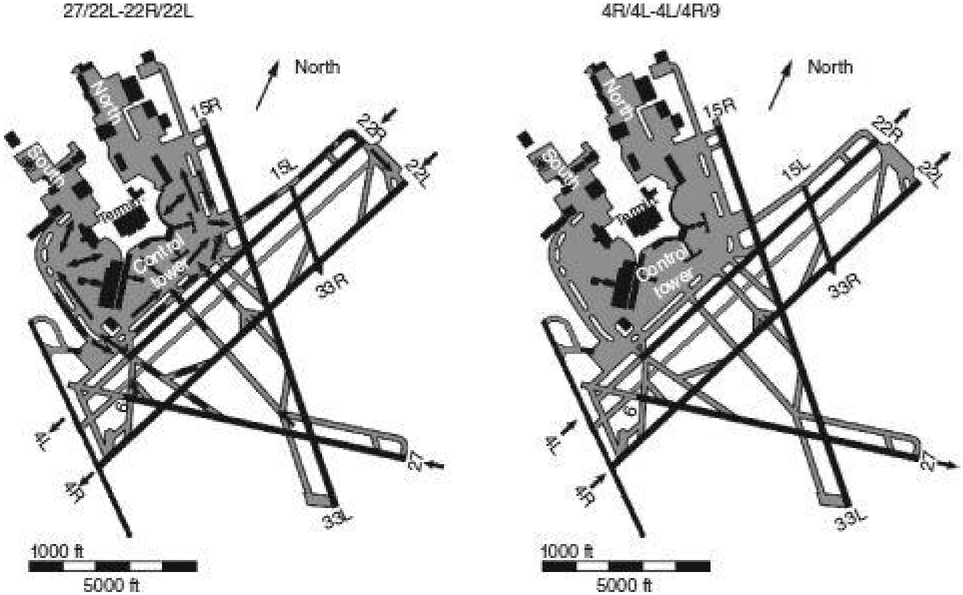


FIGURE 10.5 Two high-capacity configurations with opposite orientations at Boston/Logan. The configuration on the left is Configuration 9; the one on the right Configuration 1.

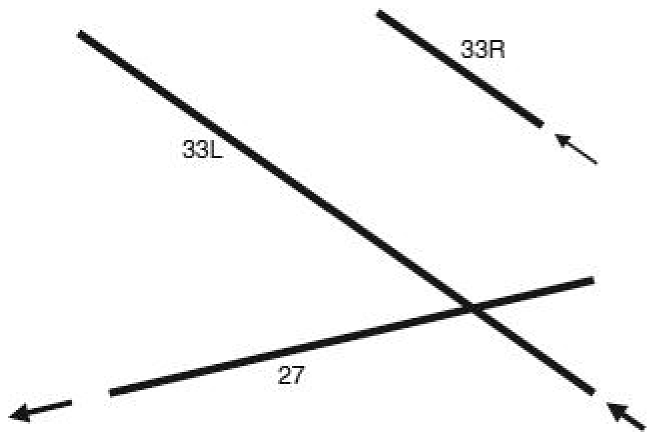


FIGURE 10.6 A low-capacity configuration in VMC at Boston Logan that motivated construction of a new runway.

When wind speed is less than 5 knots ("calm"), air traffic controllers have considerable latitude as to which runways will be used (if more than one exist) and in which direction. Such decisions must be made often, as calms prevail a large percentage of the time at most airports. To select the active runway(s) and the direction(s) of operations on these occasions, a combination of criteria such as maximizing runway capacity or minimizing noise impacts may be used (see also as follows and Chap. 6).

Mix of Aircraft

Tables 10.1 and 10.3 suggest why the mix of aircraft is another important factor in determining runway capacity. Consider, for example, a runway used only for arrivals and assume that the mix of aircraft consists of 50 percent "heavy" (H) and 50 percent "small" (S). Everywhere in the world, arriving aircraft are currently sequenced for access to a runway according to a first-come, first-served (FCFS) queue discipline.[12] With FCFS sequencing, about 25 percent of aircraft pairs [(0.5)·(0.5) = 0.25] in our example will then be "H followed by S" (H-S) pairs and, according to Table 10.1, will be separated by 6 nmi at the runway threshold. Another 25 percent of the pairs will be separated by 4 nmi (H-H), and 50 percent will be separated by 2.5 nmi (S-H and S-S). By contrast, if the traffic consisted of 80 percent L-class aircraft and 20 percent S-class, 16 percent [(0.8)·(0.2) = 0.16] of all possible pairs of aircraft (i.e., the L-S pairs) would require a separation of 4 nmi and the other 84 percent would require 2.5 nmi. The runway capacity, as measured by the expected number of landings performed per hour, will then clearly be considerably greater in the second case than in the first, as can be confirmed by using the simple mathematical model of Sec. 10.5.

In general, a relatively homogeneous mix of aircraft (i.e., a mix consisting of one or two dominant classes) is preferable to a nonhomogeneous mix from the point of view of runway capacity. Moreover, a homogeneous mix also offers advantages for ATM purposes, as it simplifies the work of air traffic controllers, who have to make fewer adjustments for wake vortex separations, for different approach speeds, and for other aircraft characteristics. In fact, when the mix of aircraft is very nonhomogeneous, air traffic controllers at multirunway airports often attempt to "segregate traffic" by assigning different aircraft classes to different runways.

This also explains why the combined capacity of two independent parallel runways, if operated well by the ATM system, can provide more than twice the capacity of a single runway: the two runways provide an opportunity to optimize the assignment of aircraft types to each runway, as well as the mix and sequencing of movements (landings and/or departures) on each runway—see as follows.

Mix and Sequencing of Movements

Another factor that influences runway capacity is the mix of movements (arrivals versus departures) at the airport as a whole, and on each runway separately. For most, but not all, ATM systems, separation requirements are such that the capacity of a runway that is used only for departures is higher than that of a runway that is used only for arrivals, given the same mix of aircraft. At some major airports in the United States, as many as 60 departures may be performed in 1 hour from a single runway when the traffic mix includes only a very small percentage of aircraft in Class H. By contrast, it is difficult to perform more than 45 arrivals per hour per runway with any aircraft mix consisting primarily of commercial jets.

At busy airports there are typically some periods of the day when arrivals dominate and vice versa.[13] Hub airports, in particular, experience surges of arrival activity several times a day, followed sometime later by surges of departures.[14] The capacity of the airport may vary correspondingly. For instance, the number of runway movements per hour that can be performed at New York/Kennedy during the early afternoon hours, when many flights from Europe arrive, is significantly smaller than can be performed late in the evening, when a similar number of flights depart for Europe.

A related issue is the assignment of arrivals and departures to runways at airports operating with more than one active runway. When given the opportunity, air traffic controllers often prefer to use separate runways for arrivals and for departures. This is especially common at European and Asian airports that operate with two parallel runways, as several do. This practice may simplify ATM operations, but is not necessarily optimal as far as overall airport capacity is concerned. It may overload one runway and underutilize another at times when the number of arrivals differs significantly from the number of departures. This may also create a serious imbalance between the delays experienced by arrivals versus those experienced by departures. In fact, a better way to operate an airport with two parallel runways is to assign, when feasible, some arrivals to a runway used primarily for departures, whenever arrivals “overflow” their primary runway, and do the reverse whenever there is an excess of departures in the mix. It may be even more efficient to mix arrivals and departures on two or more runways at airports where the ATM system is sufficiently advanced to sustain this mode of operation well. Munich Airport, with its two independent parallel runways, as well as several airports in the United States, achieve high processing rates through such a mixed runway-use strategy. The mathematical models discussed in Secs. 10.5 and 10.6 can be helpful in assessing the benefits that can be obtained from alternative assignments of operations to runways for any particular set of local conditions.

The sequencing of movements on a runway also influences runway capacity, especially whenever a runway is used for mixed operations (arrivals and departures). As noted earlier, arriving aircraft are generally sequenced in roughly FCFS order for access to a runway, and so are departing aircraft. Air traffic controllers, however, have considerable latitude regarding the sequencing of arrivals versus departures on the runway. It is possible to maintain an approximate FCFS discipline and sequence arrivals and departures roughly according to the time when they can first make use of the runway—the earlier the time, the higher the priority.[15] More typically, though, when vying for use of the same runway(s), arrivals are given priority over departures for reasons of safety, controller workload, and aircraft operating cost. However, the strictness with which this practice is applied in practice varies considerably from one ATM system to another and from airport to airport. Quite often, for example, air traffic controllers will process a string of several consecutive landings until the queue of arriving aircraft is practically exhausted and will then process a string of several consecutive departures. Air traffic controllers will also look for some “free departures”; that is, they will try to insert one or more departures between two arrivals without seriously disturbing the arrival stream and, thus, without reducing the arrival processing rate. This can often be done when there is a long gap between two landing aircraft, for example due to a 6-nmi separation between a leading aircraft of type H and a trailing one of type S. There are also occasions when a long queue of departures may form on the ground because the runway is continually busy with arrivals. In such instances, ATC may decide to interrupt the arrival stream for a while, assigning temporary priority to takeoffs until the departure queue returns to a reasonable length.

Finally, alternating arrivals and departures on the runway can be a very effective strategy for maximizing overall runway capacity, as measured by the total number of movements per unit of time. This sequencing strategy can be implemented by “stretching,” as necessary, the separation between a pair of consecutive arriving aircraft, in order to create a gap that is just sufficiently long to allow insertion of a departure between the two arrivals. In a number of countries, ATM separation requirements make it possible to achieve such insertions with only a relatively modest amount of stretching of the required A-A separations. Indeed, this happens to be the case with the separation requirements shown in Table 10.1 for the United States. Thus, by “sacrificing” only a modest amount of arrival capacity per unit of time, the number of departures served by the runway per hour becomes roughly equal to the number of arrivals. However, this type of separation-stretching procedure is more demanding from the ATM point of view and requires skilled air traffic controller teams. Thus, its application is still limited primarily to some of the busiest airports of the United States and of Europe.

Type and Location of Runway Exit

The *runway occupancy time* of an arriving aircraft is defined as the time between the instant the aircraft touches down on the runway and the instant it is on a runway exit, with all parts of the aircraft clear of the runway. Because the location of runway exits (“exit taxiways”) plays a significant role in determining runway occupancy times, it may also have an impact on runway capacity. In particular, it can be seen from Table 10.1 that reducing runway occupancy times will contribute to increasing runway capacity in the following:

- The A-D case, where the earlier the arriving aircraft leaves the runway, the earlier the trailing departure's takeoff run can begin, provided the departing aircraft is set to go
- The A-A case, but only if the requirement that two arriving aircraft should not occupy the same runway simultaneously is the more restrictive of the two requirements listed in Table 10.1—the other requirement being the longitudinal separation of 2.5, 4, 5, or 6 nmi on final approach.[16]

Well-placed high-speed exits can be helpful in reducing runway occupancy times and increasing capacity. However, as noted in Sec. 9.7, the cost of constructing a high-speed exit may be considerably higher than that of a conventional exit forming a 90° angle with the runway centerline. When the benefits, in terms of more runway capacity, of a high-speed exit are compared with this additional cost, it is difficult to justify the construction of more than two or three high-speed exits for any single direction of runway operations.

In general, it is useful to remember that high-speed exits offer essentially no capacity benefits at runways used only for departures, limited capacity benefits at runways used only for arrivals (primarily at those airports where visual separations are in use under VMC), and significant capacity benefits, under some movement-sequencing strategies, at runways used in a mixed operations mode.

State and Performance of the ATM System

A high-quality ATM system with well-trained and motivated personnel is a fundamental prerequisite (but not a sufficient condition by itself) for achieving high runway capacities. To use a simple example, tight separations between consecutive aircraft on final approach (i.e., separations that are as close possible to the minimum required in each case) cannot be achieved unless (1) accurate and well-displayed information on the positions of the leading and trailing aircraft is available to air traffic controllers, and (2) the controllers themselves are skilled in the task of spacing aircraft accurately during final approach. ATM systems for airports and for terminal airspace are reviewed in Chap. 13.

Air traffic controllers are the core element of ATM systems and will continue to be so in the more advanced ATM systems currently being planned for the next 20 or more years. Human factors and ergonomics therefore play a central role in determining airport capacity. Air traffic controllers at most of the busiest airports in the world are highly qualified and, typically, well-paid professionals. The synergy between air traffic controllers and aircraft pilots is also very important. If air traffic controllers perceive that a pilot is inexperienced or has difficulty understanding instructions, they will slow down operations considerably to allow for additional margins of safety, thus reducing airport capacity.

Environmental Considerations

Last, but certainly not least, environmental considerations, especially noise impacts, exert an important influence in determining runway system capacity at an ever-growing number of airports. In the daily course of airport operations, noise is one of the principal criteria used by air traffic controllers to decide which one among several usable alternative runway configurations to activate. As indicated earlier, a choice among two or more alternative configurations exists whenever weather and wind conditions are favorable. As a simple example, at a single-runway airport, air traffic controllers can choose to operate in either of the two directions of the runway when the weather is fair and there is little wind. The noise impacts associated with each direction will then often be the principal criterion that will determine the choice between the two.

Noise-related considerations work, in general, as a constraint on airport capacity, because they tend to reduce the frequency with which certain high-capacity configurations may be used. Example 10.2 illustrates the types of noise-related restrictions and configuration-selection practices that one increasingly encounters at major airports worldwide.

Example 10.2 At Boston/Logan, no turbofan or turbojet departures are permitted on Runway 04L, except in special cases, despite the fact that this runway is sufficiently long to accommodate the landing and takeoff requirements of most jet flights. The reason for this policy is noise mitigation for densely populated areas under the takeoff paths at the end of Runway 04L. Few jet arrivals are also assigned to 04L, again to avoid noise-related complaints from airport neighbors living under the 04L approach paths. Similarly, jet landings are generally not permitted on Runway 22R, because of noise considerations.

Another type of noise-related constraint with an impact on capacity at Boston/Logan takes the form of a set of long- and short-term goals for runway utilization and related restrictions. Specifically, the Massachusetts Port Authority ("Massport"), owner and operator of Boston/Logan, has agreed with representatives of the communities surrounding the airport on the following noise-related operating guidelines:

- Annual goals have been set for the utilization of every runway end. The goals are stated in terms of the desired percentage of "effective jet operations" that should be performed annually over each runway end (nonjets are not considered). The number of effective jet operations is obtained by multiplying the number of nighttime (22:00–07:00) operations by 10 and then adding this product to the number of operations during the rest of the day. For instance, a goal might state that Runway 33L may be used for 42 percent of effective jet arrivals and 12 percent of effective jet departures in a year. The reason for the high goal of 42 percent would be that aircraft landing on 33L approach the airport over the Atlantic Ocean and thus have little noise impact on neighboring communities. The overall objective of the annual goals is to "distribute" noise among neighboring communities in a way that is considered fair by the parties involved.
- No runway can be used continually for more than 4 hours in any single direction. This so-called "persistence" restriction is aimed at preventing the continuous exposure of any single community to noise on any particular day.
- No runway can be used for more than 24 hours in any 72-hour period. This restriction, too, is intended to prevent excessive, even if intermittent, exposure of a community to noise within a relatively short time span of 3 consecutive days.

It should be noted that these noise-related restrictions are applied only if weather conditions permit. When weather conditions are unfavorable, air traffic controllers have little or no choice as to the runway configuration to be used.

The overall effect of these restrictions is to inject noise as the second criterion (in addition to making optimal use of runway capacity) in the selection and use of active runways. For example, during periods when weather permits a choice among a number of alternative runway configurations, air traffic controllers may elect to use that configuration which will bring Logan closer to meeting the annual goals for the use of the runway ends, rather than the configuration that will offer the highest feasible capacity. Indeed, this is very often the case, especially at times when demand is relatively low. While adherence to the three restrictions mentioned previously is currently voluntary, community and Massport representatives meet regularly to review how well the airport meets each of its environmental goals.

10.4 Range of Airfield Capacities and Capacity Coverage

The capacities of runway systems of major airports around the world span a wide range. Some single-runway commercial regional airports in Europe have a capacity as low as 12 (!) movements per hour, because of inadequate air traffic control systems or other local factors. At the opposite end, a few airports in the United States operate with as many as 4 to 7 simultaneously active runways and accommodate more than 200 (and in the case of Dallas/Ft. Worth close to 300) movements in 1 hour. At locations with reasonably advanced ATM systems—and absent noise-related or other restrictions—capacities range from about 24 per hour per runway to as many as 60, depending on the many factors that were reviewed in the previous section. Airports in the United States are typically at the high end of the capacity-per-runway range. At many airports, the capacity of the runway system may also be highly variable over time, primarily due to sensitivity to weather and wind conditions.

To illustrate these points, Table 10.4 shows the FAA's estimates in 2004 of the maximum throughput capacities of 34 of the busiest airports in the United States under optimum weather conditions and under weather conditions that lead to reduced capacity.[17] In the former case, VMC permit (a) visual separations between landing aircraft and (b) procedures that result in reduced separation requirements between aircraft operating on different runways. In the latter, IMC necessitate IFR separations. The capacities shown are for the most commonly used runway configurations under these conditions. Note that 24 of the 34 airports had a capacity under optimum conditions that exceeded 100 movements per hour. In contrast, very few airports outside the United States (only Amsterdam/Schiphol, Paris/de Gaulle, Madrid/Barajas, and Toronto/Pearson, as of 2011[18]) could regularly handle more than 100 movements per hour! Similarly, very few airports outside the United States ever operate with three or more simultaneously active runways.

Airport	Conditions	
	Optimum	Reduced
Atlanta	180–188	158–162
Boston/Logan	123–131	90–93
Charlotte	130–131	102–110
Chicago/Midway	64–65	61–64
Chicago/O'Hare	190–200	136–144
Cincinnati	120–125	102–120
Cleveland	80–80	64–64
Dallas/Ft. Worth	270–279	186–193
Denver/International	210–219	159–162
Detroit/Metro	184–189	136–145
Houston/Bush	120–143	108–112
Las Vegas	102–103	70–70
Los Angeles/Internat'l	137–148	117–124
Memphis	148–181	120–132
Miami/Fort Lauderdale	60–62	52–56
Miami/International	116–121	92–96
Minneapolis/St. Paul	114–120	112–114
Newark	84–92	61–66
New York/Kennedy	75–87	64–67
New York/LaGuardia	78–85	69–74
Orlando/International	144–164	104–117
Philadelphia	104–116	96–96
Phoenix	128–150	108–118
Pittsburgh	152–160	119–150
Portland	116-120	77–80
Salt Lake City	130–131	110–113
San Diego	56–58	48–50
San Francisco/Internat'l	115–110	68–72
Seattle/Tacoma	80–84	57–60
St. Louis	104–113	64–70
Tampa	102–105	74–75
Washington/Baltimore	106–120	60–71
Washington/Dulles	135–135	105–113
Washington/Reagan	72–87	48–70

Source: FAA, 2004.

Table 10.4 Approximate Capacities of 34 of the Busiest Airports in the United States

The difference between the optimum and the reduced capacities for some of the airports in Table 10.4 is also remarkable. In several cases, this difference is of the order of 30 percent or more. As explained in Sec. 10.3 (see discussion of the effects of visibility, ceiling, and precipitation), these large differences stem primarily from the geometric configuration of the runways that forces a reduction in the number of runways that can be operated simultaneously for approaches in IMC, as well as from the use of IFR in IMC. Note that certain airports, such as Cincinnati and Minneapolis/St. Paul, have runway configurations that are little affected by reduced visibility.

A particularly convenient way to summarize the range of capacities at an airport and the frequency with which various levels of capacity are available is the CCC. The CCC shows how much runway capacity is available for what percentage of time at a given airport under the assumptions that (1) the operations mix is 50 percent arrivals and 50 percent departures and (2) the runway configuration in use at any given time is the one that provides the highest capacity under the prevailing weather conditions.

Example 10.3 Figure 10.7 shows the CCC of Boston/Logan Airport. A maximum capacity of 132 movements per hour is available for approximately 60 percent of the time, a maximum capacity of 120 movements per hour for approximately 18 percent of the time, and so on. The airport's capacity declines to the range of 60 to 70 movements per hour—that is, to about half of the peak capacities of 132 and 120 movements—for approximately 15 percent of the time. For approximately 1.5 percent of the time, when the airport is closed due to snowstorms or severe thunderstorms, the capacity is zero. This CCC is obtained by looking at historical statistics regarding the frequency with which each of the possible combinations of visibility, ceiling, and wind conditions at Boston/Logan occur during the course of a year and identifying the runway configuration that provides the highest capacity for each set of weather conditions. For instance, Configuration 1, which consists of arrivals on runways 4R and 4L and departures on runways 4R, 4L, and 09 in VMC (on the right in Fig. 10.5), is the Boston/Logan configuration with the highest capacity. When the mix of movements is 50 percent arrivals and 50 percent departures, this capacity has been estimated at 132 movements per hour, a number that can be obtained either from empirical data or from mathematical or simulation models such as the ones described in Secs. 10.5 and 10.6. Weather records indicate that Configuration 1 can be used approximately 60 percent of the time. From the assumption that the available configuration with the highest capacity will be selected at all times [assumption (b) in the definition of the CCC], it follows that Configuration 1, will be used whenever possible. This is shown in Fig. 10.7, where a capacity of 132 movements per hour for 60 percent of the time is indicated at the left-hand part of the CCC, along with the indication that this capacity is associated with Configuration 1. Proceeding now toward the right in Fig. 10.7, it can be seen that, when weather conditions do not permit use of Configuration 1, the configuration with the next highest capacity, 120 movements per hour, is Configuration 9 (shown in Fig. 10.5 on the left), which consists of arrivals on runways 22L and 27 and of departures on runways 22R and 22L in VMC. From weather statistics at Boston/Logan, the percentage of time when Configuration 9 is available and Configuration 1 is not is equal to 18 percent. This is once again shown in Fig. 10.7. Continuing in the same way toward the right part of the CCC, one encounters ever-smaller capacities, as the airport "runs out" of high-capacity configurations, until "100 percent of the time" is accounted for.

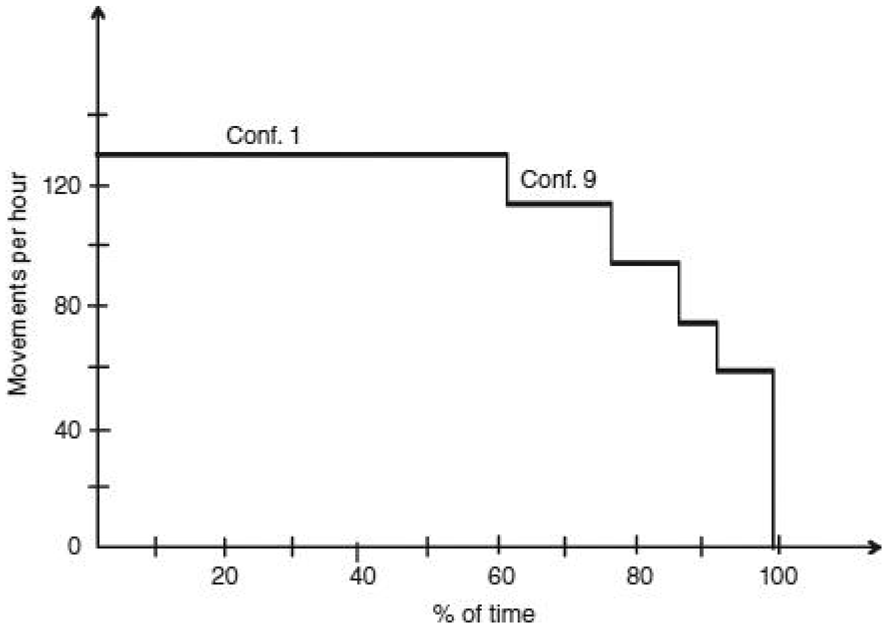


FIGURE 10.7 Capacity coverage chart for Boston/Logan.

The CCC obviously provides very useful information for airport planners and managers. However, it is also important to keep in mind its underlying assumptions. By assumption (a) in the definition of the CCC, the capacities of 132, 120, etc., for each configuration in Example 10.3 are computed for an operations mix of 50 percent arrivals and 50 percent departures. By assumption (b), the operator of the runway system (i.e., the FAA or other ATM service provider) is presumed to choose at all times the available runway configuration with the highest capacity. For instance, as noted previously, in VMC and with calm winds, both Configurations 9 and 1 (see Fig. 10.5) are available for use at Boston/Logan. According to assumption (b), Configuration 1 will always be selected in such cases, because it has the higher capacity.

In practice, assumptions (a) and (b) are only rough approximations to reality. The mix of movements is rarely exactly 50 percent arrivals and 50 percent departures. When the mix is significantly different from that (e.g., 65 percent arrivals and 35 percent departures), the capacity of the runway system may also differ significantly from the number indicated on the CCC. Fortunately, the operations mix at busy nonhub airports during peak hours typically falls in the range between 40 percent arrivals, 60 percent departures and 60 percent arrivals, 40 percent departures. Therefore, the capacities indicated under the 50–50 percent assumption are usually fairly representative of the capacities available during peak hours. However, at hub airports, where waves of arrivals are followed by waves of departures, the CCC may have to be supplemented by an analysis of the capacity to handle these surges of arrivals and departures. Mathematical or simulation models can be used for this purpose.

Regarding assumption (b), noise considerations may dictate use of a configuration other than the one with the highest capacity, especially during hours when demand is not at its peak, as already seen in the previous section. In the Boston/Logan case, Configuration 9 is often chosen over Configuration 1 during periods when they are both available, to "distribute the noise more equitably" among the airport's neighboring communities and meet annual noise-related goals (see Example 10.2). In this light, the CCC can better be viewed as showing the upper limit of how much runway capacity is available at an airport over time.

This last point is underscored by Fig. 10.8 (Idris, 2001), which summarizes the usage of runway configurations at Boston/Logan during January 1999. A very low-capacity configuration[19] that uses Runway 33L for arrivals and Runway 15R for departures (i.e., the same runway in opposite directions) is utilized heavily during the six hours of 00:00 to 05:59. This is because the configuration in question has the least noise impact of any at Boston/Logan, as both arrival and departure paths are over the sea and avoid populated areas. That the capacity of this configuration is low does not really matter, because traffic is also very low during the period when it is used. An intermediate-capacity configuration with arrivals on Runways 33L and 33R and departures from Runway 27 is used quite intensively during the early morning to noon hours, when traffic demand is not very heavy (see Fig. 10.8). Finally, the two highest-capacity configurations that figure so prominently in the CCC, Configurations 9 and 1, are utilized very heavily during the peak traffic hours between 14:00 and 21:00, when their high capacity is truly needed. Figure 10.8 confirms that the noise impact of a runway configuration is often the dominant selection criterion during periods of low demand, whereas the CCC is a good indicator of what configurations will be used during peak demand periods.

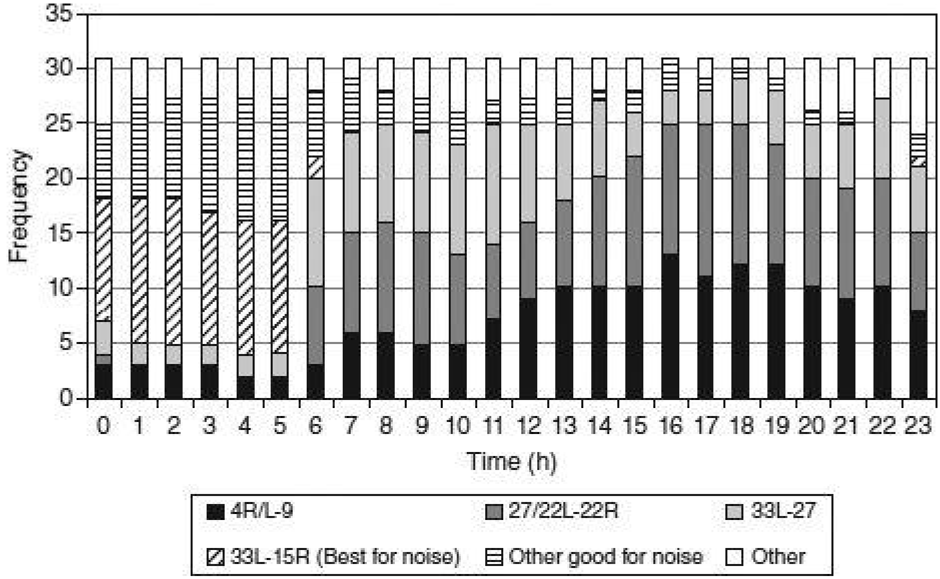


FIGURE 10.8 Runway configuration usage at Boston/Logan, January 1999 (from Logan FAA tower logs).

Thus, the CCC essentially provides a summary description of the relative frequency with which different values of capacity are available at an airport during periods of high demand. For complex, multirunway airports, the computation of the CCC may require considerable effort. For example, in the case of Boston/Logan one needs to compute the capacity of each of the more than 20 different runway configurations, along with the percentage of time when each is available on the basis of historical weather/wind data.

An "uneven" CCC, like the one of Fig. 10.7, indicates an airport where the supply of runway capacity is not reliable. This may result in long delays and serious operational problems when the typical demand levels during peak periods are close to the higher-capacity values of the CCC. Consider again Boston/Logan. Because the airport's capacity is 120 or more for about 78 percent of the time, airlines have consistently been scheduling 100 or more movements per hour for several hours each day during the peak summer season.[20] This means that for about 22 percent of the time (see Fig. 10.7), or on one out of every 5 days on average, the available capacity may fall considerably short of demand during peak periods, resulting in serious delays. On truly bad days, when the capacity may be 60 or lower for several hours in a row, very long delays and many flight cancellations occur. An extreme alternative to this scheduling practice would be to restrict airport demand to a low level, for example, to a maximum of 60 movements per hour. This would guarantee that demand is always (or almost always) exceeded by available capacity. While this ensures the virtual absence of delays and a high level of service, it also means wasting a great amount of available capacity for 80 or 90 percent of the time. This type of dilemma is discussed further in Chap. 12, which covers the subject of airport demand management.

A “flat” (or “even”) CCC, on the other hand, is characteristic of airports where the runway capacity stays relatively constant over time. For example, a single-runway airport, which almost always operates under good weather conditions, would have an almost completely even CCC for essentially 100 percent of the time. A flat CCC means more predictable airside performance and more effective utilization of airport resources and facilities, as the number of operations at the airport can be scheduled with reference to a stable level of runway capacity.

Recall now that runway capacity is defined as the expected (“average”) number of movements that can be handled per hour. This is what the CCC shows. In practice, the actual number of movements that can be performed during each hour may be greater or less than the expected number shown. For example, instead of the capacity of 132 movements per hour shown in Fig. 10.7, the actual number performed during a particular hour when Configuration 1 is in use may be 127 or 140, depending on the exact traffic mix during that hour, the performance of the team of air traffic controllers on duty at the time, the strength and variability of the prevailing winds, etc.

The reader who is familiar with probability theory will recognize that the CCC is essentially a graphical representation of the probability distribution of an airport’s (maximum throughput) runway capacity. Figure 10.7 indicates that, at any randomly chosen instant, Boston/Logan’s maximum available runway capacity will be equal to 132 movements per hour with probability 0.6, to 120 movements per hour with probability 0.18, etc., and to 0 with probability 0.015.

10.5 A Model for Computing the Capacity of a Single Runway

In addition to understanding qualitatively the definitions and complex relationships that determine capacity, it is essential in practice to have access to computing tools that provide reasonably accurate estimates of the capacity of runway systems under any set of specified conditions. A number of mathematical and simulation models have been developed over the years that make this possible. In this section one such mathematical model will be described in some detail because, despite its many simplifying assumptions, it offers insights into the physical process that drives runway capacity, as well as yields good approximations to the capacities observed in practice. The model is also particularly convenient for sensitivity analyses that explore the effects on capacity of many of the factors reviewed in Sec. 10.3. Finally, this model illustrates well the conceptual approach taken by virtually all the computer-based mathematical models that are now used widely to estimate capacity and delays at runway complexes.

This simple mathematical model is originally due to Blumstein (1959). It estimates the capacity of a single runway used solely for arrivals. The same approach, however, can be readily extended to runways used solely for departures or runways used for mixed movements.

Consider a single runway, shown schematically in Fig. 10.9, which is used for landings only. Aircraft descend in single file along the final approach path until they touch down on the runway, whereupon they decelerate and exit onto the taxiway system. The paths of arriving aircraft merge in the vicinity of the "gate" to the final approach, typically 5 to 8 nmi from the runway threshold. Throughout the final approach, aircraft must maintain a safe longitudinal distance from each other, in compliance with the ATM system's separation requirements, as explained in Sec. 10.3. Moreover, single occupancy of runways is required (see Table 10.1): each aircraft must be safely out of the runway before the next landing can touch down. These safety rules impose limits on the maximum acceptance rate of the runway, that is, on its maximum throughput capacity.

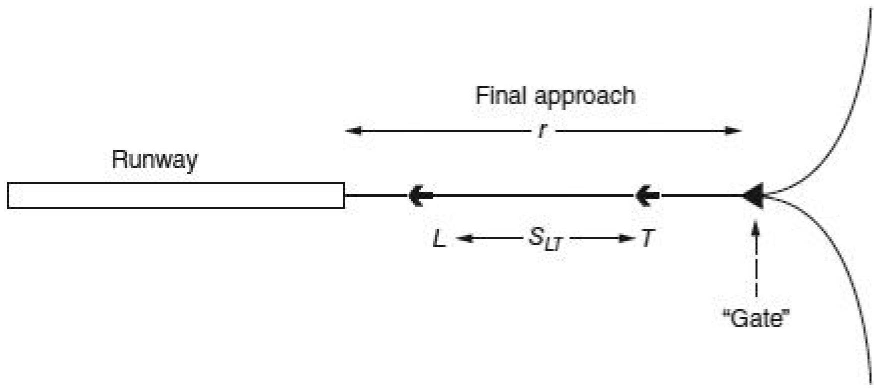


FIGURE 10.9 A simple representation of a runway used for arrivals only under IFR.

Define now the following quantities:

r = the length of the common final approach path

v_i = the speed on final approach of an aircraft of type i assuming, as a reasonable approximation, that the aircraft maintains a constant speed throughout the approach

o_i = the runway occupancy time of an aircraft of type *i*, that is, the time that elapses from the instant when the aircraft touches down on the runway to the instant when it leaves the runway at one of the runway exits

Consider the case in which an aircraft of type *i* is landing, followed immediately by another aircraft of type *j*. Denote by s_{ij} the minimum separation required by ATC between the two aircraft while they are both airborne. For example, in Fig. 10.9, s_{LT} indicates the minimum separation required between a leading aircraft of type L and a trailing aircraft of type T. Let T_{ij} denote the minimum possible time interval between the successive arrivals at the runway of the type *i* and type *j* aircraft, that is, the minimum time separation between the two landings that can be achieved without violating any ATM separation requirements. The two fundamental equations that determine T_{ij} can be written as follows:

$$T_{ij} = \max\left[\frac{r+s_{ij}}{v_j} - \frac{r}{v_i}, o_i\right] \quad \text{when } v_i > v_j \tag{10.1a}$$

$$T_{ij} = \max\left[\frac{s_{ij}}{v_j}, o_i\right] \quad \text{when } v_i \le v_j \tag{10.1b}$$

The situation in which $v_i > v_j$ is known as the "opening case" because the distance between the two aircraft keeps increasing as they fly in single file along the final approach path on the way to the runway. In this case, the two aircraft are closest to each other at the instant when the first of the two, of type *i*, is at the gate of the final approach path, a distance *r* from the threshold of the runway (see Fig. 10.9). If at that instant the two aircraft are separated by the minimum allowable separation s_{ij}, the type *j* aircraft will be a distance $r + s_{ij}$ from the runway. The difference between the times when the leading aircraft (type *i*) and the trailing aircraft (type *j*) will touch down on the runway will then be equal to

$$\frac{r+s_{ij}}{v_j} - \frac{r}{v_i}$$

However, the time interval between the successive arrivals at the runway must also be at least o_i, to allow enough time for the leading aircraft (type i) to exit the runway before the trailing aircraft touches down. The minimum time interval, T_{ij}, between the successive arrivals at the runway is then equal to the larger (maximum) of the quantities

$$\frac{r+s_{ij}}{v_j}-\frac{r}{v_i} \quad \text{and} \quad o_i$$

and is thus given by Eq. (10.1*a*). By contrast, in the closing case $v_i \le v_j$ the two aircraft are closest to each other at the instant when the first aircraft is at the runway threshold. The minimum interval, T_{ij} is then given by Eq. (10.1*b*).

Suppose now that the probability of the event "a type i aircraft is followed by a type j aircraft" is p_{ij}. Then

$$E[T_{ij}]=\sum_{i=1}^{k}\sum_{j=1}^{k}p_{ij}\cdot T_{ij} \tag{10.2}$$

where $E[T_{ij}]$ denotes the expected (or "average") value of T_{ij}, while K is the number of distinct aircraft classes ($K = 4$ in Example 10.4). A numerical example illustrates the application of the model.

Example 10.4 As noted in Sec. 10.3, the FAA subdivides aircraft into three classes for the purpose of determining the separations s_{ij} required on final approach: "heavy" (H), "large" (L), and "small" (S). (The special case of the B757 will not be considered here.) Because different types of aircraft in Class S have quite different approach speeds, this class will be subdivided in this example into two more homogeneous subclasses, S1 and S2, as is often done in airport capacity analyses. Denote the Classes H, L, S1, and S2 with the indices 1 through 4, respectively.

Assume now that, at a major airport, a runway, which is used for long periods of time for arrivals only, serves an aircraft population with the characteristics given in Table 10.5. Note that the probabilities, p_i, indicate the traffic mix at this runway (e.g., 20 percent of the aircraft are of type H and 35 percent of type L). Assume, as well, that the IFR separation requirements, s_{ij}, in use are as shown in Table 10.6. Note these are the same as in Table 10.1 with some simplifications.

I (a/c type)	p_i (probability)	v_i (knots)	o_i (s)
1 (H)	0.2	150	70
2 (L)	0.35	130	60
3 (S1)	0.35	110	55
4 (S2)	0.1	90	50

TABLE 10.5 Data for Example 10.4 (1 knot = 1 nmi/h = 1.15 statute miles/h = 1.852 km/h)

		Trailing Aircraft		
		H	L	S1 or S2
Landing aircraft	H	4	5	6*
	L	2.5	2.5	4*
	S1 or S2	2.5	2.5	2.5

*Indicates that the separation applies when the leading aircraft is at the runway threshold.

TABLE 10.6 Separation Requirements (in nautical miles) on Final Approach for Example 10.4

Let now the length, r, of the final approach path be equal to 5 nmi. Applying Eqs. (10.1a) and (10.1b), one can compute the matrix, **T**, of minimum time separations, T_{ij}, in seconds, at the runway, shown as Table 10.7. To obtain Table 10.7, Eq. (10.1a) has been used to compute T_{12} and T_{34} and Eq. (10.1b) to compute all the other elements of the matrix. Note that the two equations give the same result when it comes to the diagonal elements of the matrix and that Eq. (10.1b) has been used in the cases of T_{13}, T_{14}, T_{23}, and T_{24} because the separation requirement in these cases applies when the leading aircraft is at the threshold of the runway.

		Trailing Aircraft			
		1 (H)	2 (L)	3 (S1)	4 (S2)
Leading aircraft	1 (H)	96	157	196	240
	2 (L)	60	69	131	160
	3 (S1)	60	69	82	136
	4 (S2)	60	69	82	100

TABLE 10.7 Matrix **T** of Minimum Time Separations T_{ij} for Example 10.4

As mentioned in Sec. 10.3, air traffic controllers use FCFS sequencing of aircraft wishing to land at an airport. This makes it reasonable to assume that, for any pair of aircraft, the probability that the leading aircraft will be of type i is simply equal to p_i, the proportion of aircraft of type i in the mix, and the probability that the trailing aircraft is of type j is equal to p_j. This means that the probability of an i-followed-by-j pair is given by

$$p_{ij} = p_i \cdot p_j \tag{10.3}$$

The matrix **P**, of aircraft-pair probabilities p_{ij}, can thus be computed in this way, as shown in Table 10.8. For example, $p_{12} = p_1\, p_2 = (0.2)(0.35) = 0.07$ is the probability of having a pair consisting of a leading aircraft of type H followed by a trailing aircraft of type L.

		Trailing Aircraft			
		1 (H)	2 (L)	3 (S1)	4 (S2)
Leading aircraft	1 (H)	0.04	0.07	0.07	0.02
	2 (L)	0.07	0.1225	0.1225	0.035
	3 (S1)	0.07	0.1225	0.1225	0.035
	4 (S2)	0.02	0.035	0.035	0.01

TABLE 10.8 Matrix **P** of Pair Probabilities p_{ij} for Example 10.4

Multiplying the corresponding elements of the matrices **T** and **P** to apply Eq. (10.2) yields an expected value $E[T_{ij}] \approx 103$ seconds. In other words, if the ATC system could somehow always achieve the minimum allowable separations between landing aircraft, the runway of this example could serve one arrival every 103 seconds, on average, or up to about 35 arrivals per hour.

In practice, it is extremely difficult to achieve the perfect precision in spacing consecutive landing aircraft on final approach implied by the matrix **T**. With human factors playing a key role in the spacing between aircraft, it is reasonable to expect some deviations from the separations suggested by the elements T_{ij} of **T**. In fact, in view of the natural tendency of both pilots and air traffic controllers to "err on the conservative side," one would expect the separations between given pairs of aircraft types to be, on average, larger than the corresponding values of T_{ij}. This is indeed the case: for example, in the United States, average spacing in IMC typically exceeds the minimum required separations by about 5 to 15 seconds. The model presented here can capture this effect, if the matrix **T** is modified carefully. A "buffer time" (BT) can be added to every element T_{ij}, with the value of the BT chosen to account for the spacing added in practice, intentionally or unintentionally, to each *i*-followed-by-*j* pair of aircraft. For instance, under a particularly simple but reasonable approximation, one could just add the same constant buffer time, b, to all the elements T_{ij}, obtaining a new matrix $\mathbf{T}'$ whose elements t_{ij} give the average (not the minimum possible) separation achieved for an *i*-followed-by-*j* pair of aircraft. In this case,

$$t_{ij} = T_{ij} + b \tag{10.4}$$

The expected value of t_{ij} gives the average time interval between consecutive landings on the runway. By analogy to Eq. (10.2), one can now write this expected value as

$$E\left[t_{ij}\right] = \sum_{i=1}^{k} \sum_{j=1}^{k} p_{ij} \cdot t_{ij}$$

(10.5)

Example 10.4 (continued) Suppose that b = 10 seconds in Eq. (10.4). This means that all intervals between consecutive landings are 10 seconds longer than the minimum, due to inaccuracies in spacing of aircraft, conservatism on the part of pilots and controllers, etc. Obviously, the expected amount of time between successive landings will also be greater than the expected minimum separation $E[T_{ij}]$ by 10 seconds. In other words, $E[t_{ij}] \approx 113$ seconds = 0.03139 hour. This leads to a capacity estimate of $\mu \approx 32$ aircraft per hour, a number typical of the service rates that might be observed at an airport in the United States with a traffic mix similar to this example's operating with IFR separations.

It is easy to use this model to assess the sensitivity of airport capacity to changes in various input parameters that may result from changes in the ATM system, airline fleet composition, terminal area procedures, etc. Consider a few instances.

First, a comparison of Table 10.7 with the o_i column of Table 10.5 indicates that the runway occupancy time of the leading aircraft is not the constraining factor for any of the 16 possible pairs of consecutive landing aircraft. All 16 values of the T_{ij} in Table 10.7 are greater than—and, in a single case, equal to—the value of the corresponding o_i. [The one case in which equality applies is T_{21}, for which the minimum separation dictated by the final approach spacing requirement of 2.5 nmi for the "type-2-followed-by-type 1" pair is equal to (2.5)(3600)/(150) = 60 seconds, the same as the 60 seconds runway occupancy time of the class 2 (or L) aircraft that leads the pair.] This means that any reductions in the runway occupancy times in Table 10.5 will not increase arrival capacity. (Such reductions in o_i could, e.g., be obtained through the construction of high-speed runway exits.) In practice, it is indeed true that, for practically all ATM systems in the world, the final approach IFR spacing requirements, such as those shown in Tables 10.1 and 10.2, are more restrictive than the runway occupancy times.[21]

Second, suppose this airport was still operating with a 3-nmi separation requirement (instead of 2.5 nmi) for the L-H, L-L, S-H, S-L, and S-S aircraft pairs, as is the case at less busy airports (cf. Table 10.1). The reader can verify that this would reduce capacity by approximately 2.5 arrivals per hour (from 32 to 29.5), or by approximately 8 percent. $E[t_{ij}]$ will be equal to approximately 122 seconds.

Similarly, air traffic controllers at the busiest airports in the United States often attempt to achieve more uniform final approach speeds, typically by recommending that pilots fly the smaller and slower aircraft at speeds more similar to those of some of the commercial jets on final approach. For instance, if, in this example, the ATM system could achieve v_3 = 130 knots and v_4 = 110 knots through higher final approach speeds of S1 and S2 aircraft, $E[t_{ij}] \approx 103$ seconds, or approximately 35 arrivals per hour, an increase of approximately 9 percent over the 32 arrivals computed with the original approach speeds.

Note that a combination of (1) reducing the 3-nmi separations to 2.5-nmi, (2) increasing the final approach speeds of S1 and S2 aircraft to 130 and 110 knots, respectively, and (3) a potential reduction in the safety buffer to b = 5 seconds, instead of b = 10 seconds, has the overall effect of reducing $E[t_{ij}]$ from 122 to 98 seconds and increasing capacity from 29.5 to approximately 37 arrivals per hour, a 25 percent increase! It is the cumulative effect of relatively small changes such as these that has prevented airport capacity from falling well behind growing demand over the past 20 years.

Finally, if the expected amount of time between consecutive landings has been computed, the maximum throughput capacity (in terms of landings per hour) is simply given by

$$\text{Maximum throughput capacity} = \mu = \frac{1}{E[t_{ij}]} \tag{10.6}$$

where $E[t_{ij}]$ is measured in hours.[22]

Other possibilities for increasing runway capacity can be assessed by exploiting our simple mathematical model. For example, inspection of the matrix **T** (Table 10.7) indicates that certain aircraft sequences are more desirable than others. For example, the sequence 1-4 (or H-S2) requires at least 4 minutes of separation between consecutive landings, whereas the sequence 4-1 (or S2-H) requires only 1 minute. This suggests the possibility of computer-aided sequencing of aircraft waiting to land at an airport, an idea that has been investigated in detail by several researchers (Dear and Sherif, 1991; Psaraftis, 1980; Venkatakrishnan et al., 1993; Balakrishnan and Chandran, 2010) and is now being partially implemented through advanced ATM decision-support systems (see Chap. 13). Note that, when sequences not based on a FCFS discipline are in use, Eq. (10.3) is no longer necessarily valid and must be replaced by an expression—or an algorithm—for computing probabilities p_{ij} that reflect the sequencing scheme actually in use.

10.6 Generalizations and Extensions of the Capacity Model

The capacity model presented in the last section can be extended in a number of ways. Its accuracy can be improved as well. Before discussing some of these extensions and improvements, it is important to summarize the basic approach that the model follows. For all practical purposes, all mathematical models of runway capacity follow essentially this same approach, consisting of three basic steps.

Step 1: For all possible pairs, *i* and *j*, of aircraft classes and for all permissible pairs of movements ("arrival followed by arrival," "arrival followed by departure," etc.) involving a type *i* aircraft followed immediately by a type *j* aircraft, compute the expected time interval t_{ij} between successive movements, such that no ATM separation requirements are violated. Note that the average time interval t_{ij} is greater than or, at best, equal to T_{ij}, the minimum interval between successive movements for that aircraft pair, because t_{ij} also accounts for deviations from optimal spacing due to human factors, operational exigencies, etc.

Step 2: Compute p_{ij}, the probability of occurrence of each of the expected time intervals, t_{ij}, obtained in step 1.

Step 3: Compute the overall expected time of the interval between any two consecutive movements,

$$E[t_{ij}] = \sum_{i=1}^{k} \sum_{j=1}^{k} p_{ij} \cdot t_{ij} \tag{10.7}$$

and from that the (maximum throughput) capacity

$$\text{Maximum throughput capacity} = \mu = \frac{1}{E[t_{ij}]} \tag{10.8}$$

Example 10.5 illustrates the application of the three-step approach, previously used in the "all arrivals" model, to the case where a runway is used only for departures.

Example 10.5 Consider a runway with the same aircraft mix as in Table 10.8, but with all aircraft now performing takeoffs from the runway. Assume that the separation requirements, s_{ij}, that apply in this case are as follows, in units of seconds:

		Trailing Aircraft		
		H	**L**	**S1 or S2**
Landing aircraft	H	90	120	120
	L	60	60	60
	S1 or S2	60	60	60

In the case of departures, it is reasonable to assume that, because of the simplicity of the control process, the average interval between the beginning of the takeoff run of two aircraft of types i and j is roughly equal to the minimum separation required between these aircraft. Thus, $t_{ij} \approx s_{ij}$ for all pairs i and j in this case. Assuming, as before, FCFS sequencing of departures on the runway, the pair probabilities, p_{ij}, are the same as in Table 10.8, since the aircraft mix is the same. Applying Eq. (10.7) (i.e., multiplying each of the probabilities in Table 10.8 by the appropriate 60-, 90-, or 120-second separation requirement) one obtains $E[t_{ij}] \approx 71$ seconds and $\mu \approx 51$ departures per hour.

An analogous three-step approach can be used to estimate the capacity of a runway used for both landings and takeoffs. As already noted in Sec. 10.3, it is important in this case to identify the strategy employed by ATC controllers to sequence landings and takeoffs on the runway. Under the strategy most commonly used, controllers during peak demand periods may serve a string of consecutive arrivals (e.g., 5–10 arrivals in a row), then a string of consecutive departures, then another string of arrivals, and so on. The runway capacity can then be approximated as a simple weighted average of μ_a, the runway capacity when the runway is used only for arrivals, and of μ_d, the runway capacity when the runway is used only for departures, the weights being equal to the fractions of time spent in serving arrivals and departures, respectively. An alternative strategy used occasionally by controllers at busy airports in the United States has arrivals alternating with departures: the separations on final approach between successive arriving aircraft are "stretched" so that a departure can take off during the time interval between the two arrivals. This is a procedure that requires considerable skill but, if performed accurately, can increase significantly the capacity of the runway, as measured by the total number of movements (landings and takeoffs) performed. A model of this operating strategy was developed by Hockaday and Kanafani (1974) and was subsequently generalized by several researchers (Swedish, 1981). In Exercise 3 at the end of this chapter, the reader is guided through the application of the three-step approach to this case.

Some improvements and extensions to the generalized three-step approach can now be reviewed briefly. To begin, it is obvious that some of the parameters that are treated as constants in the examples presented so far, such as the approach speeds, v_i, and the runway occupancy times, o_i, for each class of aircraft, can be viewed more realistically as random variables with associated probability distributions. Most important, the distances between successive aircraft on final approach are random variables whose probability distribution depends on the ATM system's separation requirements and on the characteristics and performance of the terminal area ATM system, including the controllers and pilots (Harris, 1972; Odoni, 1972).

Computer-based mathematical models developed more recently address all these possibilities (Lee et al., 1997; Andreatta et al, 1999; Stamatopoulos et al, 2004). These are generalized probabilistic models for computing capacity, when a runway is used for arrivals only or for departures only or for mixed operations. For instance, much as a controller would do, the models compute the spacing required between landing aircraft as they enter the common approach path so that, with reasonable confidence, no violations will occur later on as the aircraft fly toward the runway.

The principal output of these models is the runway capacity envelope (Gilbo, 1993), that is, a boundary that defines the envelope of the maximum throughput capacities that can be achieved at the runway under the entire range of possible arrival and departure mixes (Fig. 10.10). Any point inside the envelope is feasible and any point outside is infeasible. The runway has sufficient capacity to serve x arrivals per hour and y departures per hour, as long as the point (x, y) is within the runway capacity envelope.

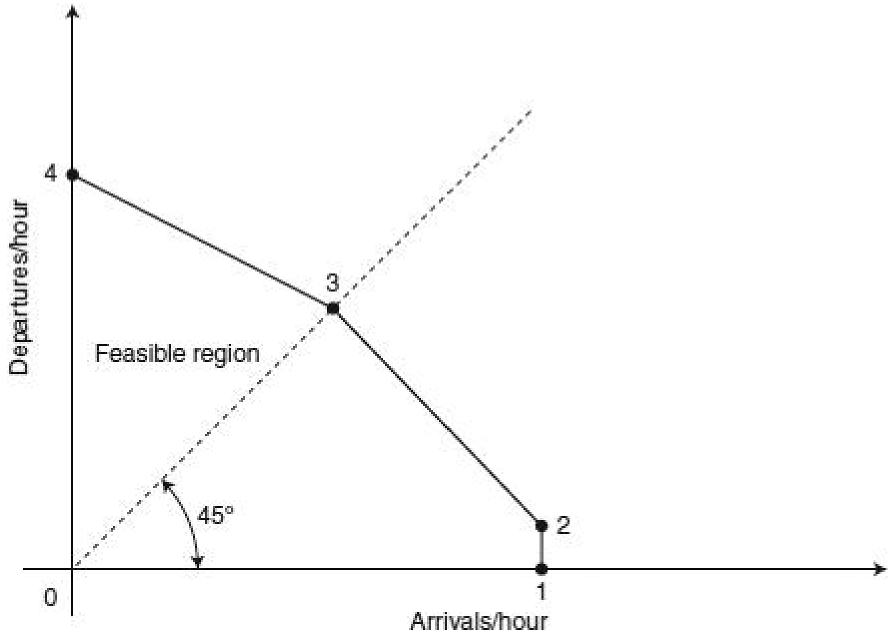


FIGURE 10.10 A typical capacity envelope for a single runway.

The models mentioned previously typically compute the coordinates of only four points on the envelope, denoted as points 1, 2, 3, and 4, and then approximate the entire envelope by interpolating between them with straight-line segments in the manner shown in Fig. 10.10. The four points are the following:

Point 1: This is the "all arrivals" point; that is, it indicates the capacity of the runway when it is used for arrivals only.

Point 2: This is known as the "free departures" (or "arrival priority") point, because it has the same capacity for arrivals as point 1 and a departures capacity equal to the number of departures that can be inserted into the arrivals stream without increasing the separations between successive arrivals and, thus, without reducing the number of arrivals from what can be achieved in the all-arrivals case. Thus, the "free departures" are obtained by exploiting large interarrival gaps.

Point 3: This is the "alternating arrivals and departures" point, that is, the point at which an equal number of departures and arrivals are performed through an A-D-A-D-A…sequence. As indicated previously, such a strategy can be implemented by "stretching," when necessary, interarrival (and inter-departure) gaps by an amount of time just sufficient to insert a departure (arrival) between two consecutive arrivals (departures).

Point 4: This is the "all departures" point, that is, the capacity of the runway when it is used only for departures.

The same modeling approach can be extended quite readily to airport configurations with two simultaneously active runways. The simplest possible case involves two parallel runways, of which one is used solely for arrivals and the other solely for departures, independently. As noted in Chap. 9, most airports outside the United States with this geometric layout typically operate in this way. In such instances, one can obtain the capacity envelope simply by computing the "all arrivals" (point 1) and "all departures" (point 4) capacities of a single runway and combining the results as shown in Fig. 10.11. Note that point 2 lies above the 45° line in Fig. 10.11, reflecting the fact that the departures capacity is higher than the arrivals capacity in most cases.

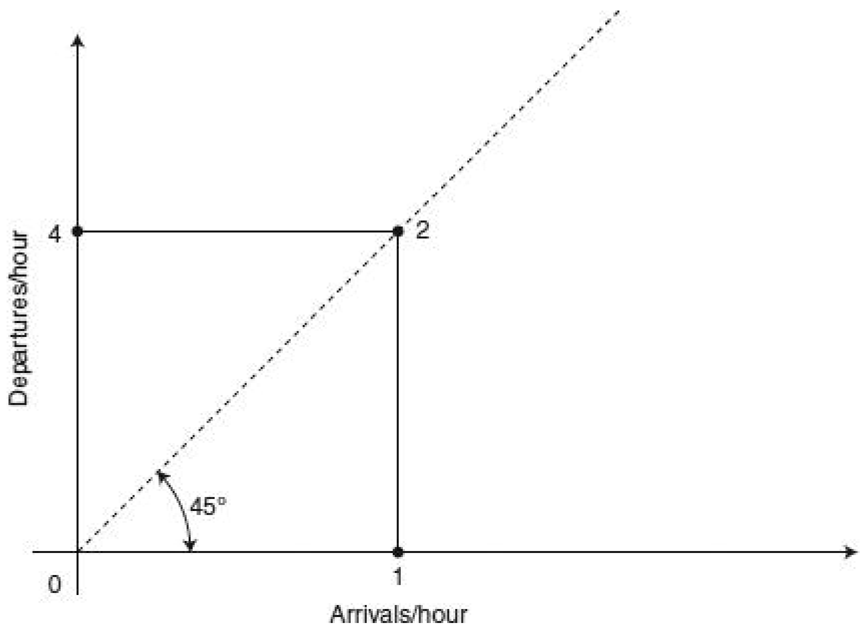


FIGURE 10.11 A capacity envelope for two independent runways: one used for arrivals only and the other for departures.

Capacity envelopes such as the ones shown in Figs. 10.10 and 10.11 provide a complete description of the capacity made available by a runway system under any specific set of conditions. Note that different capacity envelopes may (and, most probably, will) apply to

VFR or IFR or LIFR operating conditions (see Sec. 10.3). This is illustrated in Fig. 10.12 that shows capacity envelopes for a hypothetical airport where visual flight rules (VFR) are applied in good weather conditions (VMC) and IFR in "poor" weather (IMC). Note that the IMC capacity envelope is (typically) fully contained within the VMC capacity envelope. Thus, some arrival-departure combinations, which can be performed within 1 hour in VMC, are not feasible in IMC.

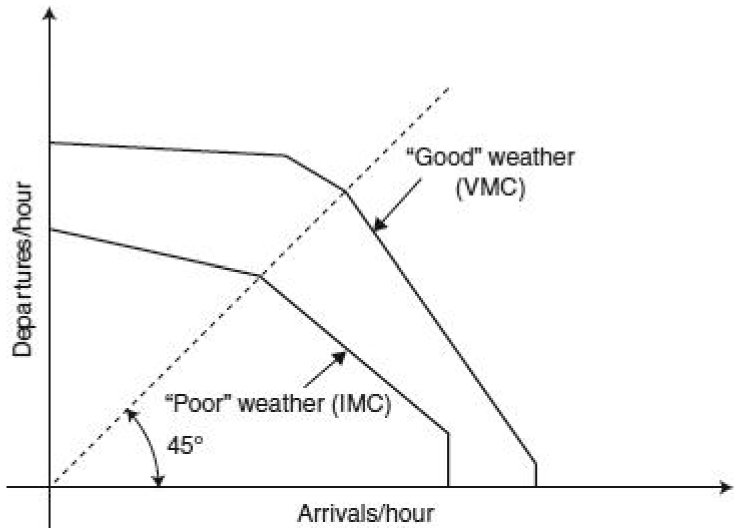


FIGURE 10.12 "Good" weather and "poor" weather capacity envelopes for the same airport.

Instead of computing a capacity envelope for 1-hour periods, one may also compute the envelope for 15-minute or 30-minute periods. This is often useful for airports that practice slot coordination (see Chap. 12), as well as in cases where one wishes to estimate more accurately the air traffic delays that will occur at an airport (see Chap. 11).

More generally, the case of two parallel runways is quite tractable using the three-step approach outlined earlier, no matter the separation between the runways and the types of movements (arrivals, departures, or mixed) served by each runway[23] (Swedish, 1981; Stamatopoulos et al, 2004). The case of intersecting pairs of runways is also tractable, as long as local procedures for operating the runways are well understood. Given a set of priority rules for sequencing operations on the two active runways, one can compute the elements, t_{ij}, of the time separation matrix for consecutive operations and approximate quite accurately the available capacity. For example, in the case of New York/LaGuardia (see Fig. 9.6) a configuration commonly used has departures on Runway 13 and arrivals on Runway 04. Air traffic controllers will typically alternate arrivals and departures in this case (a departure from Runway 13, then an arrival on Runway 04, then a departure from Runway 13, etc.). Given the location of the runway intersection, one can then compute the airport's capacity for this particular assignment of operations to runways. Note, however, that the assignment of landings and takeoffs to runways may change at New York/LaGuardia (or other airports with similar intersecting runway geometries), depending on wind direction, thus giving rise to additional configurations. Because of the consequent change of the location of the runway intersection relative to the points where takeoffs are initiated or where landing aircraft touch down, the capacities of these different configurations (all involving two active runways) may be far from equal.

With configurations involving three or more active runways, the three-step approach may become very cumbersome. The interactions among the runways may be too numerous and complicated to permit development of matrices of separations between all possible pairs of movements on all active runways, as called for under step 1. Instead, one of two alternative methodologies may be used: "decomposition" of the configuration or simulation. The former involves decomposing the configuration in use into parts, each of which consists of either a single runway or a pair of runways. This is followed by estimation of the capacity of each of the parts, using the one- and two-runway models just described. For example, Atlanta (see Fig. 9.7) can be viewed as consisting of three quasi-independent sets of runways: two independent pairs of close-spaced parallel runways and a single runway (the shorter fifth runway) that is used for mixed operations. The capacity of the full runway system can then be approximated by computing, first, the capacity of each of these three components separately, and then adding the results. Simulation models are also used extensively to estimate runway system capacity (see Sec. 11.5).

Through these two approaches (decomposition or simulation) it is also possible to estimate approximately the capacity envelopes of airports with more complex runway configurations. Figure 10.13 sketches what a capacity envelope for a multirunway airport might look like. Approximate capacity envelopes for the 34 airports listed in Table 10.4 are given in (FAA, 2004).

FIGURE 10.13 A hypothetical capacity envelope for a multirunway airport with mixed use of the runways.

10.7 Capacity of Other Elements of the Airfield

This section reviews briefly the capacity of the other elements of the airfield, namely, the taxiway system and the apron. Some general statements can be made about the capacity of these elements. However, airport-specific factors also play a primary (and often dominant) role in determining taxiway and apron capacity.

Capacity of the Taxiway System

The overall capacity of the taxiway system can be determined, in theory, by the number of aircraft per hour that the taxiway system can deliver from the apron areas to the runway system and vice versa. From a practical viewpoint, however, it suffices to know that a fully developed and reasonably well-designed taxiway system, like those that one is likely to encounter at major airports, will not, in general, be a factor limiting airport capacity. This can be seen by considering the most fundamental component of a taxiway system, a full-length taxiway, that is, a taxiway that runs parallel to the entire length of a runway and serves aircraft moving to/from the apron areas from/to the runway (see Chap. 9). These long taxiways are typically used as one-way traffic lanes for any given runway configuration. The flow capacity of a full-length taxiway typically exceeds by a considerable margin the capacity of the associated runway. For example, if aircraft travel on the taxiway at a speed of 36 km/h (~ 22 mi/h) and the separation between (noses of) successive aircraft on the taxiway is a conservative 400 m, the flow capacity of the taxiway is 90 aircraft per hour, far more than a runway can typically handle. The flow capacity will, of course, be higher if taxiing speeds are higher or if headways between successive airplanes on the taxiway are smaller.

This does not mean that a taxiway system may not have local bottlenecks—points where aircraft may sustain some taxiing delays which are additional to the delays suffered while waiting to use the runway system. Taxiway intersections, short taxiway segments between two intersections, points where taxiing aircraft must cross an active runway, and locations where high-speed runway exits merge with taxiways can all be potential local "hot points" on a taxiway system. Figure 10.14, for instance, identifies pictorially a set of points where arriving aircraft must cross the departures runway 22R, when 22L and 27 are used for arrivals and 22R and 22L for departures at Boston/Logan (Idris, 2001). During periods when 22R is busy with departures, delays at these crossing points can be significant. Equally important, air traffic controllers may occasionally have to interrupt the flow of departures on runway 22R to give waiting arriving aircraft an opportunity to cross 22R and reach the apron areas. Such interruptions may reduce the departures capacity of the runway system at Boston/Logan when the configuration of Fig. 10.14 is in use.

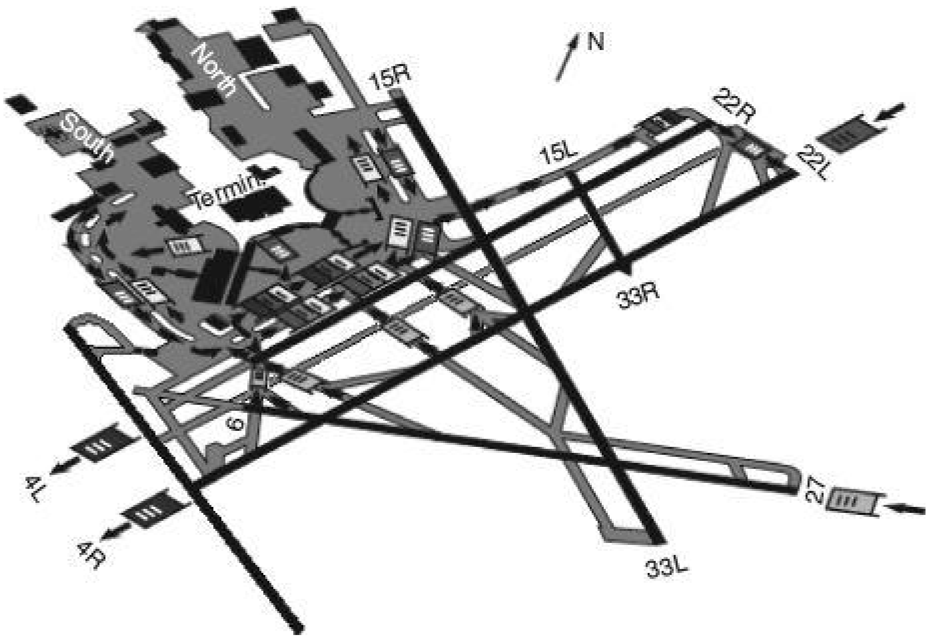


FIGURE 10.14 Potential congestion points at Boston/Logan when Runways 22L and 27 are used for arrivals and 22R and 22L for departures; locations of queues on the surface are noted.

Several such flow-constraining points typically exist on the taxiway systems of older, space-constrained airports. These points can usually be identified easily, essentially by inspection of the layout of the airfield. Taxiway flow problems, arising from the airfield geometry, can be solved only through location-specific measures. Ground controllers, that is, the air traffic controllers responsible for directing traffic on the airport's surface, are generally well aware of the presence of these potential bottlenecks on taxiway systems and try to anticipate and prevent localized delay problems from spreading throughout the airfield.

In conclusion, the overall capacity of the taxiway system of major airports almost always exceeds the capacity of the runway system and does not constitute a significant constraint on runway capacity. Delays sustained at specific "hot points" are typically much smaller than those experienced due to the capacity limitations of the runway system. Some exceptions may exist at older, space-constrained airports. The general rule is that taxiway capacity problems are airport-specific and must be resolved in the context of local conditions.

Capacity of the Aprons

In contrast to the taxiway system, the capacity of aprons can occasionally be a constraining factor on the overall airside capacity of space-constrained airports. Aprons consist of areas reserved for remote and contact aircraft stands and for taxilanes, that is, the corridors in which the aircraft utilizing these stands circulate. Stands can be further subdivided into those designated for exclusive use by a single airline (or possibly a small group of affiliated or allied airlines) and those for shared (or common) use. At many major airports in the United States, most of the stands are for exclusive use, whereas the opposite is usually true elsewhere (see Chaps. 14 and 15). When stands are for exclusive use, the scheduling of stand occupancy times and the assignment of aircraft to stands are managed by the airlines themselves or by a contractor responsible for ramp handling in that part of the apron area (see Chap. 8). When stands are shared, it is either the airport operator or a handling contractor who performs these tasks. Each stand is also characterized by its size—the dimensions of the largest aircraft it can accommodate.

Some general statements can be made about the capacity of aprons, but airport-specific conditions usually dominate. At the most obvious, a good indication of the available apron capacity is given by the number of stands at hand. This is sometimes referred to as the *static capacity* of the apron, because it indicates the maximum number of aircraft that can be occupying simultaneously the apron at any given instant. The static capacity is usually also broken down according to the maximum size of aircraft that can be accommodated ("X stands for Group V aircraft, Y for Group IV, Z for Group III, etc.").

Static capacity, while informative, provides only a "snapshot" of the capacity of the apron, that is, how many aircraft can be parked at any given instant. Static capacity cannot be readily compared to the runway capacity of the airport—or the capacity of other parts of the airport—which is specified in terms of number of aircraft movements (or of passengers, bags, etc.) that can be processed per unit of time. For this reason, the *dynamic capacity* of aprons is also a widely used measure. Dynamic capacity is defined as the number of aircraft per hour that can be accommodated at the stands and is more consistent with the notion of runway capacity.

To compute dynamic capacity, it is necessary to consider the time interval between successive occupancies of a stand by two different aircraft. By analogy to the approach used to compute runway capacity, a minimum interval and an average interval should be determined. The minimum interval consists of the sum of two components:

1. The amount of time that an aircraft is scheduled to spend at the stand; this will be referred to as the scheduled occupancy time (SOT) and is also known as the scheduled "turnaround" time of an aircraft.
2. The time needed to position the aircraft into and out of the stand; during that positioning time (PT) the stand is unavailable to other aircraft.

Typical values of SOT range from 20 minutes for flights of regional airlines, making a stop at an airport to unload some passengers and take new ones on board with no aircraft servicing, to 4 hours for wide-body aircraft turning around on an intercontinental route. Note this does not cover occupancy times as long as 10 to 12 hours for "overnighting" aircraft. Typical values for PT are of the order of only 2 to 4 minutes for remote stands,[24] to as much as 10 minutes or more for contact stands (due primarily to the time-consuming pushback maneuver).

To determine an average interval between successive stand occupancies, one must also consider the "buffer times" built into the schedules of stands at all major airports. Managers of stands, whether airport operators or airlines, are forced to make allowances for such buffer times, especially at contact stands, because air traffic is always subject to delays and short-term schedule changes. If a given stand is to be assigned to two aircraft consecutively, there should be sufficient time between the scheduled departure time of the first and the scheduled arrival time of the second to ensure that, with high probability, deviations from schedule will not necessitate a change in stand assignments. Airport operators and airlines prefer to avoid last-minute changes in stand assignments, because they are disruptive and costly: they inconvenience departing passengers in the case of contact stands and they require reassignment or repositioning of ramp equipment and of aircraft handling and passenger handling personnel. The buffer times (BT) actually used will depend greatly on local circumstances, such as the length of typical flight delays, airport policies vis-à-vis stand assignments (e.g., a preference for assigning a particular flight to the same stand every day versus variable assignments from day to day), stand type (remote or contact), apron geometry, passenger terminal configuration (for contact stands), aircraft handling agreements, and exclusive or shared use of stands. Typical BTs can range from a few minutes for stands serving remotely parked regional aircraft to 1 hour or more for contact stands for intercontinental flights. Example 10.6 illustrates how dynamic capacity is affected by some of these parameters.

Example 10.6 Consider a simplified situation in which all stands at an airport are of the same size and can accommodate all aircraft using the airport. Assume there are 60 stands and the average SOT for all aircraft is 50 minutes. The naive approach in this case is to estimate that each stand can serve an average of 1.2 aircraft per hour, so that the dynamic capacity of the apron is (1.2)(60) = 72 aircraft per hour.

If the positioning time, PT, were also taken into account—by adding, for example, 8 minutes to the 50 minutes of SOT—the apron capacity would be substantially reduced to approximately 62 aircraft per hour [= (1.03)(60)]. This would be the maximum achievable dynamic capacity, assuming that the schedule of flights was executed perfectly every day.

For purposes of this example, let us assume a buffer time, BT, of 30 minutes is now added to the 58 minutes previously allowed for SOT and PT, giving a total of 88 minutes during which access to a stand by other aircraft is "blocked." Each stand can then serve about 0.68 aircraft per hour and the dynamic capacity is reduced to approximately 41 aircraft per hour, 43 percent less than the naive estimate of 72.

The procedure outlined in Example 10.6 can now be summarized and generalized somewhat. Assume that a set of n stands exists at an airport, with each stand capable of accommodating all types of aircraft.

Step 1: Subdivide arriving aircraft into a small number K of classes according to an appropriate combination of criteria such as aircraft size and/or type of flight and/or airline. For instance, class i might consist of wide-body aircraft on long-range (5 hours or more) international flights. Note that the classes specified for the purpose of computing apron capacity are not necessarily the same as the classes (e.g., "heavy," "medium," "small") specified for the purpose of computing runway capacity.

Step 2: For each class i estimate the typical ("average") time between occupancies of the stand as the sum of SOT, PT, and BT for that class. For class i, call this sum the "stand blocking time," SBT.

Step 3: Compute the expected ("average") SBT for the airport's stands as

$$E[\mathrm{SBT}] = \sum_{i=1}^{k} p_i \mathrm{SBT}_i \tag{10.9}$$

where p_i is the fraction of arriving aircraft that belong to class i.

Step 4: The dynamic capacity of the apron is then approximately equal to $n/E[\mathrm{SBT}]$ aircraft per hour.

It should be emphasized that this procedure will yield only a rough estimate of dynamic capacity. This is especially true when, as happens at practically every airport, all stands cannot accommodate all aircraft, either because of physical limitations (e.g., size of the stand) or because of operational constraints (stands reserved for international flights vs. stands for domestic flights, etc.). For a more accurate estimate of apron capacity under these more complicated conditions, the best approach is to subdivide the stands into groups to which reasonably homogeneous conditions of use apply. One can then perform a separate capacity analysis for each such identifiable group of stands using the procedure outlined previously and then combine the results. However, this analysis can be quite tedious, even when the number of stands is relatively small. It can be facilitated through use of several available computer-based tools that can assign aircraft to stands by taking into consideration many of the constraints, operational rules, and airline priorities and preferences that are typically encountered in practice. It should be noted, however, that these tools are intended primarily for the task of fitting a specified daily schedule of arrivals and departures into the available set of stands at an airport. Thus, they can assist only indirectly in estimating the apron's dynamic capacity by indicating whether a given hypothetical daily schedule of flights can be accommodated by an airport's set of stands. To determine the dynamic capacity, one then needs to test many variations of daily schedules with increasing numbers of flights and varying BTs between stand occupancies.

A last related question concerns the comparison of the dynamic capacity of the apron with the capacity of the runway system. This question often arises in the context of determining the number of slots that many airports around the world use for "schedule coordination" purposes (see Chap. 12). Note that apron capacity is measured in terms of number of aircraft per hour and runway capacity in terms of movements per hour. Obviously, as a quick approximation, one can simply multiply the dynamic capacity of the apron by 2 to convert it to a number that can be compared to runway capacity, as the occupancy of a stand is associated with two movements on the runways, an arrival and a departure.

Example 10.6 (continued) Consider again the situation described earlier, in which E[SBT] was set equal to 88 minutes, about 1.5 hours, and the dynamic capacity of the 60 stands was consequently estimated as 41 aircraft per hour. Multiplying by 2 gives an estimate of 82 as the number of movements per hour on the runway system that can be accommodated in the apron area. Stated differently, if the runway system has a (maximum throughput) capacity of 82 movements per hour, the runway system is able to "feed" the apron about 41 arriving aircraft per hour, a number equal to the rate at which the apron can serve aircraft.

The better approach, however, calls for scanning the schedule of arrivals and departures during the busy part of a typical day (e.g., from 07:00 to 21:00 local time) to identify the most "arrival-intensive" 1.5-hour interval of the day. Suppose this interval occurs between 08:10 and 09:40 local time and that, during the interval, arrivals constitute 62 percent of the scheduled runway movements. An estimate of 66 [$\approx 41/(0.62)$] is then obtained for the apron's capacity, expressed in terms of runway movements per hour. The logic here is that, if a runway system with a maximum throughput capacity of 66 movements per hour is available and if the airport's strongest surge of arrivals over a 1.5-hour period results in 62 percent arrivals and 38 percent departures during that period, the runway system will send about 41 aircraft per hour to the apron (i.e., a number equal to the capacity of the apron) if the runways are working at full capacity. Note that the new equivalent capacity of 66 movements per hour is significantly lower than the 82 obtained through the "simple method."

A more prudent approach, however, takes into consideration the fact that the daily flight schedule at any airport contains periods during which there are considerably more arrivals than departures, and vice versa. Surges in arrivals may "flood" the apron with aircraft. The approach consists of two steps. First, the schedule of runway movements during the busy hours of the day is scanned to identify the largest fraction of arrivals in the traffic mix during any time interval of length comparable to $E[\text{SBT}]$, as defined in step 3 of the procedure described previously. The apron's dynamic capacity is then divided by this fraction to obtain the equivalent capacity expressed in terms of runway movements per hour. The advantage of this approach (illustrated below in the continuation of Example 10.6) is that it does not overestimate the apron's ability to cope with the unavoidable fluctuations in arrivals and departures during the day. It is particularly useful for hub airports that experience several major surges in arrivals and departures in the course of a day. Note, however, that when applying this approach to project the need for apron stands at a future time, one requires both good historical data on the dynamic mix of arrivals and departures over the course of a day and a reasonable guess as to what this dynamic mix will look like in the future.

The following practical rule of thumb can also be stated: To convert the dynamic capacity of the apron to an equivalent number of runway movements per hour, multiply the dynamic capacity by 1.67. Note that the coefficient 1.67 [$\approx 1/(0.60)$] implies an approximately "60 percent arrivals, 40 percent departures" mix during the peak arrivals surge of the day. This is reasonable and works quite well for busy, nonhub airports.

Exercises

10.1. The following information is given about air traffic at a particular runway of an airport:

a. Aircraft are classified into three types: heavy (H), large/medium (L), and small (S).

b. Some relevant aircraft characteristics are as follows:

Aircraft Type	Approach Speed (knots)	Mix (%)	Runway Occupancy Time on Landing (s)
H	150	20	70
L	135	40	60
S	105	40	50

c. The length of the final approach to the runway is 6 nmi.
d. The minimum separation requirements (in nautical miles) between successive landing aircraft on final approach are given by the matrix below (rows indicate the leading aircraft and columns the trailing aircraft):

	S	L	H
S	2.5	2.5	2.5
L	4*	2.5	2.5
H	6*	5	4

*These separations apply only when the leading aircraft is at the runway threshold; all other separations apply throughout the final approach.

e. A "buffer time" (BT) of 15 seconds (see Sec. 10.5) is added to all the minimum separation times between successive landings to account for uncertainties.
f. The minimum separation requirements (in seconds) between successive departing aircraft are given by the matrix below (rows indicate the leading aircraft and columns the trailing aircraft):

	S	L	H
S	45	45	45
L	60	60	60
H	120	120	90

Part 1: Suppose this runway is used for departures only. Find its (maximum throughput) capacity for departures. (No BTs are added for departures.)
Part 2: Suppose this runway is used for arrivals only. Find its (maximum throughput) capacity for arrivals.

10.2. Consider the model for the capacity of a single runway with arrivals only presented in Sec. 10.5. Let W be a constant greater than 1. Assume that the runway occupancy times are negligible compared to the time intervals between arrivals dictated by the longitudinal separation requirements on final approach. Assume that you have been assigned the task of assessing two alternative proposals to improve the capacity of a runway: (a) multiply the final approach speeds of all aircraft types by W; and (b) divide the length of the final approach path by the same constant W. With the exception of the proposed changes, everything else in the model remains the same.

Would the two proposals lead to the same improved runway capacity? If not, which of the two proposals would lead to the higher runway capacity? Justify your answer using the analytical model of the runway capacity.

10.3. The airport of Exercise 10.1 is sometimes forced to use only a single runway during IFR weather periods. The runway must accommodate both landings and takeoffs during these periods. The data in Exercise 10.1 for arriving and departing aircraft apply here, unless noted otherwise. The following rules/assumptions apply:

a. The local air traffic controllers use an operations sequencing strategy of alternating landings and takeoffs on the runway; that is, during periods of continuous demand, a landing is always followed by a takeoff, which is then followed by a landing, etc. Thus, when the minimum required time gap between two landing aircraft, i and j, is not sufficient to insert a takeoff, the time gap will be increased by ATC appropriately.

b. There is no uncertainty about the position of aircraft on final approach. Thus, this is an entirely "deterministic" problem and no buffers are added to minimum aircraft spacing; ignore assumption e of Exercise 10.1.

c. Takeoffs wait next to the threshold of the runway. As soon as a landing aircraft crosses the runway threshold, the next departing aircraft enters the runway and prepares for the takeoff run. It takes 40 seconds for a departing aircraft to enter the runway and set up for takeoff. (Note that, in the meanwhile, the arriving aircraft that just landed is moving down the runway toward a runway exit.)

d. A takeoff run cannot begin until the preceding landing aircraft has cleared the runway.

e. Once a takeoff run begins, the runway occupancy time for all departing aircraft (time from the beginning of the takeoff run to clearing the runway) is 60 seconds.

f. The takeoffs of successive aircraft must be separated by at least 90 seconds in this case (i.e., disregard the separation matrix shown in f of Exercise 10.1).

g. A landing aircraft is not allowed to cross the runway threshold unless the runway is clear of all landing or departing aircraft. (Note that this is assumed to be the only "departure followed by arrival" separation requirement.)

Part 1: Find the capacity of this runway (total number of landings and takeoffs per hour) when it is used for both arrivals and departures in the manner described.

Part 2: Part e stated that all aircraft, independent of type, occupy the runway for 60 seconds on departure. In practice, the three different types of aircraft [H, L, or S] have different (typical) runway occupancy times on takeoff. Without doing any calculations, write a couple of paragraphs explaining how you would modify your work in Part 1 to take this into consideration. Make sure to indicate what your "time separation" matrix and your matrix of probabilities would look like.

References

Andreatta, G., Brunetta, L., Odoni, A., Righi, L., Stamatopoulos, M., and Zografos, K. (1999) "A Set of Approximate and Compatible Models for Airport Strategic Planning on Airside and on Landside," *Air Traffic Control Quarterly, 7*, pp. 291–317.

Balakrishnan, H. and Chandran. B. (2010) "Algorithms for Scheduling Runway Operations under Constrained Position Shifting," *Operations Research, 58*, pp. 1650–1665.

Blumstein, A. (1959) "The Landing Capacity of a Runway," *Operations Research, 7*, pp. 752–763.

Dear, R. and Sherif, Y. (1991) "An Algorithm for Computer Assisted Sequencing and Scheduling of Terminal Area Operations," *Transportation Research A, 25*, pp. 129–139.

FAA, Federal Aviation Administration (1981) *Airfield and Airspace Capacity/Delay Analysis*, Report FAA-APO-81-14, Office of Aviation Policy and Plans, Washington, DC.

FAA, Federal Aviation Administration (2004) *Airport Capacity Benchmark Report*, Washington, DC.

FAA, Federal Aviation Administration (2011) "Air Traffic Control: Order JO 7110.65M," Air Traffic Organization Policy, Washington, DC.

FAA, Federal Aviation Administration (2012) *Airport Design*, Advisory Circular 150/5300-12, U.S. Government Printing Office, Washington, DC.

Gilbo, E. (1993) "Airport Capacity: Representation, Estimation, Optimization," *IEEE Transactions on Control Systems Technology, 1*, pp. 144–154.

Harris, R. M. (1972) *Models for Runway Capacity Analysis*, Report FAA-EM-73-5, The MITRE Corporation, McLean, VA.

Hockaday, S. and Kanafani, A. (1974) "Developments in Airport Capacity Analysis," *Transportation Research, 8*, pp. 171–179.

Idris, H. (2001) "Observation and Analysis of Departure Operations at Boston Logan International Airport," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

ICAO, International Civil Aviation Organization (2009) *Aerodromes, Annex 14 to the Convention on International Civil Aviation, Volume I: Aerodrome Design and Operations*, 5th ed., ICAO, Montreal, Canada.

Lee, D., Kostiuk, P., Hemm, R., Wingrove, W. and Shapiro, G. (1997) *Estimating the Effects of the Terminal Area Productivity Program*, Report NS301R3, Logistics Management Institute, McLean, VA.

Odoni, A. (1972) "Efficient Operation of Runways," in *Analysis of Public Systems*, Drake, A., Keeney, R. and Morse, P. (eds.), MIT Press, Cambridge, MA.

Psaraftis, H. (1980) "A Dynamic Programming Approach for Sequencing Groups of Identical Jobs," *Operations Research, 28*, pp. 1347–1359.

Stamatopoulos, M., Odoni, A., and Zografos, K. (2004) "A Decision Support System for Airport Strategic Planning," *Transportation Research C, 12*, pp. 91–118.

Swedish, W. (1981) *Upgraded FAA Airfield Capacity Model*, Report MTR-81W16 and FAA-EM-81-1, The MITRE Corporation, McLean, VA.

Venkatakrishnan, C., Barnett, A., and Odoni, A. (1993) "Landings at Logan Airport: Describing and Increasing Airport Capacity," *Transportation Science, 27*, pp 211–227.

[1]The idea of associating capacity with some acceptable LOS has great merit. As Chaps. 14 through 16 show, this is a key concept when it comes to defining the capacity of airport passenger buildings. For example, while one could, in principle, jam four people in an area of 1 m^2 (~ 2.5 ft^2 per person), no one would dare claim that the capacity of a 2000-m^2 lobby in a terminal is 8000 passengers, as the crowding would be intolerable under such conditions. The capacity, in this case, would undoubtedly have to be determined with reference to an acceptable level of personal comfort (i.e., to a LOS).

[2]When first used, the performance targets were called "engineered performance standards."

[3]Sections 10.3 and 10.4 provide details about high- and low-capacity runway configurations.

[4]This is also true, in most cases, for declared capacity. However, there exist a few instances of airports that have declared a capacity roughly equal to their maximum throughput capacity. If the number of scheduled movements at these airports is then set to be approximately equal to the declared capacity, very long delays will ensue on a routine basis.

[5]The meaning of "well-located" and "well-designed" in this context is discussed in several parts of this book, including a number of relevant points later in this section.

[6]The 2-nmi requirement is usually more restrictive than the requirement that the departing aircraft be clear of the runway before the arriving aircraft touches down. In visual meteorological conditions (VMC) only the latter requirement applies. A more precise statement of the former requirement is that "a departure should be separated from a trailing arrival on final approach by 2 nmi, if separation will increase to 3 nmi within 1 minute after takeoff." Controllers refer to this as "2 increasing to 3."

[7]A somewhat less strict requirement may apply at a number of airports; it states that the trailing aircraft cannot touch down on the runway unless the leading aircraft is more than a specified distance (e.g., 8000 ft) ahead on the runway and heading toward a runway exit.

[8]An additional consideration is that these separations can be applied only when the departure courses of the aircraft involved diverge by 15° within 1 mile of the runway end (7110.65 para. 5-8-3). Controllers refer to this as "fanning" the departures.

[9]Additionally, certain conditions must be satisfied regarding availability of adequate airspace to ensure safe separation in case of missed approaches (FAA, 1989).

[10]High capacities are also attained in the United States when the intersection is at the far end of two long runways, if both runways are used for arrivals; this is achieved through use of the "land and hold short" of the intersection (LAHSO) procedure.

[11]Even in VMC, however, operations on parallel runways with centerline spacing under 2500 ft (762 m) are treated as in Table 10.3 whenever wake turbulence may be a factor.

[12]Departures are also typically sequenced according to FCFS. However, advanced ATM systems occasionally deviate from the FCFS order so as to avoid particularly wasteful sequences of aircraft, such as a stream of consecutive H-S arriving pairs. In fact, some of the new ATM automation systems currently being installed at busy terminal areas in the United States and in Europe include software that assists controllers in performing a limited amount of aircraft resequencing to increase capacity and efficiency, as described in Chap. 13.

[13]Such surges are often referred to as “an arrival push” or “a departure push.”

[14]Europeans use the term “wave” and Americans the term “bank” to refer to surges of connecting arrivals and departures by any particular airline at a hub airport (e.g., “American Airlines schedules eight banks a day at Dallas/Ft. Worth” or “KLM schedules five waves a day at Amsterdam”).

[15]This is approximately the case in practice at times when the airport is not heavily utilized.

[16]As Sec. 10.5 shows, the longitudinal separation on final approach is, with few exceptions, the more restrictive of the two requirements. This means that the principal benefits from high-speed exits usually come from the A-D case, not the A-A.

[17]The “benchmark” capacities in Table 10.4 should be interpreted as only partial indications of the overall capacities at these airports.

[18]Several more airports outside the United States (e.g., Frankfurt/International, Beijing, Delhi, and the new Dubai Airport) are scheduled to join this group within a few years because of the addition of new runways.

[19]The capacity is so low that the configuration does not even appear in the Boston/Logan CCC of Fig. 10.7.

[20]This describes the situation during the late 1990s.

[21]It is possible that if runway occupancy times were reduced considerably from their current values, the longitudinal separations required on final approach might also be reduced.

[22]The Greek letter μ has been used to denote capacity. This is in line with standard notation in queuing theory (see Chaps. 11 and 20).

[23]As seen in Sec. 10.3, the capacity of a pair of parallel runways depends on the distance between runway centerlines, cf. Table 10.3.

[24]Remote stands with taxilanes behind and in front of them have the shortest positioning times.

CHAPTER 11

Airfield Delay

Airport delays and congestion constitute a major threat to the future of air transportation. The dynamic characteristics of airport delays are difficult to predict accurately. Advanced computer-based tools are typically needed to obtain good estimates of delay-related measures at a busy airport over time. In general, delays

- May be present even during periods when the demand rate is lower than capacity.
- Depend nonlinearly on changes in demand and/or capacity, becoming very sensitive to even small changes when demand is close to or greater than capacity.
- Exhibit a complex dynamic behavior over any time span (e.g., a day of operations) when the runway system is utilized heavily.

In the long run, both the expected length and the variance (a measure of variability) of airport delays increase nonlinearly with increases in the utilization ratio of the runway system, that is, the ratio of the demand rate divided by the (maximum throughput) capacity. Delays will be very long and highly variable from day to day at runway systems operated with utilization ratios in excess of the 0.85 to 0.9 range during the most active 15 to 18 traffic hours of the day.

It is difficult, in practice, to attribute air traffic delays to specific causes. In addition to airport congestion, such factors as poor weather conditions, mechanical problems with aircraft, slow processing in terminal buildings or at airport stands, etc. may contribute to delays. Moreover, delays propagate across the air transportation system, so that the delays observed at an airport may be due to congestion at a different airport. It is very important to have a clear understanding of these complexities when measuring and attributing delays.

A quantity of great practical interest is the approximate annual capacity of a runway system. To estimate this number, it is necessary not only to compute the capacity coverage chart (CCC) of an airport but also to make projections or assumptions regarding daily demand patterns, day-of-the-week demand patterns, seasonal demand patterns, and acceptable levels of delay. Historical evidence provides some useful guidelines regarding these parameters.

A number of simulation models and mathematical, computer-based models are available to assist in investigating issues related to airfield capacity and delay. A crucial decision on the part of the prospective user concerns the level of modeling detail needed for the analysis. The more detailed (and more costly and complex) models are not necessarily better suited for many of the questions that come up in practice.

11.1 Introduction

The principal consequence of the lack of adequate airside capacity at an airport is delays to landings and takeoffs, with their attendant economic and other costs. When delays become large, other undesirable consequences such as missed flight connections, flight cancellations, and flight diversions may also become commonplace. Airport and air traffic congestion is a growing problem on an international scale and is widely viewed as one of the principal constraints to the future growth of the global air transportation industry. In the United States alone, the total cost of air transportation delays was estimated by the most detailed study to date (Ball et al., 2010) to have been $29 billion in 2007, the worst year ever for aviation delays. This was the sum of the direct cost to the airlines ($8.3 billion) and passengers ($16.7 billion), and of the welfare loss incurred by passengers who avoid air travel as the result of delays ($3.9 billion). To provide some perspective on these numbers, the U.S. airline industry's highest annual profits in the decade 2001–2010 amounted to about $8 billion (in 2006).[1] About 25 percent of scheduled arrivals at the 34 busiest airports in the United States, where more than two-thirds of all U.S. passengers were enplaned, were more than 15 minutes late in 2007.[2] The corresponding figure at the 34 busiest airports in Europe was just as bad at 23 percent. Several airports in Asia, which were reasonably delay-free until recently, such as Singapore and Hong Kong, now routinely report significant flight delays.

The major U.S. airlines have been required by law since 1987 to report statistics on flight delays to the U.S. Department of Transportation (DOT). This requirement reflects the great interest that the subject of air traffic congestion has for travelers and the general public. The U.S. government processes and publishes these statistics on a monthly basis (DOT, monthly) and the media review them carefully. These statistics also provide a rich data source for researchers and for developers of computer-based models of airport delays. The European Union has also initiated recently a program that requires airlines to provide detailed data on delays at European airports—a practice that was voluntary previously.

This chapter briefly reviews some fundamental points concerning the estimation, characteristics, and measurement of airside delays at airports. Section 11.2 qualitatively reviews the characteristics of airside delay and congestion in a short-term, dynamic sense. Section 11.3 discusses long-term characteristics and some of their policy implications. It provides several important practical guidelines that airport operators should follow for the purpose of maintaining an adequate level of service on airside. Section 11.4 addresses the question of the annual capacity of a runway system. Its estimation requires consideration of local demand patterns and of level-of-service issues. Finally, Sec. 11.5 covers briefly some issues related to the computation of delays and to the availability of delay data. Chapter 20 provides a more quantitative discussion of certain aspects of airport congestion, as well as a short introduction to queuing theory—the mathematical theory of waiting lines. Chapters 12 and 13 describe approaches for reducing the magnitude and/or cost of airside delays through demand management and better air traffic management (ATM).

11.2 The Characteristics of Airside Delays

It is useful to begin with a qualitative look at the relationship between airside demand and capacity, on the one hand, and delays, on the other. Figure 11.1 shows schematically a typical weekday demand profile at a major U.S. airport and compares it with three different levels of (maximum throughput) capacity, associated with "good," "reduced," and "unfavorable" weather conditions. As Chap. 10 notes, weather is the principal factor that determines the level of capacity at which the airport operates at any given time.[3] The demand profile, which has two peaks, as is typical of many busy airports with large volumes of business traffic, shows the total number of aircraft movements *scheduled* per hour. Note that the number of movements scheduled is not necessarily the same as the number that airlines will *actually* operate on a day-to-day basis during that hour. Because of mechanical or logistical problems with aircraft, flight cancellations, late-boarding passengers, late-arriving crews, delays at other airports, etc., the number of movements actually requested at a given airport during any particular period of time will fluctuate around the number scheduled. In this sense, just like capacity, the number of movements scheduled for an hour can be viewed as only an *expected* (or "average") *value*. This expected value is henceforth referred to as the *demand rate* for that hour.

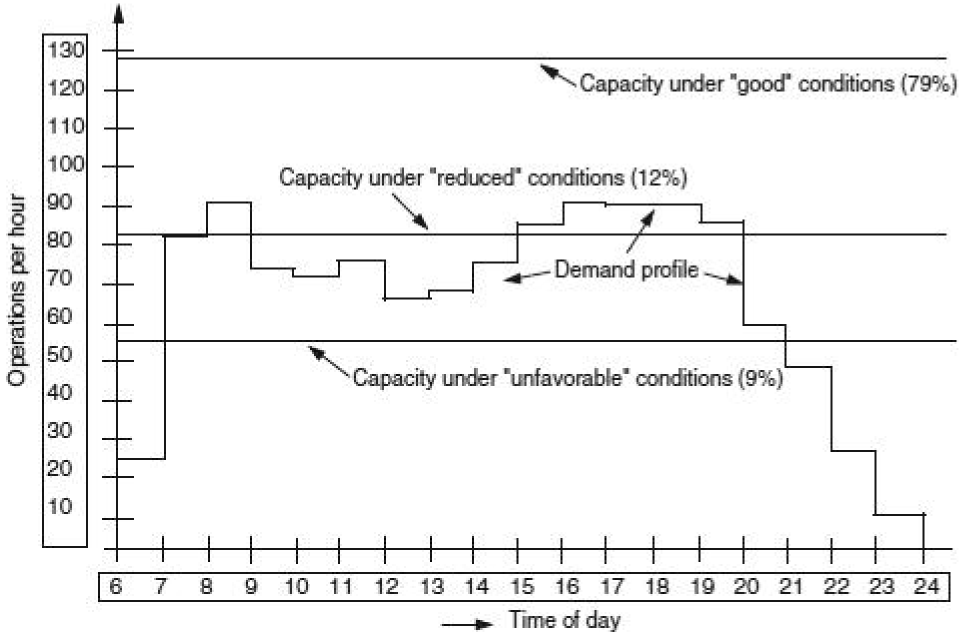


FIGURE 11.1 Weekday demand profile at a major U.S. airport compared with three different levels of the capacity of the airport's runway system.

A few observations can now be made about the delays associated with each of the three levels of capacity. First, queues of landing and departing aircraft will almost certainly form and delays (often called *overload delays*) will occur during those parts of a day when the demand rate exceeds the capacity for any significant length of time. This is because aircraft will seek to use the runway system at a rate greater than the system's capacity. Colloquially, "demand exceeds capacity" during such periods. In general, if there is an interval of length T during which the demand rate continually exceeds the service rate, both the expected length of the aircraft queue and the expected waiting time per aircraft during that interval will grow in direct proportion[4] to T and to the difference between the demand rate and the capacity during T. In Fig. 11.1, the period during which demand exceeds the "unfavorable" level of capacity runs from 07:00 to 21:00 and the "reduced" level from 08:00 to 09:00 and from 15:00 to 20:00. As one can infer, on days when the weather is "unfavorable," congestion will keep building up throughout the day and aircraft scheduled to arrive or depart during the afternoon and evening hours may be subjected to horrendous delays. Indeed, the airlines may be forced to cancel many flights during such days because of the size of the expected delays.

Less obviously, significant delays may also be observed when the demand rate is *less than but reasonably close* to the service rate—a notion that is sometimes confusing to airport operators. For example, one would typically observe significant delays between 09:00 and 15:00 on days when the airport is operating in "reduced" conditions. Some delays may even occur on days when the demand rate is less than the capacity for the *entire day*, as is the case for "good" conditions in Fig. 11.1. Such delays are due primarily to the variability of the time intervals between consecutive requests for use of the runways, as well as to the variability of the time it takes to process ("serve") each landing and takeoff. The sources of this variability are several:

- The time instants at which demands (arrivals and, especially, departures) are *scheduled* to take place are generally not evenly spaced but are often "bunched together" around certain times, which aircraft schedulers prefer (e.g., "on-the-hour" or "on-the-half-hour" departure peaks).
- The instants at which demands *actually* occur on a day-to-day basis are "randomized" as a result of the inevitable deviations from schedule due to the many reasons already mentioned (mechanical problems, delays at other airports, etc.).
- The amount of time it takes to serve departures and arrivals on the runway system is not constant, but varies with the many factors Chap. 10 discusses (type of aircraft, separation requirements from preceding aircraft, runway exit used, etc.).

The net effect is the presence of time intervals during which "clusters" of several closely spaced demands and/or of longer-than-usual service times occur. Queues of airplanes will then form on the ground and/or in the air.[5] When the demand rate is smaller than the capacity but close to it, a long time may pass before such queues dissipate. In fact, new clusters of demands or of long service times may come along before the previously formed queue has dissipated and the waiting line(s) may get longer for a while, not shorter. The resulting delays are often called *stochastic* (or "probabilistic") to distinguish them from overload delays. In summary, long queues may form even if the demand rate is smaller than capacity in cases where (1) there are "spikes" in the scheduled demand or there is considerable variability in the times between consecutive demands on the runway systems and/or in the service times at the runway system and (2) the demand rate is close to the runway system's capacity (see Example 11.1). What "close" means in this context is discussed shortly.

Example 11.1 Figure 11.2, based on a study of delays at a major airport, illustrates all these points. The curve Dem shows the demand profile[6] during a day when the total number of scheduled runway movements is 1200 and the peak-hour demand rate is 94 movements during the hour between 15:00 and 16:00. (The scale for the demand profile is shown next to the vertical axis on the right-hand side of Fig. 11.2.) Practically all the demand, with the exception of 32 nighttime movements, is concentrated in the 16 "busy" hours of the day between 06:00 and 22:00 local time.

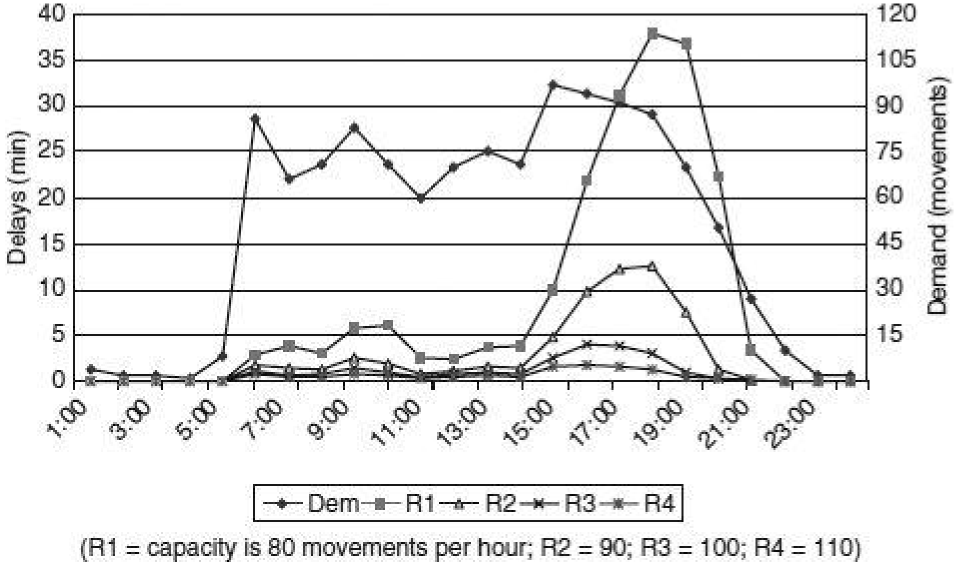


FIGURE 11.2 The dynamic behavior of aircraft delay for four different levels of capacity ranging from 80 to 110 movements per hour; the scale for the demand profile is on the right and for the expected delay on the left.

Figure 11.2 also shows the estimated expected waiting time in queue ("average delay") when the maximum throughput capacity at the airport is 110, 100, 90, or 80 movements per hour throughout the day. Specifically, for each time *t* on the horizontal axis, each of the four graphs shows how much delay a landing or takeoff that requests use of the runway system at time *t* would be expected to suffer when the capacity is at the level (110, 100, 90, or 80) corresponding to that graph. The delay estimates have been generated through DELAYS, a computer-based queuing model (see Sec. 11.5).

Figure 11.2 makes several noteworthy points. First, as already mentioned, delays occur not only under overload conditions, when the demand rate exceeds capacity (as happens for short or for longer parts of the day when the capacity is equal to 90 and 80 movements per hour, respectively), but also when the demand rate is less than capacity throughout the day (as happens when the capacity is equal to 100 or 110). Note that the overall shapes (i.e., the peaks and valleys) of the four graphs for the expected waiting time during the day are not fundamentally different in the former (some overloads) and in the latter (no overloads) cases. The major difference lies in the fact that the overall magnitude of the expected waiting time and its peaks increase as one transitions from a capacity of 110 to a capacity of 100, 90, and then 80.

Second, observe that, as capacity decreases, the magnitude of the delays grows dramatically, in a *nonlinear* way. Table 11.1 highlights this aspect of the behavior of this queuing system (ignore for now the two rightmost columns). The second column from the left shows the maximum value, to the nearest minute, that the expected waiting time takes during the 24-hour day. With a capacity of 80 movements per hour, an aircraft requesting access to the runway system shortly after 18:00 can expect a delay of roughly 39 minutes (see also Fig. 11.2). The third column indicates the estimated overall expected waiting time per movement over the entire day. Note that, when the hourly capacity is reduced from 110 to 100 movements per hour (a 9.1 percent reduction), the overall expected delay per movement doubles (from 0.8 to 1.6 minutes), whereas an 11.1 percent reduction, from 90 to 80, leads to a *tripling* of overall expected delay per movement, from 4.3 to 12.8 minutes. In this sense, Fig. 11.2 and Table 11.1 illustrate one of the principal results of queuing theory: Expected delay changes nonlinearly as the demand rate and/or the capacity change—and the closer the demand rate is to capacity, the more sensitive delay is to even small changes in demand and/or capacity. From Table 11.1 we can estimate that, on days when the capacity per hour is 80, aircraft using this airport incur a total of about 15,000 minutes [≈ (12.8)(1200)] or 250 hours of delay.

At a typical direct operating cost of about $3600 per aircraft hour, this is equivalent to a daily cost of $900,000! It should be emphasized that the waiting times shown in Fig. 11.2 and in the second and third columns of Table 11.1 are *expected* ("average") *values*. At each capacity level and at each time of the day, the delay that will actually be observed on any given day is a random variable. Thus, some aircraft may suffer, on a day-to-day basis, considerably longer or shorter delays than indicated by the expected values in Fig. 11.2. In fact, another fundamental result of queuing theory states that the *variability of delay*—as measured by the *variance* of the delay or its *standard deviation*—also increases nonlinearly as the demand rate gets closer to capacity. One might therefore expect that, on certain days, especially when the capacity is low (e.g., 80 movements per hour), the actual delays experienced will be considerably longer than shown[7] in Fig. 11.2.

Capacity (Movements per hour)	Expected Waiting Time (minutes)		Utilization Ratio	
	Maximum	Per Movement	24 h	06:00–21:59
110	2	0.8	0.455	0.664
100	4	1.6	0.500	0.731
90	13	4.3	0.556	0.812
80	39	12.8	0.625	0.913

TABLE 11.1 Some Queuing Statistics for Example 11.1

Figure 11.2 also illustrates some of the complex dynamic characteristics of queues and delays: the delay (and queue length) aircraft experience during any particular time interval depends strongly on the waiting times and queue lengths during previous intervals. For example, it can be seen in Fig. 11.2 that, for each of the four levels of capacity, the expected delay per aircraft during the morning hour of 06:00–07:00 is very small compared to that during the three afternoon hours that begin at 15:00. This is despite the fact that the demand rate between 06:00 and 07:00 is only about 10 percent smaller than the average demand rate during the 3 hours that begin at 15:00 (86 per hour vs. about 94 per hour). Part of the explanation is that the morning peak hour of 06:00–07:00 is preceded by a period of practically no demand (and no delays), whereas a queue has already started building up well before the beginning of the afternoon peak period. Moreover, the morning peak lasts for only about 1 hour, whereas the one in the afternoon persists for several hours in a row.

Another interesting aspect of the dynamic behavior of airside queues is that a lag often exists between the time when the demand rate peaks and the time when delays reach their peak. This time lag may be long on days when the demand rate exceeds the capacity continuously for a significant period of time.[8] For example, in the case when the capacity is 80 movements per hour in Fig. 11.2, one can see that the peak of the expected delay occurs between 18:00 and 19:00, while the peak demand hour is between 15:00 and 16:00. The reason for this time lag is that queues "build up" during periods of high traffic demand. Thus, those flights that request access to the runways near the end of these periods must necessarily join the end of the queues that have already formed and, as a result, may experience the worst delays. This phenomenon often occurs at busy airports. It is also one that motorists often experience when driving *after* the peak of the morning or afternoon "rush hour" on congested highways.

11.3 Policy Implications and Practical Guidelines

Of the many characteristics and properties of airside delays discussed in connection with Example 11.1, the one that has the most important implications for the long-term growth of congestion at busy airports is the nonlinear relationship between delays, on the one hand, and demand and capacity, on the other. The key parameter in this regard is the *utilization ratio*, typically denoted in queuing theory by the Greek letter ρ ("rho"), which is defined as the average demand rate over a specified period of time divided by the average capacity over that time. For instance, suppose that, in Example 11.1, the specified "period of time" is a 24-hour day. Because the demand is 1200 movements per day, one obtains $\rho = 0.625$ (= 1200/1920) for the 24-hour day in the case when the hourly capacity is 80, meaning that the daily capacity is 1920 (= 24 × 80). This value of ρ is shown in the last row of column 4 of Table 11.1.

An untenable situation would result if anyone attempted to operate an airport with a daily utilization ratio, ρ, greater than 1. This would mean that the number of movements requested per day would be greater than the daily capacity of the airport. Thus, on average, some movements would not be served by the end of each day and left in a queue to be "processed" on the next day.[9] The leftover demand would then be added to the already scheduled demand for the next day—which, by assumption, was greater than the daily capacity, in the first place. The queue of movements accumulated at the end of the day would thus grow *ad infinitum* from day to day in the long run.

Clearly, the situation just described is so extreme that it makes little practical sense. However, it points to a general condition that must be satisfied if a queuing system is to function in a stable way over a long period of time. This condition states that a queuing system *cannot* be operated *in the long run* with a utilization ratio greater than 1, because delays for its use will never reach equilibrium and will grow without limit.

Having established that, in the long run, ρ must be less than 1, the following fundamental property of queuing systems can now be stated informally as follows: *In the long run, both the expected waiting time and the expected queue length at any queuing system that reaches equilibrium* ($\rho < 1$) *increase nonlinearly with ρ, in proportion to the quantity* $1/(1-\rho)$.

This very important relationship is shown schematically in Fig. 11.3 for a typical queuing system and can be interpreted as follows. Suppose a queuing system has been operating for a sufficiently long time that its long-term equilibrium characteristics can be observed and measured statistically. Let W_q denote the time that a random user spends waiting in queue before being served by the system. The quantity plotted in Fig. 11.3 is the *expected value,* $E[W_q]$, of W_q. For example, in the context of a runway used for arrivals only, $E[W_q]$ would be the expected (or "average") time an aircraft requesting to land would spend waiting for its turn to approach the runway.[10] Note that as ρ approaches the value of 1 in Fig. 11.3—or, as the demand rate approaches the service rate or, more colloquially, as "demand approaches capacity"—$E[W_q]$ grows rapidly in proportion to the quantity $1/(1-\rho)$. $E[W_q]$ finally becomes infinite for ρ greater than or equal to 1, as suggested by the necessary condition for equilibrium described previously. Let us also define the quantity N_q, to be the number of users waiting in queue and let $E[N_q]$ denote its expected value. A plot of

$E[N_q]$, the expected number of users in queue, versus ρ would have a similar shape to that of $E[W_q]$ in Fig. 11.3. The exact mathematical expressions for $E[W_q]$ and $E[N_q]$ depend on several parameters. In general, the higher the variability of user service times (i.e., of the processing time per movement in the case of runways) and of the intervals between consecutive demands by users (i.e., between successive aircraft requesting use of the runways), the faster $E[W_q]$ and $E[N_q]$ will increase as ρ increases. However, the dominant characteristic of both $E[W_q]$ and $E[N_q]$ remains the same: they increase nonlinearly in proportion to $1/(1 - \rho)$.

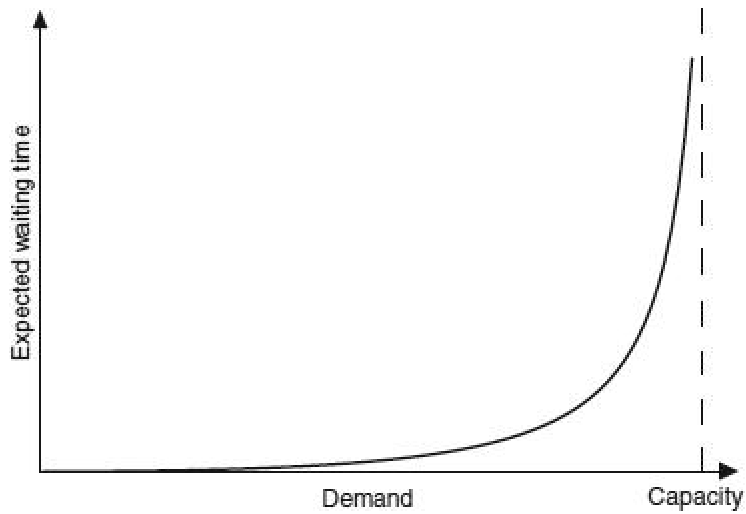


FIGURE 11.3 Typical relationship between the expected waiting time and the demand at a queuing system.

From the practical point of view, this observation and the generic "shape" of the function $E[W_q]$ shown in Fig. 11.3 (as well as of $E[N_q]$) have important implications for airports at the policy level.

First, they provide a warning to airport managers and ATM policy makers not to operate runway systems at levels of utilization that are very close to 1 over an extended time. If they do, long aircraft delays and queues will occur routinely. Moreover, queuing theory has shown that the variability of W_q and N_q, as measured by their standard deviations, $\sigma(W_q)$ and $\sigma(N_q)$, also increases in proportion to $1/(1 - \rho)$. This means that, when ρ is close to 1, a queuing system not only experiences serious congestion on average but is also subject to large fluctuations over time. Under the same set of *a priori* conditions (similar demand rates, weather conditions, etc.), delays to landings and takeoffs may be modest and tolerable on a particular day and extremely long and unacceptable on the following day.

This phenomenon occurs often at the busiest airports throughout the world. A good rule of thumb is that runway systems should not be operated at more than 85 to 90 percent of their capacity for the duration of the consecutive busy traffic hours of the day. The number of consecutive busy traffic hours in a day is considerably less than 24 at the great majority of airports: typically even the busiest of them have at most 16 to 18 hours per day of high traffic activity, whereas little happens during the remaining 6 to 8 hours, usually the nighttime.

Table 11.1 illustrates these ideas clearly. In column 5, note that, with a capacity of 80 movements per hour, the utilization ratio during the 16 busiest hours of the day is equal to 91.3 percent. The associated delays are essentially unacceptable, with the average waiting time for all movements during the day and during the peak hour equal to about 13 and 39 minutes, respectively. Column 4 shows that the utilization ratio for the entire day is quite modest (0.625) in this case, but this is deceptive and simply reflects the fact that there is little demand between 22:00 and 6:00, or for one-third of the day.

A second major point for policy-making purposes is that, when a runway system operates at high levels of utilization, small changes in demand or in capacity can cause large changes in delays and queue lengths. This is a direct consequence of the fact that both the expected value and the standard deviation of W_q and N_q are proportional to $1/(1 - \rho)$: when ρ is close to 1, $1/(1 - \rho)$ is large and its value is highly sensitive to even small changes in ρ. Figure 11.4 illustrates this point. Note that the same amount of change in demand (an increase in the value of ρ by 0.05) has very different consequences in terms of increases in expected delay, depending on whether the initial demand was high (0.85) or low (0.6) relative to capacity. This observation motivates much that is being done today at major airports to contain delays.

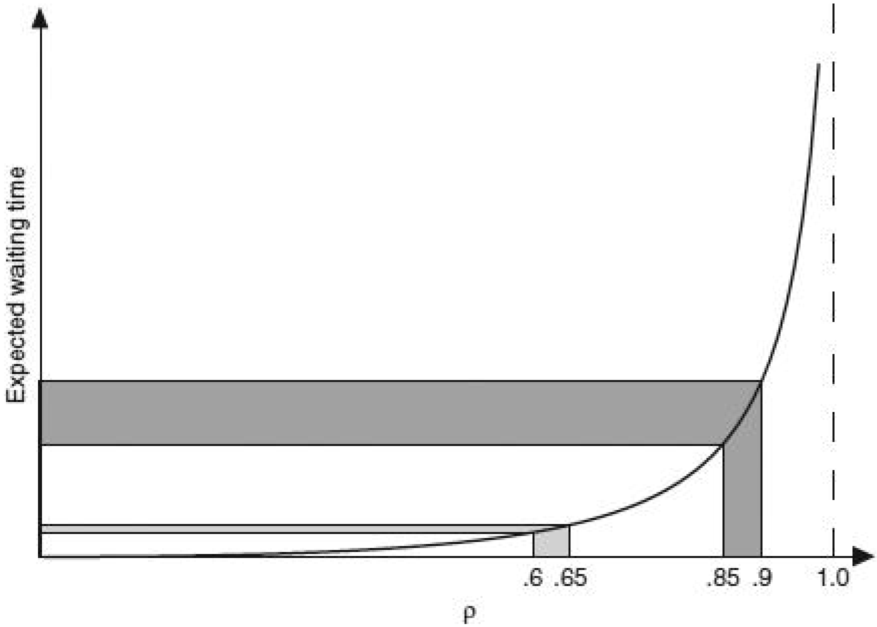


FIGURE 11.4 Nonlinear response of delay to demand changes; when an airport operates close to its capacity, delays are very sensitive to changes in demand or in capacity.

Many initiatives are currently under way aimed at either managing demand (see Chap. 12) or at increasing the airside capacity of busy airports through improvements of the ATM system (see Chap. 13). Airport operators generally recognize that most of these initiatives will produce only small changes in demand or in capacity. However, these small changes are still expected to produce significant reductions in delays, because many of the facilities and services at these airports operate at very high utilization ratios. It is hoped that such reductions in delay will make it possible to maintain acceptable levels of service until more dramatic improvements in capacity can be achieved. Experience as well as queuing models and simulations show that, at very congested airports, a 1 percent reduction in daily demand for airside operations or a 1 percent increase in runway system capacity may result in a 5 percent or more reduction in delay (Fan and Odoni, 2002).

One can also revisit some of the definitions of capacity in Chap. 10 in light of this discussion. Remember that *practical hourly capacity* (PHCAP) is defined as the number of movements at which the average delay for use of a runway is equal to 4 minutes. The 4-minute criterion was derived from graphs like those of Figs. 11.3 and 11.4, which the FAA prepared in the early 1960s. These suggested that expected waiting time at a typical airport would start to increase rapidly at levels of airport utilization corresponding to about

4 minutes of delay per aircraft (Fig. 11.5). It was therefore decided to use 4 minutes as the threshold value at which a runway would be said to have reached its "practical capacity."

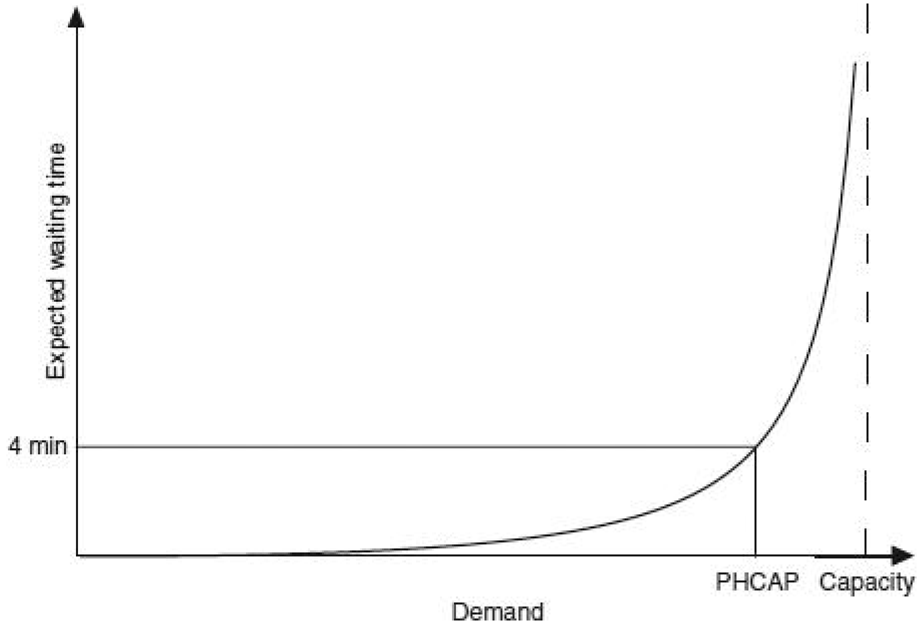


FIGURE 11.5 Determination of the practical hourly capacity (PHCAP) for a runway.

Similarly, the notion of *declared capacity*—which is used widely outside the United States, as Chap. 10 notes—is tied directly to the observation that delays will reach unacceptable levels at airports operated at their maximum throughput capacity for long periods. By "declaring" capacities that are equal to about 85 to 90 percent of the maximum throughput capacity,[11] airport operators seek to maintain an adequate level of service as well as utilize their runway systems intensively.

Figure 11.6 illustrates similar ideas, but from a somewhat different perspective. It shows the estimated average delay per operation as a function of the number of annual operations at a major airport—Orlando/International in this case. These estimates are based on a model known as the *annual service volume* (ASV), which the FAA often uses to compute approximately the delays associated with any annual volume of operations, taking into consideration all possible runway configurations and their frequency of use (see Chap. 10), as well as aircraft fleet mix, and other factors (Chin et al., 2012). Note again how average delay increases rapidly as annual demand increases beyond roughly 600,000 movements per year.

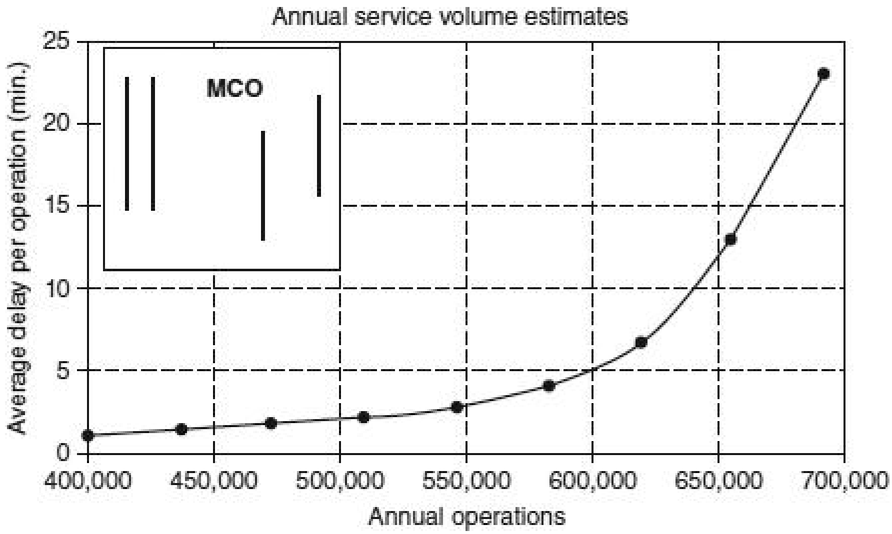


FIGURE 11.6 Average delay as a function of the number of annual operations at an airport. (*Source:* Chin et al., 2012.)

Finally, the fact that airside delay is highly variable from day to day at airports that are intensely utilized has important implications for performance measurement and assessing level of service. Specifically, airport managers should adopt metrics that describe not only the average magnitude of delays but also their dispersion around these average values. All of the following are examples of the types of performance metrics that should be of interest in an analysis of airside delays:

- Expected delay per movement during a typical day
- Variance (or standard deviation) of the delay per movement during a typical day
- Probability that delay will exceed some specified high value (e.g., 15, 30, or 45 minutes)
- Expected delay during the peak hour of the average day of the busiest month of the year.

Expected delay per movement is undoubtedly the principal measure of interest. However, airlines are almost equally concerned about the predictability of delays, that is, about how "tightly" delays are distributed around that average. If delay at an airport has a large variance (or standard deviation), the implied high variability in operating conditions means low reliability in executing daily airline schedules. Delays that exceed certain large values are also particularly damaging, as they result in missed passenger connections and in flight delays that propagate throughout an airline's network. Thus, the probability that such extreme delays will occur is another important measure of performance. Last, some specially "targeted" measures, such as the expected delay during the peak demand hours of the year, can be very helpful in assessing performance at times that are particularly critical to an airport's operations.

11.4 The Annual Capacity of a Runway System

The reviews of CCCs in Chap. 10 and of the characteristics of airside delays in the last two sections provide the necessary background for a discussion of the *annual capacity* of the runway system of an airport. The annual capacity is a quantity of great practical interest. Airport demand forecasts are typically given in terms of annual figures ("X million passengers and Y thousand aircraft movements are forecast for 2025"). By comparing such demand forecasts with estimates of annual capacity, airport operators can determine roughly the timing of major capital investments and plan accordingly.

Annual capacity can be estimated only in very approximate terms, because of two principal difficulties. First, to compute annual capacity, planners and managers must make a number of rather subjective choices concerning the minimum acceptable level of service. There are no international standards at this time to assist in making these choices. Second, annual capacity depends in large part on daily and seasonal demand patterns that are difficult to predict far into the future. These points are best explained through a detailed example.

Example 11.2 Consider again the case of Boston/Logan airport, whose CCC was described in Chap. 10. From the CCC, one can compute the average long-run hourly capacity of the airport's runway system. This capacity for Boston/Logan (see Example 10.3) is equal to approximately 115 movements per hour [(132)(0.6) + (120)(0.18) + … + (0)(0.015) ≈ 115]. One could thus infer that the airport can accommodate up to roughly 1,000,000 movements over an entire year [(115) × (24 h/day) × (365 days/year) = 1,007,400].

In truth, however, this number is entirely theoretical, far from being attainable in practice. This is because the estimate of 1,000,000 movements per year implies that the airport will be utilized at 100 percent of its average capacity for 100 percent of the time. It is necessary to adjust this estimate downward following the type of reasoning indicated as follows.

Adjustment 1. Very limited or no runway activity takes place for between 6 and 8 hours in a day (or for one-quarter to one-third of the time) at almost every major airport in the world. Passengers, especially on domestic flights, generally prefer not to arrive at or depart from airports between midnight and 6 a.m., and the marketplace usually accommodates this preference. In addition, to alleviate environmental impacts, most major airports discourage all but a small number of flights in late nighttime and post-midnight hours. In a few cases, they impose curfews that ban all nighttime flights outright. At most airports in the United States, for example, passenger flights essentially cease from roughly 23:30 to about 06:00 local time, every day. Moreover, traffic is usually low between 22:00 and 23:30. Because the issue here is the *ultimate* annual runway capacity of the airport, it may then be reasonable to assume that the airport will eventually "stretch" intensely utilized hours to the equivalent of roughly 16 to 17 hours per day. It is doubtful that a runway system could be utilized heavily for much longer than that every day. This assumption will reduce the initial estimate of annual capacity at Boston/Logan to the range of 670,000 to 710,000 movements [(115) × (16) × (365) = 671,600; (115) × (17) × (365) = 713,500].

Adjustment 2. Even this range, however, is unrealistic. It posits demand of about 115 movements per hour for 16 or 17 consecutive hours every day, 365 days a year. This would certainly lead to intolerable delays on the many days when the runway system operates for several hours in a row at capacities below 115 per hour. Airside operations would also experience serious problems on those days when average hourly capacity is somewhat higher than 115. The utilization ratio, ρ, will be very close to 1 for extended periods of the day, and as Secs. 11.2 and 11.3 indicated, this would result in long expected delays and high delay variability. It is therefore necessary to assume that demand will not exceed a certain percentage of the average hourly capacity of 115 during the 16 to 17 "useful" hours of the day, if delays are to be kept at an acceptable level. As Sec. 11.3 discussed, this means a roughly 85 to 90 percent utilization of the average maximum throughput capacity of 115. The approximate estimate of the annual runway capacity of Boston/Logan would then be reduced further (and the range of the estimate broadened) to 570,000 to 640,000 movements per year [(671,600) × (0.85) = 570,860 at the low end; (713,500) × (0.90) = 642,150 at the high end].

Adjustment 3. However, seasonal variations in demand have not yet been considered! At most airports in the Northern Hemisphere, for instance, demand during the summer season exceeds demand during the winter season, often by a considerable margin. If during the summer, Boston/Logan is utilized at 85 to 90 percent of its capacity during the 16 to 17 useful hours of the day, the utilization during the winter season will necessarily be less because some airlines will reduce their schedules, especially to international destinations and to summer resorts. (The terms "summer season" and "winter season" are used here to denote 6-month periods from May to October and from November to April, respectively.) At Boston/Logan, the number of movements on an average day during the summer season is approximately 15 to 20 percent higher than during the winter season. (This is a rather small difference in comparison to many other locations, as the number of summer season movements per day often exceeds winter season movements by 25 percent or more, especially at smaller, seasonal airports.) Thus, the estimates of the range of annual capacity must be updated once more to account for seasonal peaking. In the case of Boston/Logan, it may be reasonable to assume that the intensity of summer peaking will decrease further as traffic grows, although by not much below its already low levels. Using a 15 percent seasonal peaking percentage, the range of the annual capacity estimate then becomes 530,000 to 600,000 movements per year [(570,860/2) + (570,860/2) × (1/1.15) = 533,630; (642,150/2) + (642,150/2) × (1/1.15) = 600,270]. The lower end of this broad range is the one that can be considered the more reasonable, as the high end is based on the rather extreme assumptions of 17 heavy traffic hours during a typical summer season day with a 90 percent utilization ratio over the entire day.

It can be noted parenthetically that the number of movements at Boston/Logan in 1998–2000 was about 500,000; that is, the airport was operating within about 10 percent of its estimated ultimate annual capacity! However, in 2010 the airport served only 353,000 movements, having experienced a remarkable 30 percent drop in number of flights in a decade.[12] Note that the estimated range of annual capacity in this example did not consider the possibility of future capacity increases due to ATM developments or to expansion/improvements of the system of runways.

Several comments can now be made regarding the annual capacity of a runway system and the procedure for its estimation illustrated in Example 11.2.

1. Just like "practical hourly capacity," "sustained capacity" and "declared capacity," *annual capacity* is a derivative measure. One must first compute the maximum throughput capacity of an airport—and, indeed, its entire CCC that indicates how much capacity is available for what percent of time—to be able to estimate the annual capacity of an airport.
2. The estimation of annual capacity requires a number of implicit or explicit assumptions concerning, at the very least: future daily demand patterns, such as the

assumption of 7 or 8 essentially idle hours in the Boston/Logan example; acceptable levels of delay, for example, limiting operations to 85 to 90 percent of full capacity during the 16 to 17 useful hours of the day; future seasonal demand patterns, for example, 15 percent more operations per day in the summer season, on average; and potential future changes in the CCC of the airport. As the appropriate assumptions are likely to vary from airport to airport, so will the relationship between the maximum throughput capacity per hour of the airport and its annual capacity. In other words, the estimation of annual capacity depends very much on subjective judgment and on local factors and considerations.

3. The naive approach of multiplying the average hourly maximum throughput capacity of an airport by the number of hours in the year [(24)(365) = 8760] to compute the annual capacity of an airport greatly overestimates the true annual capacity. After all the necessary adjustments are made for level of service and for daily and seasonal demand patterns, the true annual capacity will be much smaller than the initial estimate (e.g., 530,000–600,000 vs. 1,000,000 in the Boston/Logan example).

4. Industry practices, as well as public perceptions of what is "acceptable" or "reasonable," may change over time. Such changes in perceptions would also mean changes in estimated annual capacities. In the Boston/Logan example, if nighttime air travel increased or if winter became a major travel season for leisure passengers—perhaps stimulated by airline price incentives—the annual capacity of the airport would increase! Indeed, experience suggests that future annual airside capacities have often been seriously underestimated in the United States. For example, in 2007 the number of movements at each of the three major commercial airports of New York exceeded the airport's purported practical annual capacity (PANCAP), as estimated by the FAA in 1980, by more than 50 percent (or by more than 130,000 annual movements in each case)! The reason was not an increase in the hourly capacities of these airports, as no new runways have been added since 1980, while capacities per runway have increased very modestly in that time span. What happened instead was that airport operations were spread into longer parts of the day, airlines increased the number of mid-day operations as they increased their flight frequencies on many routes, and both airlines and passengers have been forced to tolerate delays that would have been considered unacceptable in the early years of aviation. Most U.S. airports have, in fact, operated at levels significantly above their FAA-estimated PANCAPs at various times since 1980.

These observations can now be generalized. Let *A* denote the number of annual movements obtained by multiplying the number of hours in a year (8760) by the average value of the maximum throughput capacity available per hour at an airport. The annual capacity of the airport will then be equal to *kA*, with the coefficient *k* usually in the range 0.50 to 0.60. The appropriate value of *k*, in each instance, depends on local demand characteristics and willingness to accept delays. Values of *k* at the low end of the range (0.5–0.55) will apply to airports with relatively sharp daily and seasonal peaking, little or no activity for 8 to 10 hours per day, and limited tolerance for long delays. Values of *k* at the high end (0.55–0.60) will apply to airports with moderate daily and seasonal peaking, intensive utilization during all but 6 to 8 (typically nighttime) hours of the day, and high tolerance for delays.

Many airports outside the United States use *declared capacity*, instead of maximum throughput capacity, as their measure of hourly capacity (see Chaps. 10 and 12). If *A* has been computed by multiplying the declared capacity—instead of the maximum throughput capacity—by the number of hours in the year (8760), the annual capacity should be estimated by using a coefficient *k* in the range 0.60 to 0.70 (instead of 0.50–0.60). The reason is that declared capacity is typically set to about 85 to 95 percent of maximum throughput capacity (see Sec. 11.3) to ensure that delays will be reasonable, if the number of movements scheduled per hour is set equal to the declared capacity. When using declared capacity as the starting point, we have already implicitly performed adjustment 2 of Example 11.2.

11.5 Estimating Delays with Models

Planners and managers of airlines, airport operators, and ATM organizations have a major stake in activities concerned with measuring air traffic delays, attributing these delays to various causes, and estimating (typically through use of computer-based models) how delays would change under a variety of current and future conditions. For the airlines, air traffic delays impose heavy costs and disrupt daily operations. Estimates of the size of delays in upcoming months also affect greatly the preparation of flight schedules. Delay is also one of the principal measures of performance of airports and of ATM systems. Proposals for investments into new airport facilities and improved ATM systems are typically largely based on the argument that such investments would reduce delays and their cost. This and the next section address briefly procedures, issues, and difficulties associated with estimating, measuring, and attributing air traffic delays.

The estimation of the delays to be expected at a heavily utilized runway system is usually a difficult task because of the complex dynamics of queuing systems and the nonlinear relationships that drive the behavior of queues, as Sec. 11.2 discusses. However, airport planners and managers must deal with this task repeatedly. A complete analysis of airside capacity and delay requires a multistep procedure that can be summarized as follows.

Step 1. Identify all possible runway configurations and the weather conditions in which they are used.

Step 2. Compute the maximum throughput hourly capacity of each of these configurations.

Step 3. From historical records of weather conditions—and after taking into consideration local policies[13] regarding selection among alternative runway configurations when more than one configuration is available—estimate the annual utilization of each runway configuration, that is, the approximate percent of time in a typical year during which each configuration is in use. At airports where the policy is to choose at all times the available configuration with the highest capacity, this step is equivalent to determining the CCC.

Step 4. Prepare typical daily profiles of demand on the runway system (hourly number of arrivals and departures, mix of aircraft types, seasonal variations in the profiles).

Step 5. Estimate the delays associated with all applicable combinations of demand profiles and runway configurations in use.

Step 6. Estimate overall delay statistics on the basis of the results of step 5 and of the frequency with which each runway configuration is used as determined in step 3.

This procedure is obviously not simple. It requires access to extensive traffic data to prepare the demand profiles and to historical weather data to estimate the utilization of the various runway configurations. Computer-based (mathematical or simulation) capacity models are usually necessary to carry out step 2 for all but the simplest runway configurations. Reasonable approximate estimates of delays for step 5 can be obtained in some simple cases through mathematical formulae available from queuing theory. (Chapter 20 reviews this topic and provides several good references.) However, these formulae assume conditions, such as constant demand and capacity for extended parts of the day and a utilization ratio of less than 1 throughout the day ($\rho < 1$), which are often not met at busy airports. Therefore, such formulae should be used with great caution. In particular, one should avoid the use of simplistic graphs provided in handbooks (e.g., FAA, 1981) for general use. Planners are strongly advised to use either computer-based mathematical queuing models or simulation models to estimate delays in all cases that involve time-varying ("dynamic") demand and capacity at a runway system.

The type of delay model that will be most appropriate in each case depends on the requirements of the analysis. If a very high level of detail is desired, such as computing delays at every part of the airfield including taxiways, apron areas, stands, etc., one of the standard large-scale, "microscopic" simulation models should be used. The two best-known and most widely used models of this type are SIMMOD and TAAM (Total Airspace and Airport Modeler). The FAA supported initial development of SIMMOD and also distributed it. Various companies now market several versions of SIMMOD, each with somewhat different features and advantages. TAAM was originally developed in Australia and is now a product of Jeppesen, a Boeing Company.

In most instances, however, questions concerning delays center on delays associated with the runway system, typically by far the most important bottleneck at an airport. In such cases, computer-based mathematical queuing models and less-detailed, "macroscopic" simulation models are usually an adequate and less expensive alternative to detailed simulations. They are often more informative as they can explore quickly a wide range of possible scenarios. Queuing or simulation models that accept demand and capacity inputs, which are both time varying and probabilistic ("stochastic") are particularly useful. Such models capture the effects of uncertainty, that is, consider explicitly the fact that the times between consecutive demands and the duration of service times on runway systems are random variables. Figure 11.7 provides an example of the output of such a stochastic and dynamic queuing model, the DELAYS model developed at MIT. Note that, in addition to the average waiting time, this model computes the probability that an aircraft will suffer a delay greater than 15 minutes (or any other user-selected value) in every time period of a day. Odoni et al. (1997) and Barnhart et al. (2003) provide detailed reviews of macroscopic and microscopic airport capacity and delay models; Hansen et al. (2009) and Stolletz (2008) concentrate on macroscopic queuing models of runway systems.

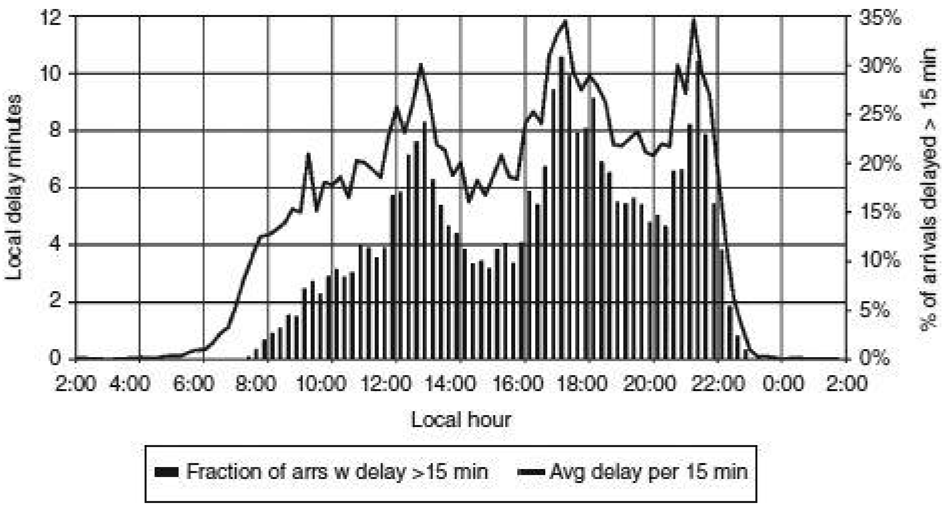


FIGURE 11.7 Estimated average delay per arrival due to local congestion (left scale) and percent of arrivals delayed by more than 15 minutes.

More recently, queuing and simulation models for the study of air traffic congestion at the *network* level have also become available. They can be used to estimate not only delays at individual airports, but also how delays propagate from airport to airport and in airspace. At the macroscopic level, AND (Airport Network Delays) and LMINET2 are examples of fast queuing network models. NASPAC (National Airspace System Performance Analysis Capability) is a simulation used extensively by the FAA. NASA has developed two highly detailed (microscopic) network models of the U.S. national airspace system: ACES (Airspace Concept Evaluation System) and FACET (Future ATM Concepts Evaluation Tool). All these network models have been applied to modeling the national airspace system of the United States. Networks consisting of the busiest airports in Europe have also been implemented in the AND model and in ATM-NEMMO, a macroscopic simulation developed in France. Pyrgiotis (2012) reviews in detail macroscopic and microscopic airport network delay models.

11.6 Measurement and Attribution of Delays

Comprehensive databases that keep track of air traffic delays on a flight-by-flight basis are essential to the task of measuring, attributing, and monitoring the evolution of air traffic delays. Such databases are increasingly being developed in many parts of the world. The U.S. DOT and the FAA maintain three such major databases: the Airline Service Quality Performance (ASQP), the Aviation System Performance Metrics (ASPM), and the Operations Network (OPSNET). Major differences exist among these three regarding the data they report, their principal areas of focus, and how they define delays and attribute them to causes. ASPM is the most comprehensive of the three, as it includes elements of the other two. A thorough discussion of the content, strengths and weaknesses of the three U.S. databases can be found in GAO (2010). EUROCONTROL also maintains a Central Office for Delay Analysis (CODA) database that assembles information on air traffic delays in all 39 of its Member States. Before 2011, the CODA database included data on only about 65 percent of all flights in the Member States because airline participation in data reporting was voluntary and most low-cost carriers chose not to participate. Many countries also maintain national databases, whose quality and comprehensiveness vary widely but are generally improving. However, all these databases should be used with caution and with full understanding of their limitations. For instance, the FAA's OPSNET database, whose statistics are often cited in the media, seriously undercounts the delays that actually occur.

The subtleties and difficulties associated with the critical tasks of measuring and attributing delay can best be explained with reference to Fig. 11.8, which schematically compares the scheduled and the actual components of a hypothetical flight of airline XYZ between two airports A and B. The scheduled gate-to-gate time (or "block time") for the flight, shown in the upper part of the figure, consists of the sum of the unimpeded taxi-out time at A, nominal flight time between takeoff from A and landing at B, unimpeded taxi-in time at B, and a "buffer" time. The latter is an additional time that airline XYZ intentionally includes in the block time, so that potential delays in performing one or more of the taxi-out, airborne, and taxi-in stages of the flight can be "absorbed" without affecting subsequent flights by the same aircraft. The actual gate-to-gate time (bottom part of Fig. 11.8) consists of the sum of the actual time it takes to complete each of the three flight stages, plus the delay, if any, incurred in leaving the flight's departure gate.

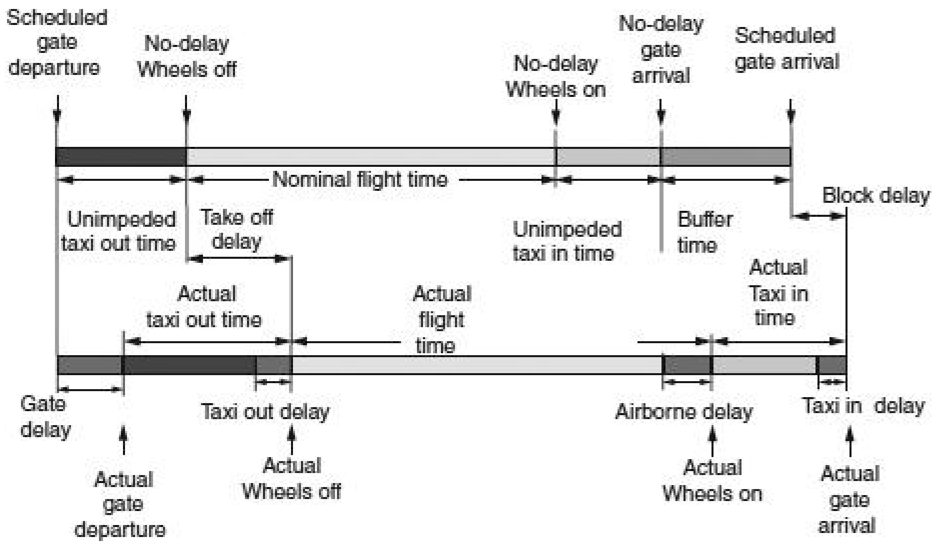


FIGURE 11.8 Scheduled gate-to-gate time versus actual gate-to-gate time for a flight.

Several important issues can now be discussed. The most fundamental concerns the meaning of the term "flight delay," which can be defined in two possible ways: *delay relative to schedule* and *delay relative to a nominal gate-to-gate time*. The former is simply the difference between the actual gate arrival time of the flight and the scheduled gate arrival time, indicated as "block delay" in Fig. 11.8. It can be negative, especially when the buffer time added by XYZ is large or in the presence of strong favorable winds or absence of delays during the actual flight. Delay relative to schedule has two important advantages. First, it is easy to measure in practice: all one needs is the actual time of gate arrival and the airline's timetable that provides the scheduled arrival time. Second, it is the measure of delay that airline passengers are primarily interested in when making their flight plans ("how late am I going to be in comparison to what I had planned for?"). For this second reason, the U.S. DOT, beginning in 1993, made it mandatory for all major carriers to publish monthly the lateness (relative to schedule) statistics for every single flight. This information is readily available to consumers and motivates, in part, the addition of buffer times (also known as "schedule padding") to nominal flight times (upper part of Fig. 11.8) (see Skaltsas, 2011). Note that, by adding a buffer time, the airline reduces the magnitude of delays relative to schedule, as well as their likelihood. By allowing for additional time to perform a flight, the airline also increases the probability that it will be able to execute its daily schedule of operations without major disruptions and flight cancellations. On the negative side, an excessive amount of buffer time will increase operating costs by reducing aircraft utilization and requiring more crews to perform any given set of flights. Thus, the setting of buffer times is an important task that forces airline managers and planners to make difficult tradeoffs between costs and schedule reliability.

However, "delay relative to schedule" has the undesirable property of (potentially greatly) understating delays: a flight that may have been 5 minutes late relative to its scheduled arrival time may, in truth, have incurred 45 minutes of actual delay, if its block time included a buffer of 40 minutes. Delay relative to a nominal gate-to-gate time is far more informative when it comes to measuring the true extent and costs of airport and air traffic congestion and planning for the capacity of aviation infrastructure (airports and the ATM system). This delay is equal to the difference between the amounts of time (1) it actually takes to complete a flight from A to B, and (2) it would have taken in the absence of any delays. Unfortunately, the computation of (1) and (2) from field data is not straightforward.

In the case of (1), the problem lies with determining what portion of the gate delay should be included in the time it takes to complete a flight. Delays in leaving the departure gate are most often associated either with events unrelated to traffic congestion (e.g., mechanical problems with the aircraft, the late boarding of passengers, etc.) or with insufficient "turnaround" time due to the late arrival at the gate of the aircraft that will perform the flight. The latter gives rise to "propagated delay" (from the previous flight)—or what is called "reactionary delay" in many parts of the world. However, delay leaving the gate may also be caused by the increasingly common air traffic flow management strategy of "absorbing" some air traffic delays at the gate, before a flight's departure (see Chap. 13 for a detailed discussion). If, for instance, it is expected that a flight from A to B will have to wait in the takeoff queue at A for 25 minutes, that flight may be held at the gate for 20 minutes, so it will spend only 5 minutes in the takeoff queue. This strategy has the dual benefit of reducing the number of aircraft on the taxiway system (which means a smaller workload for air traffic control) and reducing fuel consumption and pollutant emissions. Note that, in this example, the 20 minutes of gate departure delay should be added to the actual time it takes to complete the flight—the 20 minutes would otherwise have been part of a taxi-out delay in Fig. 11.7. Unfortunately, the existing databases do not attribute gate delays to specific causes.

In the case of (2), it is far from easy to determine unimpeded or nominal ("delay-free") travel times for each of the three stages of a flight. For example, taxi times depend on the active runway configuration at the time of arrival or departure, the airline operating the flight (determines the terminal where the flight terminates or originates), and the type of aircraft involved (determines which gates can accommodate the aircraft and, occasionally, the routing of the aircraft through the taxiway system). Similarly, total gate-to-gate time may vary greatly as shown in Fig. 11.9, depending on winds aloft, orientation of departure and arrival runways, altitude flown, etc., as well as air traffic congestion. The most common approach to the estimation of unimpeded/nominal travel times is to first obtain a distribution of such travel times from historical data, as in Fig. 11.9, and then select a value for the unimpeded/nominal travel time such that only a small fraction (e.g., 15 percent) of observed instances have travel times smaller than this value. In Fig. 11.9, for instance, one may choose the value of 340 minutes as the nominal gate-to-gate travel time from New York/Newark to Los Angeles/International on the premise that the great majority of observed travel times exceed that value and the relatively small percent of flights for which the actual travel time was smaller than 340 minutes were performed under particularly favorable flight conditions. This approach is obviously quite arbitrary and necessitated by the fact that existing databases (see as follows) do not include sufficiently detailed information to permit a more accurate reconstruction of the circumstances associated with every flight. Note that, when it comes to predicting through a model the delays that might occur under any particular future scenario (see Sec. 11.5), no actual flight-by-flight data will be available: one must compare model-estimated travel times (that include delays) with unimpeded/nominal travel times, which are based on historical data and the statistical procedure just described.

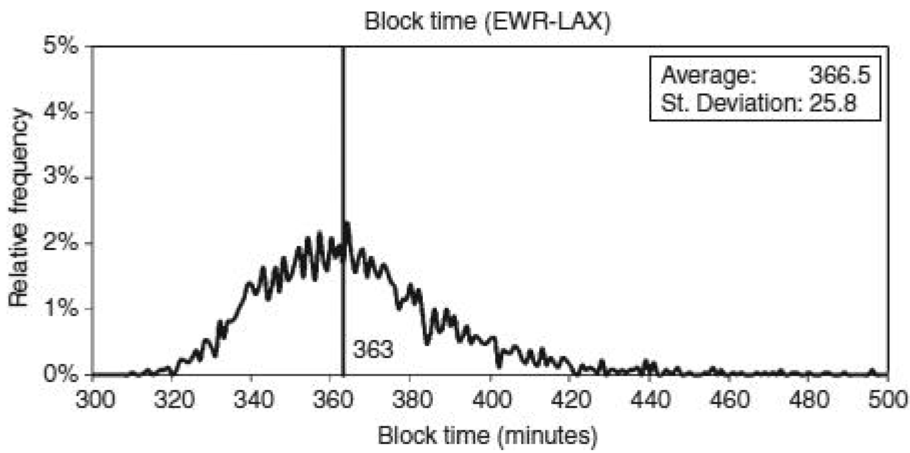


FIGURE 11.9 Distribution of actual gate-to-gate times ("block times") for flights from New York/Newark to Los Angeles/International.

An analogous set of difficult issues arises when delays must be attributed to specific causes. Important examples are given as follows:

1. There is no internationally agreed list of causes to which air traffic delays can be attributed. The International Air Transport Association (IATA) has prepared a long and detailed list of such causes, but different air navigation service providers (ANSP) around the world use different subsets and combinations for their own purposes. Table 11.2 shows the lists used by the FAA and EUROCONTROL in connection with their ASQP and CODA databases. Such lists of attributes tend to change with accumulated experience over time (see EUROCONTROL and U.S. Federal Aviation Administration, 2009).

FAA ASQP Database	EUROCONTROL CODA Database
Extreme weather	Weather
Late aircraft	Reactionary delay
Carrier	Airline
National Airspace System (NAS)	En route constraints
	Airport constraints
Security	Government
	Miscellaneous

Note: The first three causes in each list are roughly equivalent—"weather" in CODA refers to unusually poor meteorological conditions. NAS refers to a broad set of circumstances (excessive traffic volume, nonextreme weather that does not prevent flying) that slow down traffic and generate delays. CODA subdivides these delays into those due to en route and airport constraints. "Government" includes all delays caused by governmental action/inaction, including security interventions.

Table 11.2 Primary Causes of Delay in Two Major Databases

2. Several different causes may contribute to any particular delay. The information available may not be sufficiently detailed to attribute parts of the delay to each of these potentially contributing causes. A prominent example, as already noted, is delay in leaving the departure gate: airline-related problems, late arriving aircraft due to congestion at other airports, and expected congestion after departure (be-

fore takeoff or en route or at the arrival airport) may all play a role in delaying an aircraft's gate departure.

3. The attribution of delay to a particular cause can be a matter of judgment. The most important example is the much-debated question of whether delays that occur in instrument meteorological conditions (IMC), when airport capacities are typically lower (see Chap. 10), should be attributed to "weather" or to "traffic volume," that is to the fact that the normal volume of traffic in such circumstances is often high compared to available capacity. One school of thought on this question argues that the only delays that traffic volume should be blamed for are those that occur when an airport is operating at full capacity, that is, in good weather (visual meteorological conditions). According to this reasoning, delays that occur in IMC should be attributed to weather, because it is the weather conditions that reduce capacity and lead to delays. The opposite side reasons that, with the exception of extreme weather (thunderstorms, snowstorms, icing, very poor visibility), IMC are part of normal weather. A well-functioning air transportation system should be able to cope with IMC and any delays that occur in these conditions should not be attributed to weather, but to the airport's and ATM's inability to handle normal demand adequately. In recent years, this second argument seems to be gaining increasing acceptance internationally. For instance, the ASQP database, as shown in Table 11.2, attributes delays to "weather" only when meteorological conditions are extreme, essentially preventing or seriously inhibiting flight. As a result, only about 5 percent of annual delays are typically attributed to weather conditions in the ASQP database. By contrast, the FAA's OPSNET database adopts the first of the two positions outlined previously and attributes the vast majority of delays (more than 70 percent in a typical year) to "weather."

4. The location where a delay is observed is not necessarily the same as the location of the cause of the delay. For example, in the United States, flights headed to congested airports may be delayed in en route airspace, mostly through path stretching and speed control (see Chap. 13), hundreds of miles away from their destination—or even at their departure gate at the airport of origin. In such cases, it may be difficult to tell from the data where the true cause of the delay was located.

Exercises

(Additional exercises on delays and congestion can be found in Chap. 20.)

11.1. Suppose that an airport has a maximum throughput capacity of 100 movements per hour in good weather, which prevails about 80 percent of the time, and of 60 movements per hour in poor weather (about 20 percent of the time). To estimate delays at this airport, consultant A computed an expected capacity of 92 per hour [= (0.8) (100) + (0.2) (60)]. He then obtained delay estimates through a computer-based queuing model that uses as inputs the daily demand profile at the airport and an airport capacity of 92 per hour. Consultant B used the same computer model as A with the same daily demand profile as A. However, she ran the model twice, once for a capacity of 100 per hour and then for a capacity of 60 per hour. She then took the weighted average of the delays computed through the two runs by multiplying the delays obtained from the first run by 0.8 and those from the second by 0.2.

a. Which consultant's approach is more correct and why? Explain with reference to Fig. 11.3 or 11.4.

b. Which consultant's delay estimates will be higher?

c. Would you use the same daily demand profile for good-weather days and poor-weather days?

11.2. This assignment asks you to assess the impact, if any, that constraints in the airport system have on a major airline with which you are familiar. You should consider the general question of whether the infrastructure of the principal airports that your airline uses is adequate to support its operations. You should also consider what the future portends at these principal airports. Make recommendations on what you think your airline should do regarding the key airports and their constraints. You do not have to address all the issues raised as follows, but the list may be helpful in structuring your report.

a. For your airline, review briefly the situation at its top two airports. Make an assessment about how congested these airports are and their potential for capacity increases.

b. How sensitive is the runway capacity of these airports to weather conditions?

c. Do the airside constraints appear mostly on the surface of the airport (aprons and taxiways, including crossing of runways), at the runway system, or in the airspace?

d. How "delay prone" are the airports you have examined and your airline? Examine some delay statistics for these airports.

e. To what extent does airline scheduling ("banks" or "waves" of connecting flights) contribute to the problems you have identified?

11.3. Consider Exercise 1 of Chap. 10, in which you computed the maximum throughput capacity of a single runway used for arrivals only. Based on that work you can easily compute both the expected value of the service time at this runway, and the variance of the service time. Assume below that the instants when the demands for landing at this runway occur can be approximated as being generated according to a Poisson process.

a. Estimate the PHCAP of the runway, using the approximate queuing expression given in Eq. (20.10).

b. Suppose that the cost of 1 minute of delay in the air is $120, $60, and $20 for H, L, and S aircraft, respectively. Suppose, as well, that this runway has a demand of 27 arrivals per hour for about 6 consecutive peak hours during the day, a demand of 20 per hour for about 10 hours, and a demand of 10 per hour for 8 hours. Estimate approximately the annual delay costs incurred at this runway due to traffic delays. (This part, of course, oversimplifies what happens in practice: in truth, capacity changes over time with weather and runway configuration changes and demand will typically be more variable than is indicated here.)

References

Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., Peterson, E., et al. (2010) *Total Delay Impact Study: A Comprehensive Assessment of the Costs and Impacts of Flight Delay in the United States*, Final Report, prepared for the Federal Aviation Administration, October 2010.

Barnhart, C., Belobaba P., and Odoni A. (2003) "Applications of Operations Research in the Air Transport Industry," *Transportation Science, 37*, pp. 368–391.

Chin, D., Meilus, A., Murphy, D. and Thyagarajan, P. (2012) "Forecasting Airport Delays," Working Paper, Federal Aviation Administration, Washington, DC.

EUROCONTROL and U.S. Federal Aviation Administration (2009) *US/Europe Comparison of ATM-Related Operational Performance*, Report 2009-AJG-333, prepared by the Air Traffic Organization Strategy and Performance Business Unit (FAA) and the Performance Review Unit (EUROCONTROL), Washington, DC and Brussels.

FAA, Federal Aviation Administration (1981) *Airfield and Airspace Capacity/Delay Analysis*, Report FAA-APO-81-14, Office of Aviation Policy and Plans, Washington, DC.

Fan, T. P. and A. Odoni (2002) "A Practical Perspective on Airport Demand Management," *Air Traffic Control Quarterly, 10*, 285–306.

GAO—U.S. Government Accountability Office (2010) *National Airspace System: Setting On-Time Performance Targets at Congested Airports Could Help Focus FAA's Actions*, Report to the Committee on Commerce, Science, and Transportation, U.S. Senate, May 2010.

Hansen, M., Lovell, D., Nikoleris, T., Odoni A., and Vlachou, D. (2009) "Use of Queuing Models to Estimate Delay Savings from 4D Trajectory Precision," *Proceedings 8th USA/Europe Seminar on Air Traffic Management Research & Development*, Napa Valley, CA, July.

Odoni, A., Deyst, J., Feron, E., Hansman, R. J., Simpson, R. W., Delahaye, D., Bowman, J., Khan, K., et al. (1997) *Existing and Required Modeling Capabilities for Evaluating ATM Systems and Concepts*, prepared for NASA's Advanced Air Transportation Program, http://web.mit.edu/aeroastro/www/labs/AATT/index.html.

Pyrgiotis, N. (2012) "A Stochastic and Dynamic Model of Delay Propagation within an Airport Network for Policy Analysis," Ph.D. dissertation, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA.

Skaltsas, G. (2011) "Analysis of Airline Schedule Padding on U.S. Domestic Routes," Master of Science Thesis, MIT Department of Civil and Environmental Engineering, Cambridge, MA.

Stolletz, R. (2008) "Non-stationary Delay Analysis of Runway Systems," *OR Spectrum 30*, 191–213.

---

[1]Only two other years in the decade were profitable, 2007 and 2010, with profits of $4 billion and $2 billion, respectively.

[2]Another way to appreciate the severity of the delays is to consider that airline delays in the United States in 2007 totaled about 2.3 million aircraft hours. With a jet aircraft of a major airline typically performing about 3200 hours of commercial service per year, this is roughly equivalent to wasting the potential services of the equivalent of 720 jet aircraft per year—about the same as the size of the largest airline in the United States (Delta) or the combined size of the fleets of Lufthansa, Air France, and British Airways!

[3]As Sec. 10.3 explains, the capacity of the runway system also varies with the mix of arrivals and departures. The capacity is typically higher during periods when the demand consists mostly of departures and is lower when the reverse is true. However, for the purposes of this example, it is assumed that the arrival–departure mix is about 50 to 50% during all hours. In this way the hourly capacity can be approximated by a single value that may vary by time of day, depending only on weather conditions.

[4]The reader will find more details in Sec. 20.3, which derives this result through an example based on the notion of *cumulative diagrams*, which can be very useful for approximate analyses of queuing phenomena during periods when demand exceeds capacity. This is often the case, for instance, when a passing weather front reduces considerably the capacity of a runway system for several consecutive hours.

[5]The exact physical location of these queues depends on ATM operating policies and on the length of the queues themselves. For example, in the case of departing aircraft, the queue(s) in most instances will form on the taxiway(s) leading up to the departure runway(s); however, when the taxiway queue becomes very long, airplanes will often be held at their departure stands, so that a second queue of airplanes waiting to enter the taxiway system is created (see Chap. 13).

[6]The demand profile graph, Dem, connects with straight-line segments the 24 points that indicate the demand rate during each of the 24 hours of the day. For example, the demand rates between 11:00 and 12:00 and between 12:00 and 13:00 are equal to 60 and 71 movements per hour, respectively. For the purpose of computing delays, these values were assigned to the midhour points (11:30 and 12:30, respectively). The instantaneous demand rate between 11:30 and 12:30 is then read from the Dem graph by interpolating between the values of the demand rate at the two midhour points.

[7]In fact, the delays may be so bad on some days that the airlines may decide to cancel some flights. Indeed, cancellations often act as a safety valve at congested airports. In effect, they reduce demand during days when delays are at their worst and consequently also reduce the delays suffered by the flights that are actually performed.

[8]However, a noticeable time lag of the delay peak may exist even when the demand rate never exceeds (but is close to) capacity.

[9]In practice, of course, the "leftover movements" would probably be cancelled each day, indicating again that, over a long period, a demand rate that exceeds capacity is unsustainable.

[10]In most instances the wait would take place while the aircraft is airborne. However, in cases of extremely long delays, the aircraft may be "held" on the ground at its airport of origin to avoid excessive waiting times in the air.

[11] As Chap. 12 indicates, most airports outside the United States use their maximum throughput capacity in IMC as the basis for setting their declared capacity.

[12]Despite the large drop in the number of annual movements, the number of annual passengers was essentially the same in 2000 and 2010 (27.7 and 27.4 million, respectively).

[13]As Chap. 10 explains, such local policies involve consideration of ATM procedures as well as of environmental impacts.

CHAPTER 12

Demand Management

Demand management refers to any set of administrative or economic measures and regulations aimed at constraining the demand for access to a busy airfield and/or at modifying the temporal characteristics of such demand. It is practiced extensively at major airports throughout the world, but its use in the United States is limited.

The available approaches to airport demand management can be subdivided into three categories: purely administrative, purely economic, and hybrids, that is, combinations of the other two. The fundamental justification for the last two categories is that they allow economics and market-based mechanisms to play a role in determining access to congested airports. Congestion pricing, for instance, forces aircraft operators to consider the delay costs they impose on other airport users—costs that are largely ignored under most circumstances at present.

Schedule coordination, conducted under the aegis of the International Air Transport Association (IATA), is currently used at the great majority of the busiest airports outside the United States. This is an administrative demand management procedure that uses historical precedent as the primary criterion for allocating airport slots. It can be effective at mildly congested airports, but may cause serious market distortion and affect competition adversely at airports where unconstrained demand exceeds available capacity by a significant margin.

All purely economic approaches to demand management involve some form of congestion pricing; that is, they are based on the principle that, to optimize use of a congested facility, users should be forced to internalize the external costs imposed by their use of the facility. Congestion pricing is an economically efficient approach, but may be difficult to implement for both technical and political reasons.

Hybrid demand management systems combine elements of administrative and economic approaches. Their common characteristic is the use of an administrative procedure to specify the number of slots available at an airport, followed by the application of such economics-based schemes as congestion pricing, slot markets, and auctions, to arrive at the final allocation of slots. The worldwide use of hybrid demand management systems will probably increase significantly in the future.

12.1 Introduction

The ever-tighter relationship between demand and capacity at many of the world's major commercial airports, and the worsening air traffic congestion have led to growing interest in and widespread application of airport demand management measures. Demand management refers to any set of regulations or other interventions aimed at constraining the demand for access to a busy airfield and/or at modifying the temporal characteristics of such demand. Examples are slot restrictions and airport pricing schemes aimed at discouraging the scheduling of flights during peak traffic hours and inducing airlines to shift some operations to off-peak hours.

Until the early 1980s, demand management was practiced at only a small number of airports around the world. Because of its potential for adversely affecting competition, demand management was viewed as only a method of last resort for reducing airport congestion and congestion costs. However, with few exceptions, notably the United States, the debate on demand management today has shifted from whether it should be used at all to how it can be applied most effectively.

The overall premise is as follows: Capacity expansion should generally be the principal means for accommodating growth in airport demand, but may require a long time or may even be entirely infeasible due to local or other constraints. In such circumstances, some form of demand management may be the only available alternative, at least in the short and medium terms, for keeping delays within reasonable bounds.

Demand management is currently practiced, in one way or another, at the great majority of the busiest airports outside the United States, as well as at many secondary ones. It is viewed as an essential complement, on the demand side, to "supply-side" efforts to increase capacity. As this chapter shows, all demand management approaches have some weaknesses, recognized even by their proponents. It is typically argued, however, that the overall effects on competition and on access to airports and markets are mild compared to the significant benefits that stem from the resulting reduction in air traffic congestion. Regardless of their own views on demand management, airport professionals need to be familiar with this subject, as its implications for current and future airport operations, and for the air transportation system in general, are potentially very important.

This chapter reviews the principal approaches to demand management, as currently practiced or proposed, and discusses some of the advantages and disadvantages of each. The approaches can be subdivided into three categories: purely administrative, purely economic, and hybrids, that is, combinations of the first two. One of the fundamental objectives of the last two categories is to force aircraft operators to consider, directly or indirectly, the external costs that stem from their use of congested facilities, including the delay costs they impose on other users. In the absence of purely economic or hybrid demand management measures, airport users can largely ignore these external costs.

By definition, purely administrative approaches do not use economic incentives, such as landing fees that may vary by time of day, to influence the choices of prospective users concerning the time when they will operate at an airport or even whether they will operate there at all. If such economic incentives are used, along with administrative measures, the approach is classified as a hybrid. In purely economic approaches, on the other hand, there is no administrative interference with airline choices: after a set of prices for accessing an airport at different times of the day has been put in place, decisions on whether and when to operate are left entirely to the airline. These three types of demand management approaches are covered in Secs. 12.3 through 12.5, whereas Sec. 12.2 recaps briefly the rationale for applying any form of demand management at congested airports. Section 12.6 presents some policy-related conclusions.

It is important to clarify at the outset that this chapter is concerned with *strategic* demand management approaches, that is, measures that become part of the institutional and regulatory framework within which an airport is operated. Some air traffic management (ATM) organizations, notably EUROCONTROL and the U.S. Federal Aviation Administration (FAA), also apply a *tactical* form of demand management to relieve airport congestion on a day-to-day basis: *air traffic flow management* (ATFM) controls the flow of traffic into congested airports and congested parts of the airspace on a dynamic, "real-time" basis. A typical tactic is to postpone for some time the departure of an airplane if it is expected that, once airborne, it will be subjected to a long delay. ATFM is described in Chap. 13.

Moreover, this chapter refers primarily to approaches for managing access to an airport's *system of runways*. This is by far the most common context in which demand management is applied. Entirely analogous ideas can be applied in the context of access to passenger terminals, aprons, or other parts of the airport. In these latter cases, passenger service charges or aircraft parking charges may sometimes be used as the instrument for managing demand (see Sec. 12.5 and, especially, Chap. 8). It should be noted, however, that the coordination and harmonization of landside demand management measures with those aimed at the airside can be a complex task and may involve difficult economic and regulatory issues. Very few airports practice simultaneous airside *and* landside demand management at this time.

A large number of scholarly writings and papers on airport demand management have appeared over the years beginning with the late 1960s. This chapter provides some references. A good starting point for a far more extensive coverage of the subject is the volume edited by Czerny et al. (2008).

12.2 Background and Motivation

The principal objective of demand management is to assist in maintaining efficient operations at airports that are threatened by congestion. This is not done through capital investments or changes in traffic handling procedures aimed at increasing capacity, but through regulations or other measures that aim at some combination of (1) reducing overall demand for airfield operations, (2) limiting demand during certain hours of the day, and (3) shifting demand from certain critical time periods to other, less critical ones. The means used to accomplish this objective is what differentiates one demand management approach from another. In all cases, the net effect is to constrain in some way access to the airport, either at all times or during targeted periods. For this reason demand management is also often referred to as *access control*.

In the absence of a demand management program, access to commercial airports throughout the world is governed by a simple rule: Any aircraft technically qualified to operate at a particular airport, that is, fulfilling the relevant ATM and airworthiness requirements, can utilize that airport by paying a landing fee proportional to the weight of the aircraft.[1] In this environment, an airline may schedule a flight at the airport for any time it wishes, outside any curfew hours that may exist.[2] A demand management program modifies these conditions. For example, an upper limit may be placed on the number of operations that can be scheduled during a particular period of the day, or a surcharge may be imposed on the landing fee during peak hours.

The motivation for demand management comes directly from a fundamental observation in Chap. 11, namely, that when the utilization of a service facility is high (or when demand approaches the capacity of a system), the relationship between delay, on the one hand, and capacity or demand, on the other, becomes *very nonlinear:* a small increase in capacity or a small reduction in the demand rate results in a proportionally much larger reduction in delay (with the reverse also being true). Demand management aims at achieving those small reductions in demand (or the shifts in demand from peak to off-peak periods) that will bring about these large delay benefits. In the process, additional benefits may be achieved, such as reduced operating costs through a more efficient utilization of available personnel, equipment, and resources.

It is sometimes argued that demand management of any form is unnecessary, even at the busiest airports, because delay will act by itself as a “natural” access-control mechanism. According to this argument, as delays at an airport increase, more and more aircraft operators will deem the situation unacceptable and will choose not to use the airport. At some point, the costs associated with delay, as perceived by individual users, will become so high that demand will cease to grow and equilibrium will be reached.

This line of reasoning misses a critical point. The equilibrium reached in this way will, in general, be inefficient economically, as (1) the associated level of delays to aircraft and passengers will be excessive and (2) the resulting mix of airport users may include a large fraction of users who have a low value of time and whose use of the airport cannot be justified on economic grounds.[3] Section 12.4 explains this in detail. Suffice to note here that a growing number of aviation experts, managers, and operators have come to realize that the "do nothing" alternative (i.e., allowing demand to grow unabated until the users themselves become discouraged by the high cost of delays) is wasteful and inefficient. This has motivated the extensive ongoing examination of the relative merits and effectiveness of the various demand management approaches.

Example 12.1 illustrates what may happen if delay is allowed to serve as the only access-control mechanism. It also indicates the large benefits, in terms of delay reductions, that may be obtained from even primitive demand management measures under the right conditions.

Example 12.1 For more than 30 years before 2000, the number of aircraft operations at New York/LaGuardia (LGA) was constrained by the number of slots authorized under the FAA's High Density Rule (HDR)—see Sec. 12.3. Early in the spring of 2000, an average of about 1050 aircraft movements (arrivals and departures) took place at LGA on a typical weekday. However, the Aviation Investment and Reform Act for the 21st Century (AIR-21), enacted in April 2000, exempted from HDR slot limitations aircraft with a capacity of 70 seats or fewer performing scheduled flights between LGA and small airports in the region.

In the first 7 months after AIR-21 was enacted, airlines sought to schedule more than 600 new movements a day at LGA. As of November 2000, about 300 of those new movements had begun operations, bringing the average number on a typical weekday to 1350. The result was unprecedented levels of delay and numerous flight cancellations on a daily basis. LGA alone accounted for more than 25 percent of the serious delays (more than 15 minutes) experienced at *all* commercial U.S. airports in the fall of 2000. Yet airlines kept announcing the scheduling of additional regional flights at the airport.

As an interim solution, the FAA, with strong support from LGA's operator, the Port Authority of New York and New Jersey, imposed a limit on the number of slot exemptions granted under AIR-21. It allocated the pool of "AIR-21 slots" among eligible flights through a lottery that took effect on January 31, 2001. The lottery was designed to impose an hourly cap of approximately 75 scheduled movements per hour, a number that the airport was deemed able to accommodate at reasonable levels of delay in good weather conditions.

Indeed, the severity of delays and the number of cancellations at LGA declined enormously after January 2001 from the levels reached in the fall of 2000. The slot lottery reduced the number of weekday scheduled flight movements (takeoffs and landings) from about 1350 in November 2000 to 1205 in August 2001, a roughly 10 percent reduction. Figure 12.1 compares the profile of hourly flight operations before and after the lottery. Note that the level of flight operations prior to the slot lottery exceeded the sustainable capacity of 75 for most of the day, with virtually no time for schedule recovery. Figure 12.2 shows the consequent impact on delays. Prior to the slot lottery, delay rose continuously from the early morning till 8 p.m., reflecting the fact that scheduled demand exceeded (good-weather) capacity throughout that period. At its daily peak, expected delay was more than 80 minutes per movement in November 2000; it declined to a peak value of about 15 minutes per movement in August 2001 (Fan and Odoni, 2001). Total expected delay for a typical weekday was about 900 aircraft-hours prior to the slot lottery, but only about 150 aircraft-hours after it. The 10 percent reduction in the number of movements imposed through the lottery thus led to a reduction of 80 percent in total aircraft delays! At a then-estimated average of about $1600 per hour in direct operating costs, this translated to savings of about $1.2 million in direct operating costs *per weekday* for the airlines, not including the savings associated with less passenger waiting time and reduced schedule recovery costs.

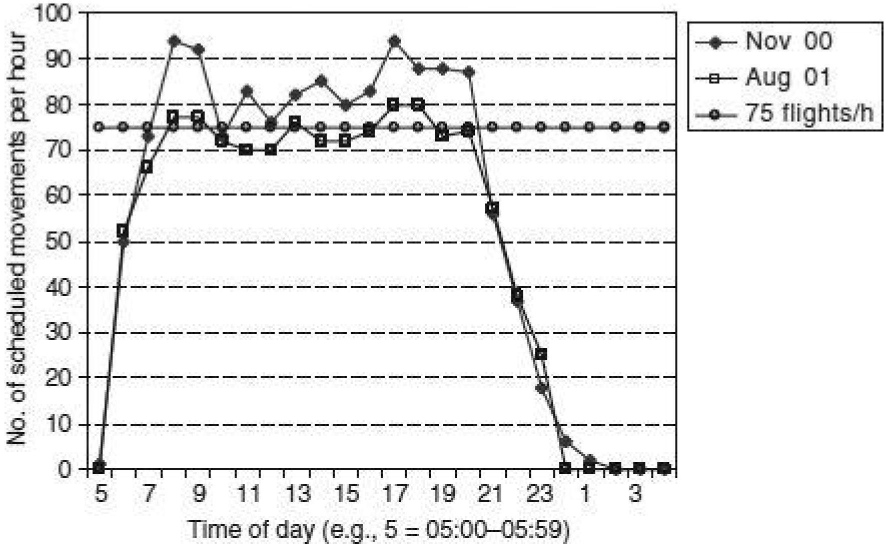


FIGURE 12.1 Profile of scheduled movements at New York/LaGuardia before and after the 2001 slot lottery.

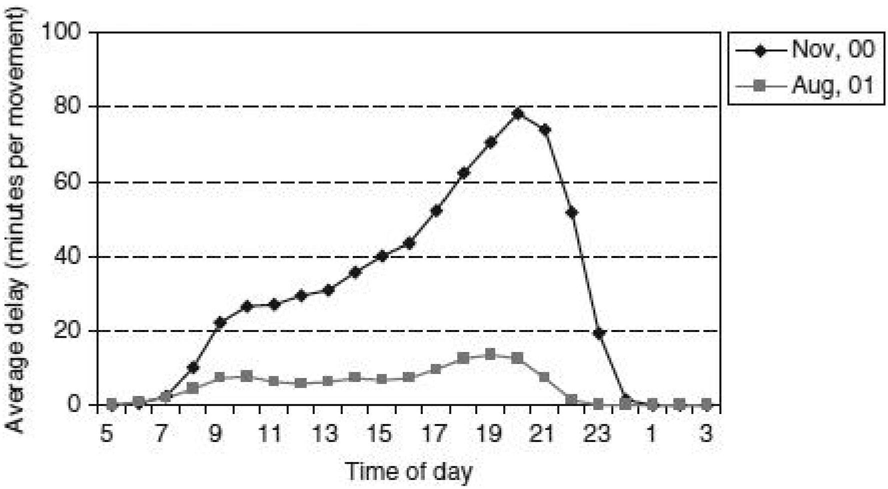


FIGURE 12.2 Profile of average delay per aircraft movement at New York/LaGuardia before and after the 2001 slot lottery.

The FAA also stated at the time its intention to replace the lottery with a more carefully designed long-term demand management mechanism for preventing excessive congestion at LGA. In June 2001 it published a Notice for Public Comment (FAA, 2001), which described several interesting potential approaches for managing demand at LGA in the future. The alternatives included schemes variously based on purely administrative allocation, on congestion pricing, and on auctioning of slots (see Secs. 12.3–12.5). These plans were abandoned after the rapid decline of traffic that LGA experienced following the events of September 11, 2001. However, interest in demand management measures at New York's airports was strong again by the end of the decade, following the traffic's recovery a few years later (see Sec. 12.5).

12.3 Administrative Approaches to Demand Management

A fundamental element of all administrative approaches to demand management is the concept of a slot. A *slot* is an interval of time reserved for the arrival or the departure of a flight and is allocated to an airline or other aircraft operator for a specified set of dates. Thus, the statement, "Airport X can offer up to 60 slots between 09:00 and 10:00 local time for the summer season of 2011," means that the administrative entity responsible for Airport X is prepared to allocate among prospective aircraft operators up to 60 time intervals for scheduling arrivals and departures during the hour in question. Once a slot is allocated, it typically becomes associated with a specific flight operation, for example, "Airline A has been given the 09:10 slot for the arrival of its Flight 124 at Airport X during the weekdays of the summer season of 2011." This does not mean that Flight 124 necessarily *has* to land at 09:10, but that Flight 124 has been *scheduled* to arrive at 09:10 and will be expected to arrive at about this time on every weekday (Monday-Friday) of the summer season of 2011.

Administrative approaches to demand management require the selection of a set of criteria for allocating slots among prospective users. Some examples of reasonable criteria and of related considerations may include the following:

- The length of time for which a flight has already been operating in the past (flights that have been operating for a long time may be deemed to deserve priority for continuation of service)
- The regularity of the flight (scheduled flights operated on a daily or weekly basis may be given preference over occasional or charter flights)
- The origin or destination of the flight (service to/from certain locations or to new markets may be deemed particularly important)
- The characteristics of the airline requesting the slot (in the interest of more competitive service, for example, "new entrant" airlines that have not previously served a particular route or a particular airport may be given priority for slots)

Schedule Coordination: The IATA Approach

All the aforementioned criteria are applied to some degree in connection with the *schedule coordination approach* of the IATA, which is used, with some regional variations, in most of the world. The following description omits many important details that can be found in the *Worldwide Slot Guidelines*[4] (IATA, 2012) henceforth referred to as WSG.

For purposes of schedule coordination, airports are classified into three categories; Level 1, Level 2, and Level 3. *Level 1* airports are those whose capacities are adequate to meet the demands of users and require no demand management interventions. *Level 2* (or "*schedules facilitated*") airports are those where "there is potential for congestion… which can be resolved by voluntary cooperation between airlines" (IATA, 2012) with assistance from an independent "facilitator" appointed by "the responsible authority." *Level 3* (or "*schedule coordinated*") airports are those satisfying the following set of conditions (IATA, 2012):

(a) Demand for airport infrastructure significantly exceeds the airport's capacity during [parts of the scheduling] period;
(b) expansion of airport infrastructure to meet demand is not possible in the short term;
(c) attempts to resolve the problem through voluntary schedule adjustments have failed or are ineffective; and
(d) as a result, a process of slot allocation is required whereby it is necessary for all airlines and other aircraft operators to have a slot allocated by a coordinator in order to arrive or depart at the airport during the periods when slot allocation occurs.

When an airport is designated as Level 3, a schedule coordinator is appointed and assigned the task of resolving schedule conflicts and allocating available slots. A local committee of experts and stakeholders usually supports the coordinator. *All* requests for slots at Level 3 airports must be reviewed and cleared by the schedule coordinator. As of 2012, approximately 170 airports worldwide have been designated as Level 3 (89 of them in Europe) and use some version of the IATA's schedule coordination approach. The list includes practically every one of the busiest airports of Europe, Asia, and the Pacific Rim, as well as many secondary ones.

Schedule coordination is carried out at Schedule Coordination Conferences (SCCs) organized by the IATA every November and June[5] and attended by numerous representatives of airports, airlines, and civil aviation organizations from around the world. Each fully coordinated airport must first specify a *declared capacity* (see Chap. 10), which indicates the number of aircraft movements per hour (or per other unit of time) that the airport can accommodate. Under the IATA system, responsibility for determining the declared capacity of each airport rests with local and national authorities. Declared capacity need not be determined solely by the capacity of the runway system. Constraints due to the availability of aircraft stands, passenger terminal processing capacity, and even aircraft ramp servicing capacity can be taken into consideration. This is one of the reasons that the list of Level 3 airports includes a number of secondary airports in countries with highly seasonal traffic, such as Spain, Italy, and Greece. Many of these airports have severe landside capacity constraints and can therefore process passengers from only a very limited number of movements per hour during peak months.

Prospective users of Level 3 airports must submit a formal request for each and every desired slot. The declared capacity is rationed according to a set of criteria, among which the principal and overriding one is *historical precedent*: an aircraft operator who was assigned a slot in the equivalent previous season ("summer" or "winter") and utilized that slot for at least 80 percent of the time during that previous season is entitled to continued use of that "historical slot." Second priority is assigned to requests for changing the time of historical slots. In addition, a slot awarded on the basis of historical precedent may be used in the new season to serve a different destination from the one served in the previous season. Slot exchanges and slot transfers between airlines are also allowed. Such exchanges may involve compensation from one airline to another, at airports where this is legally permitted.

Any unassigned slots, or slots for services that have been discontinued, or slots not used at least 80 percent of the time when they were supposed to be used,[6] become part of a "slot pool" for reallocation. Any new slots made available through increased airport capacity (e.g., as a result of the opening of a new runway or of improvements in ATM) are also placed in the slot pool. All requests for new slots are served from the slot pool.

To encourage competition, establish new markets, and strengthen previously underserved ones, at least 50 percent of the slots in the pool within each coordinated time interval must be allocated to airlines designated as *new entrants*, assuming that a sufficient number of such slot requests exists. However, the definition of a "new entrant" is restrictive: An aircraft operator qualifies for this designation as long as it does not hold more than four slots in a day, *after* receiving any new slots from the slot pool (IATA, 2012). Note that this amounts to a severe regulatory constraint: A new entrant is essentially limited to at most two flights (or four runway movements) per day, hardly sufficient to establish a significant foothold at a major airport. This constraint precludes, in effect, a type of competitive strategy that airlines have used frequently and effectively in the United States during the deregulation era. The strategy consists of setting up, within a short time, a large number of flights at an airport where an airline was not previously operating or had a minor presence. In this way, the airline in question reaches "critical mass" overnight and becomes a major competitor at the airport.

After new entrants, priority for new slots is given to requests for extending seasonal scheduled service (previous winter or previous summer) to year-round scheduled service. Any further remaining slots after this step are distributed according to a number of additional criteria, such as the size and type of market involved, contribution to competition on routes, the existence of any curfews at the airports of flight origin or flight destination, etc. The WSG includes many other detailed guidelines, such as provisions for shared use of a slot (e.g., in the case of airline alliances), disposition of slots held by an airline that ceases operations, obligations of slot holders (e.g., reasonable adherence to the timing of the slots), etc.

The schedule coordinator obviously plays a central role in this process. The manner in which the coordinator is selected varies from country to country and even from one airport to another in some countries. In many cases, the national airline or a major airline of the country is asked to designate an experienced employee or a team of employees to serve in this capacity.

There are also significant variations in the level of sophistication with which these demand management procedures are applied at different airports. For example, some airports utilize a simple limit on the number of movements that can be scheduled in any single hour of the day, whereas others use combinations of limits that may restrict the number of movements for intervals smaller than an hour. Brussels Airport, for example, imposes limits on the number of movements that can be scheduled over 5-, 10-, 30-, and 60-minute intervals, as shown in Table 12.1. The objective is to "even out" any intrahour peaks in the traffic schedule. Moreover, limits are specified on the number of arrivals, the number of departures, and the total of the two. Note that the limit on total operations is very often not equal to the sum of the limits on arrivals and departures. This is because of the interdependence of operations on different runways at many airports (see Chap. 10). Environmental constraints may drive limits specified for the nighttime hours between 23:00 and 6:55. At some airports limits may also be set for environmental reasons on the total number of slots in a day (e.g. at Tokyo/Narita), or even in a week or an entire year.

Times/Period		05 min			10 min			30 min			60 min		
From	Until	Arr	Dep	Total	Arr	Dep	Total	Arr	Dep	Total	Arr	Dep	Total
00.00	05.55	5	5	8	9	9	9	16	16	16	30	30	30
06.00	06.55	5	5	9	10	9	13	24	27	35	35	40	45
07.00	22.55	6	5	10	10	9	15	30	27	40	48	44	74
23.00	23.55	5	5	8	9	9	9	16	16	16	30	30	30

Source: Morisset, 2010.

TABLE 12.1 Declared Capacities Table for Brussels Airport, 2009

Declared capacities are typically adjusted from year to year and may increase gradually as a result of airport infrastructure expansion, air traffic control improvements, and airport operator experience with the slot coordination system.

Possibly the most advanced procedures for setting declared capacities are the ones in use at London/Heathrow. The declared hourly capacity may change by time-of-day and also depends on the mix of departures and arrivals in each hour. Table 12.2 compares the number of available slots in the summer seasons of 2008 and 2009. Note that, while the average number of slots per hour was 80.4 in both years, a total of 88 slots were made available during the peak demand hour of the day (17:00–18:00) in both years. The exact number of slots is determined through an extensive set of simulation runs with the objective of maintaining the average delay per flight to 10 minutes or less, taking historical weather conditions into consideration. The number of slots available in the summer and winter seasons thus changes markedly, reflecting less favorable conditions in the winter. The declared capacity is also adjusted from year to year in response to ATM developments and changes in aircraft mix. Note how marginal the changes were between 2008 and 2009. Overall, the number of daily slots at London/Heathrow has increased by about 10 percent between the summer seasons of 1991 and 2010.

Hour	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	Total	Average
Arrivals																			
Summer 2008	38	39	37	39	39	42	40	43	43	43	42	42	44	43	38	44	20	676	39.8
Change				+1		–1				–1		+1							
Summer 2009	38	39	37	40	39	41	40	43	43	42	42	43	44	43	38	44	20	676	39.8
Departures																			
Summer 2008	27	43	43	42	40	42	41	43	42	44	44	43	44	43	39	40	30	690	40.6
Change	–1	+1	–1		+1												+1	+1	
Summer 2009	26	44	42	42	41	42	41	43	42	44	44	43	44	43	39	40	31	691	40.6

Note: A complex set of restrictions limit traffic severely between 22:00 and 05:00.

Table 12.2 Declared Capacity at London/Heathrow: Summer of 2008 vs. Summer of 2009

In summary, the main appeal of the IATA's purely administrative schedule coordination approach is that it "has been singularly successful in maintaining a high degree of coherence and stability in the international air transport system"[7] (IATA, 2000). Indeed, this demand management mechanism has worked well in practice in instances where demand exceeds the supply of airport capacity by a relatively few operations and for only a small number of hours in a day.

However, when a significant excess of demand over capacity exists, there is a clear risk that the IATA approach may lead to serious distortions of the marketplace. Indeed, it can be argued that, at some of the most congested airports in the world, the schedule coordination process currently serves as a means for preserving the *status quo*, effectively acting as a regulatory device at the airport level. New competitors may be prevented from entering markets effectively: they may be denied slots altogether or be relegated to slots at inconvenient times of the day or awarded far fewer slots than necessary to establish a truly competitive schedule of flights. Some airports have in fact gone as far as segmenting airport capacity to serve perceived or alleged public policy goals. Examples include the designation of blocks of slots that are reserved for international, domestic, and general aviation traffic at Tokyo/Narita or for regional services within the State of New South Wales at Sydney Airport.

EU Modifications of IATA's WSG

In response to such concerns, a number of governments around the world have been examining closely the IATA's schedule coordination procedures. This has been particularly true in the European Union (EU), whose air transport system was largely deregulated in 1993. Beginning with Regulation 95/93, which became effective in January 1993, and especially through Regulation 793/2004, which amended 95/93 in 2004, the EU has attempted to modify the IATA schedule coordination approach and make it more impartial and more accommodating of change and competition at airports. These attempts at improving the schedule coordination process provide a good illustration of the many complications that must be dealt with. Noteworthy provisions in the EU regulatory framework concern: schedule coordinators, new entrants, "ownership" of slots, and a process of appeals and enforcement.

The EU explicitly assigns to national governments responsibility for appointing schedule coordinators, thus aiming to take slot allocation out of the hands of the airlines. This has led to the gradual evolution in some EU countries of state-sponsored units that specialize in schedule coordination. For example, the German Airport Coordinator (GAC) has a staff of 15 full-time people who coordinate slot allocation at 16 German airports, including Frankfurt/International and Munich, with a total of about 2 million annual movements. On the other hand, several EU Member States have continued to assign to their flag carriers, in various ways, a central role in schedule coordination in their national territories, reasoning that these airlines possess the necessary expertise for the task.

The definition of new entrant airlines in Regulation 793/2004 is significantly more "liberal" than in the IATA guidelines. An airline *may* qualify as a new entrant if it holds fewer than 5 percent of the slots[8] at an airport on the day for which it requests a new slot (European Commission, 2004). If it passes this test, route-specific criteria are applied to determine whether the airline is entitled to new entrant status. Consider, for example, a route between the subject airport and a regional airport in the EU—defined as one with fewer than 5 million passengers per year—and suppose no carrier currently provides nonstop scheduled service on that route. In this case, according to 793/2004, a carrier that requests slots to provide nonstop scheduled service on this route is considered a new entrant if, after that carrier's request was accepted, it would hold fewer than five slots at the subject airport on that day *for that nonstop service*. Note that under this rule, the carrier in question may operate two flights a day on this particular route and may still qualify for new entrant status on other routes between the subject airport and other regional airports in the EU. For the complete set of rules concerning new entrants, see European Commission (2004).

Regulation 793/2004 also importantly addressed the crucial question of slot "ownership." Historically, airlines have claimed grandfathered airport slots as their own assets, whereas airport operators have argued that slots constitute airport property, because airport infrastructure enables the existence of such slots. The position of Regulation 793/2004 is that slots are an "entitlement"; that is, they represent only an airline's right to use an airport's infrastructure at the time and for a period specified. In other words, a slot is a public good allocated to an airline under a specific set of conditions. This allocation may thus be withdrawn "without giving the concerned airline a legal claim" and the slots may be returned to the slot pool (Kilian, 2008).

Finally, the EU regulations establish a process for appealing slot allocation decisions, first, to the schedule coordinator, then to the Member State where the airport is located, and eventually to the European Commission itself. The Commission, in fact, retains the right to cancel some or all of the slot allocations made by a schedule coordinator. More significantly from a practical viewpoint, the EU regulations provide schedule coordinators with considerable powers of enforcement. For instance, the GAC has the right to impose penalties on airlines for "misbehavior"[9] in complying with schedule coordination rules. These penalties may range from lower prioritization of the carrier in the slot allocation process, to imposition of a fine for repeated misbehavior, to withdrawal of the slot from the airline. The GAC monitors closely slot adherence, receiving relevant data weekly from DFS (the German Air Navigation Service Provider or ANSP) and sending to the airlines written inquiries to which a response is mandatory.

The EU has continued to address issues related to the allocation of slots at EU airports beyond Regulation 793/2004. In a major proposal issued in December 2011, the European Commission (2011) indicated its intent to move toward adopting some market-based mechanisms for slot allocation by establishing a "transparent framework" within which airlines may engage in "secondary trading," that is, the exchange or transfer of slots with monetary compensation. It has also suggested a further "liberalization" of the conditions that new entrants must satisfy, and some tightening of the grandfathering rules, such as raising the use-it-or-lose-it limit to 85 percent from the current 80 percent.

Experience in the United States

As demand for slots at a congested airport increases, so does the complexity of the task of the schedule coordinator. The situation may eventually become untenable: no matter how slots are allocated administratively, the interests of several carriers will be materially damaged. The rejection of many slot requests will certainly lead to distortion of market forces and dilution of competition.

A case in point is the experience with schedule coordination in the United States during the years immediately following airline deregulation in 1978. [For a detailed description see U.S. Department of Transportation (1995).] In November 1968, the United States initiated the airport HDR. Five busy airports, New York/Kennedy, New York/LaGuardia, New York/Newark, Washington/Reagan, and Chicago/O'Hare, were designated as HDR airports and hourly limits were placed on the number of operations that could be scheduled at each of them.[10]

Two schedule coordination committees were established at each HDR airport, one to administer and coordinate slots designated for use by air carriers and the other to do the same for slots designated for commuter airlines. After a difficult first meeting in 1969, which required a full month of intense negotiations before consensus was reached, subsequent HDR meetings generally went smoothly for more than a decade, with participants making minor adjustments from year to year to the schedules and slot assignments. This was before the deregulation era, when there were few, if any, new entrants from season to season into routes served by the airports involved. In fact, HDR made no special provisions for transferring slots to new entrants.

The situation changed soon after the 1978 deregulation of the airline industry. In the new environment, prospective new entrants pressed for the acquisition of large numbers of slots at the four HDR airports. The schedule coordination committees were unable to satisfy these requests, despite some changes that the Department of Transportation (DOT) made to the HDR.[11] After several years during which committees were deadlocked virtually continuously, the DOT finally abolished the committees in 1985. The Department declared that slots at the four HDR airports were to be available for use by the airlines that held them as of that date, and established a "buy-and-sell" (or "secondary trading") environment for the slots (see Sec. 12.5). Schedule coordination committees have not operated in the United States since 1985.

The HDR was abolished on January 1, 2007, in accordance with legislation passed by Congress in 2000 and after a 7-year transitional period. Slot limits were maintained at only two airports after that date, New York/LaGuardia and Washington/Reagan. Unfortunately, the abolition of HDR coincided with the year (2007) when demand for runway movements reached its highest point ever at many airports in the United States. In particular, two former HDR airports, New York/Kennedy and New York/Newark, operated in near-gridlock conditions for the entire year. This episode illustrates what may happen in the absence of scheduling limits (see Sec. 12.4) at airports where demand is greater than or nearly equal to available capacity. Acting independently of one another, airlines arrived at daily schedules of movements whose numbers often exceeded even the visual meteorological condition (VMC) capacities of these two airports for several hours on a typical day. This is shown in Fig. 12.3 for New York/Newark. Note that the demand exceeded the IMC capacity of the airport (about 72–76 movements per hour) for roughly 7 consecutive hours in the afternoon and evening, meaning very large delays on any day when weather was less than optimal. Indeed, delays relative to schedule at New York/Newark in 2007 were extremely large, with *daily* averages of roughly 20 and 30 minutes per arrival and departure, respectively, and much higher values in the afternoon and evening.

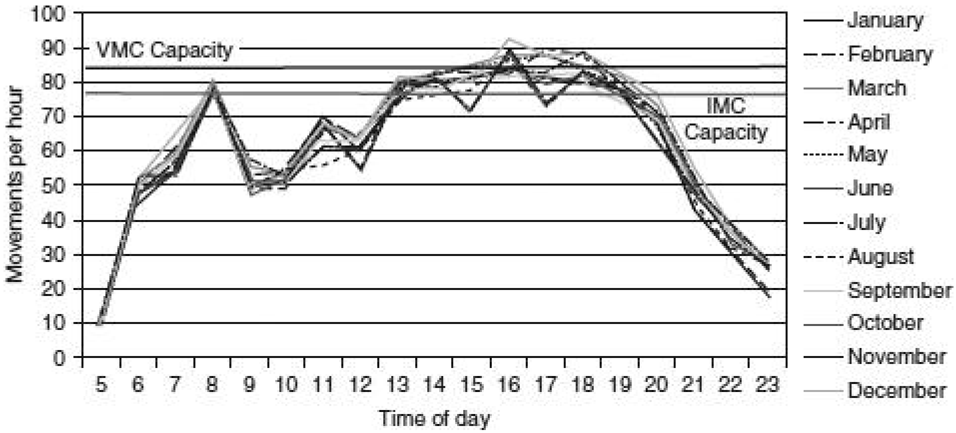


FIGURE 12.3 Average daily demand profile per month at New York/Newark, 2007. (*Source*: Morisset, 2010.)

Table 12.3 casts more light on the situation. It shows the average delay relative to schedule experienced by all arrivals at these airports in all of 2007 for selected times of the day (Morisset, 2010). It also compares these values to delays in Frankfurt/International, which is considered one of the most congested airports in Europe, but also one where strict slot limits are applied and enforced. As Table 12.3 shows, both New York airports simply could not keep up with demand over the course of the day. Whereas the average delay (relative to schedule) for aircraft scheduled to arrive between 8 and 9 a.m. at New York/Kennedy was 7.6 minutes, this had increased to 34.7 minutes by 8–9 p.m. Moreover, the variability of the delay, as measured by its standard deviation, increased rapidly throughout the day, reaching more than 1 hour (!) between 8 and 9 p.m. Thus, in addition to the very large delays that arriving passengers experienced on average, schedule reliability at both airports was extremely low in the afternoon and evening hours—with all the attendant negative consequences for people, airlines, immigration and customs services, etc. In sharp contrast, delays were far more modest and stable throughout the day at Frankfurt/International.

	Frankfurt/International		New York/Kennedy		New York/Newark	
Hour	Average	Std. dev.	Average	Std. dev.	Average	Std. dev.
8–9 a.m.	0.9	24.9	7.6	46.6	–0.2	29.0
12–1 p.m.	4.1	30.6	10.8	49.3	8.1	38.1
4–5 p.m.	1.7	28.4	23.2	56.9	30.1	55.0
8–9 p.m.	7.3	23.5	34.7	65.4	32.4	52.9

Source: Morisset, 2010.

TABLE 12.3 Average and Standard Deviation of Delays (in min) for All Flights Scheduled to Arrive during the Indicated Hours in 2007

In response to these nearly disastrous conditions, the FAA in the beginning of 2008 imposed scheduling limits at both Kennedy and Newark, in addition to the preexisting limits at New York/La Guardia. The limits for the two airports were set to 81 movements per hour plus two slots held in reserve for unscheduled flights (air taxi or general aviation). The total of 83 is very close to the VMC capacities of both Kennedy and Newark and has been strongly criticized as too high in a report of the Office of Inspector General of the U.S. DOT (2010).[12]

12.4 Economic Approaches to Demand Management

Economic approaches to airport demand management utilize various forms of *congestion pricing* to influence the choices of aircraft operators and, in general, airport users concerning airport access. A system of access fees based on congestion pricing principles takes into consideration the pattern of delay at an airport over time and attempts to reduce delays to an economically efficient level. The access fees typically vary with time of day, as well as possibly by season and even by day of the week, with higher fees during peak demand periods and lower fees during off-peak periods. For this reason, congestion pricing is sometimes called *peak period pricing*.

Congestion pricing can serve as either an alternative or a complement to the purely administrative slot allocation procedures described in the previous section. When used as a complement (as is currently done at a few of the busiest airports in Europe), the pricing scheme is designed to discourage users from requesting slots during the most "popular" hours of the day by specifying higher charges for those hours. This kind of system is a "hybrid" and is discussed in Sec. 12.5. This section concentrates on purely economic demand management approaches.

Congestion Pricing in Theory

As noted in Sec. 12.1 and in Chap. 8, aircraft typically pay for access to airports through a landing fee proportional to the weight of the aircraft. The landing fee per unit of weight is determined through the average-cost pricing method, described in detail in Chap. 8.[13]

Two aspects of this virtually universal practice are problematic. First, as the amount of traffic at a congested airport increases, the landing fees will decrease, because the cost of the airfield is divided among more users (and more aircraft weight). Thus, with average-cost pricing, access to the airport becomes cheaper as congestion worsens! Second, there is a tenuous, at best, relationship between the landing fee that an aircraft pays and the true costs imposed by that aircraft's operation. A fee based solely on the landing weight of the aircraft essentially charges aircraft according to their "ability to pay," rather than in proportion to the true costs associated with operating at an airport. A 360-ton wide-body commercial jet will pay a landing fee 60 and 18 times greater, respectively, than a 6-ton general aviation aircraft and a 20-ton aircraft of a regional airline, using the same runway. Yet, as noted in Chap. 10, all three aircraft will occupy the runway and associated final approach path for roughly similar amounts of time—and occupancy time is what really counts in the case of congested facilities.

These fundamental inconsistencies between the price charged and the true cost of using congested airport facilities has long been pointed out by economists [see, e.g., Levine (1969), Carlin and Park (1970), Morrison (1987)] and is being increasingly recognized by airport and civil aviation experts and administrators. With growing congestion, the use of airport landing fees as a policy instrument for reducing delays and maximizing efficiency at these very expensive and scarce facilities may be just as important as the role that these fees currently play as generators of airport revenues. This is especially true at a time when landing fees, which had once been the principal source of revenue for airports, are becoming far less dominant in this respect[14] (see Chap. 8). The strongest economically and most congested airports can afford to experiment with alternative structures of the landing fees that draw on the principles of congestion pricing.

The fundamentals of congestion pricing, in general and in the context of airports, are well understood [see, e.g., Hotelling (1939), Vickrey (1969), Carlin and Park (1970)]. Only the main points are summarized here. Consider any facility that experiences congestion, all or part of the time. Every user who obtains access to the facility during periods when delays exist generates a congestion cost that consists of two components: (1) an "internal delay cost," or "private delay cost," that is, the cost that this particular user will incur due to the delay the user suffers; and (2) an "external delay cost," that is, the cost of the additional delay to all other prospective facility users which is caused by this particular user. For example, if airplane A, which uses a runway during a peak period, will delay 30 other aircraft queued at the runway by 2 minutes each—a very common occurrence at congested airports today—then the external cost generated by airplane A is the cost of the 60 minutes of delay to the other aircraft. (The 2 minutes correspond roughly to the "service time" of aircraft A, that is, to the time during which A occupies the runway/final approach, excluding all other aircraft from it—see Chap. 10.) At a cost of $80 per minute of delay to an airborne aircraft—a cost typical of airports with a significant fraction of wide-body airplanes in the traffic mix—this comes to $4800. This is an amount much greater than either the internal delay cost experienced by most individual aircraft using the runway system or the weight-based landing fee that a narrow-body aircraft, such as a Boeing 737 or Airbus 320, would pay at most of the world's busy commercial airports (see Chap. 8).

When an aircraft pays only the traditional, weight-based landing fee to operate at an airport (no matter how congested that airport might be), the only cost, in addition to the landing fee, that the aircraft's operator will perceive is the internal delay cost. Those airport users with the highest tolerance for internal delay costs, that is, those with a low cost of delay time, will be the ones who will persist the longest in using an airport as congestion and delays grow. By contrast, high-value-of-time operations, such as airline flights with large numbers of passengers, tight connections, and short turnaround times on the ground, are the ones that will be the most sensitive to worsening congestion.

The fundamental principle that the theory of congestion pricing applies in such cases is that, to achieve an economically efficient use of the facility, one must impose a congestion toll on each user equal to the external cost associated with that user's access to the facility. This is what economists refer to as forcing users to "internalize external costs." The underlying rationale is clear. Those who are willing to pay the congestion toll, that is, to compensate "society" for the external costs they impose, must be deriving an economic value from the use of the facility that exceeds these external costs. In other words, their use of the facility increases total economic welfare. Conversely, a user who is not willing to pay the congestion toll must be deriving a net economic benefit from the use of the facility that is less than the cost imposed on others. (Otherwise, the user would pay the toll.) Prospective users in this second category are thus denied access to the facility through the device of the congestion toll: such access would reduce total economic welfare.

The congestion toll then serves as a device for optimizing the use of the facility: absent any constraints, optimal use is achieved through a toll equal to the external cost associated with each additional ("marginal") user. Such a congestion toll not only contributes to a socially desirable result, but also is necessary to reach such a result. In the case of airport runways, the congestion toll is paid through the landing fee.

Figure 12.4 illustrates the situation. Curve *D* is the demand "curve" for the runway system of an airport facing capacity constraints. Curve *I* shows how the expected cost of delay suffered by each individual aircraft movement ("internal delay cost") increases as a function of demand.[15] Curve T shows the sum of the internal delay cost and the external delay cost generated by each additional aircraft movement at every level of demand. The difference between curves T and I at each level of demand is equal to the external delay cost generated by an additional (or "marginal") aircraft movement at that level of demand. Note the nonlinear increase of curves T and I with the number of movements, consistently with the principles outlined in Chap. 11.

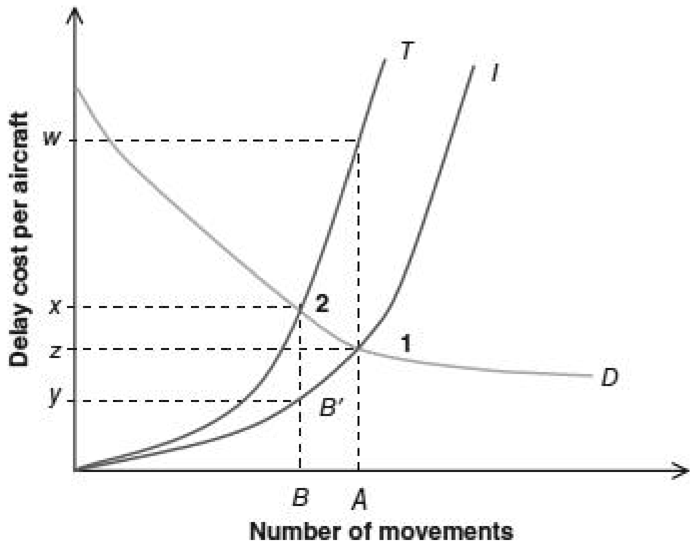


FIGURE 12.4 The effect of charging for external delay costs.

When there is no congestion toll, that is, when aircraft do not pay for any part of the delays they cause to other aircraft, the equilibrium point is at 1, where the demand curve D intersects the internal delay cost curve I. The equilibrium level of demand is equal to A, the projection of point 1 on the demand axis. When a congestion toll is imposed with a value equal to the external delay cost caused by the marginal runway system user, the new equilibrium is at 2, the point where the demand function D intersects the total cost curve T. Thus, the new equilibrium level of demand is equal to B. The demand is reduced by an amount equal to $(A - B)$ because of the congestion toll. In turn, the congestion toll is equal to the external cost corresponding to a demand equal to B, that is, to $(x - y)$ in Fig. 12.4. The total amount collected from congestion tolls is $B(x - y)$, the area corresponding to the rectangle $yx2B'$ in Fig. 12.4.

Figure 12.4 and this discussion explain the statement in Sec. 12.2 regarding the economic inefficiency of the "do-nothing" alternative. As mentioned then, the use of delay as the sole "natural" mechanism for access control results in an equilibrium (point 1 in Fig. 12.4) where (1) the delay level will be excessive and (2) the resulting mix of airport users may include a large fraction who have a high tolerance for delays because of their low value of time lost. Access to the airport by these low-value-of-time users cannot be economically justified, as they cannot compensate others for the external delay costs they cause.

An important side benefit of congestion pricing is that it provides information to decision makers about the need for investing in additional capacity and the value that users attach to such capacity. Equilibrium ("market-clearing") congestion tolls help establish a market price for airport capacity (Morrison, 1983; Oum and Zhang, 1990).

Congestion Pricing in Practice

The application of the theory of congestion pricing to airports is far from simple. At the technical level, it is not easy to estimate accurately the marginal external costs for any given level of demand, although considerable progress has been made in this direction in recent years. It is even more difficult to predict the exact effects of any proposed system of congestion tolls on demand, because existing information about the elasticity of airport demand to the landing fee is limited. Consequently, it is also difficult to determine the size of the landing fees that will lead to a stable situation ("equilibrium"), that is, will not drive away too many or too few users.

The principal practical problem, however, is more often political. The impact of congestion pricing is most severe on general aviation operators and on regional airlines. These two classes of users can least afford to compensate others for external costs and oppose congestion pricing as being discriminatory. Such opposition, when politically strong, as is the case in the United States, can arrest or slow down attempts at implementation. Small and remote communities, which depend on regional airlines for access to major airports and to the national and international aviation systems, typically join in opposition.[16]

The major airlines often find themselves in an ambivalent position in this respect. In principle, these carriers stand to benefit the most from congestion pricing. When the traffic mix includes flight operations by general aviation, regional airlines, and the major carriers, the application of congestion pricing is likely to reduce significantly delays to major carriers, by "driving away" from the peak traffic hours many general aviation and regional airline operations. As a result, major carrier operations would face reduced costs, even after paying the higher landing fee. Yet many airlines, especially in the United States, have to date assumed a stance on congestion pricing that ranges from guarded to adversarial. The reasons are several. For one, many major carriers have alliances with regional and commuter airlines (or even own such airlines) and are reluctant to support measures that are perceived as detrimental to them. Major carriers at busy airports also benefit from "feeder" traffic carried by smaller aircraft to/from smaller communities. Some of these flights might be canceled or scheduled at less convenient times if congestion pricing was applied. It is also probable that major carriers are, in general, uneasy about congestion pricing because they perceive it as a significant change to the *status quo*, with conceivably broader implications for the existing relationship between airports and airlines. The principal concern about this is the possibility that airports will abuse their inherent monopoly power or significant market power.

As a consequence of such practical and political considerations, the congestion pricing mechanisms that have been proposed or implemented to date are far less sophisticated than the theory suggests. They also generally impose or propose congestion tolls that are much lower than the true marginal external costs at congested airports. Typically, these congestion-pricing schemes involve one of the following approaches.

- A *surcharge* is applied to the weight-based landing fee during the airport's peak period(s). For example, all aircraft landing at an airport between 07:00 and 10:00 and between 16:00 and 19:00 local time might be required to pay a surcharge of $250, in addition to the weight-based landing fee to which they are subject.
- A *flat fee*, entirely or partly independent of the aircraft's weight, is imposed on all aircraft operating during the peak period. For example, all aircraft, no matter what their weight, may be required to pay a landing fee of $1000 if operating during peak hours. Or, aircraft under 50 tons may be required to pay $800 and aircraft over 50 tons $1500.
- A *multiplier* is applied to the weight-based landing fee charged to aircraft operating during the peak period. For example, with a multiplier of 1.25, an aircraft that would be subject to a $400 landing fee during off-peak hours would pay $500 if operating during the specified peak period.
- A *minimum landing fee* is specified for aircraft operating during peak hours, to be applied only to aircraft that would otherwise have paid less than that amount. For example, if the minimum landing fee is specified as $200, an aircraft that would have paid $100 during off-peak will be charged $200 if operating during the peak period; however, an aircraft that would have paid $250 during off-peak will still be charged the same $250 for a peak-period operation.

Each of these approaches has different impacts on different categories of airport users. For instance, a minimum landing fee affects only light aircraft that would have otherwise paid a smaller fee. By contrast, a multiplier increases the fee paid in direct proportion to an aircraft's weight, thus "penalizing," in absolute terms, heavier aircraft more than light aircraft for operating during peak periods. One can, in fact, question whether an approach based on multipliers bears any relationship to the principle of charging in proportion to external costs.

A number of additional practical problems must be resolved when developing congestion pricing schemes. A particularly thorny one concerns the use to which an airport puts the funds collected through the congestion tolls. Presumably these funds should be used locally to support projects, such as construction of new runways, aimed at relieving congestion by increasing airport capacity. However, at many congested airports around the world, there is little that can be done to increase capacity, other perhaps than waiting for improvements in the traffic handling efficiency of the ATM system (see Chap. 13). Airports such as New York/LaGuardia, Washington/Reagan, Boston/Logan, London/Heathrow, London/Gatwick, etc., already fully utilize their "real estate" and would face enormous political and other difficulties if they tried to expand substantially their airside facilities by acquiring adjacent parcels of land or by other means. In such cases, the collection of congestion tolls might be viewed by some as just a punitive measure or, worse, as a way for the airport operator to increase its revenues by taking advantage of the presence of congestion. It has been suggested that one possible approach to blunting this criticism is to reduce off-peak period charges so that the total revenue the airport derives from landing fees remains the same as before the imposition of a congestion toll.

Example 12.2 describes a demand management system that illustrates many of the points discussed in this section, as well as additional issues that may come up in practice. Example 12.3 illustrates the use of congestion pricing in passenger terminal buildings and on aprons.

Example 12.2 Massport has examined a broad range of congestion pricing options for Boston/Logan. Peak-period pricing was one of the principal alternatives explored in its 1999 Draft Environmental Impact Statement, which proposed a number of airport improvements. Figure 12.5 summarizes the structure of one of the pricing schemes examined in detail at Massport (Barrett et al., 1994). Its main features are outlined as follows, along with the underlying rationale.

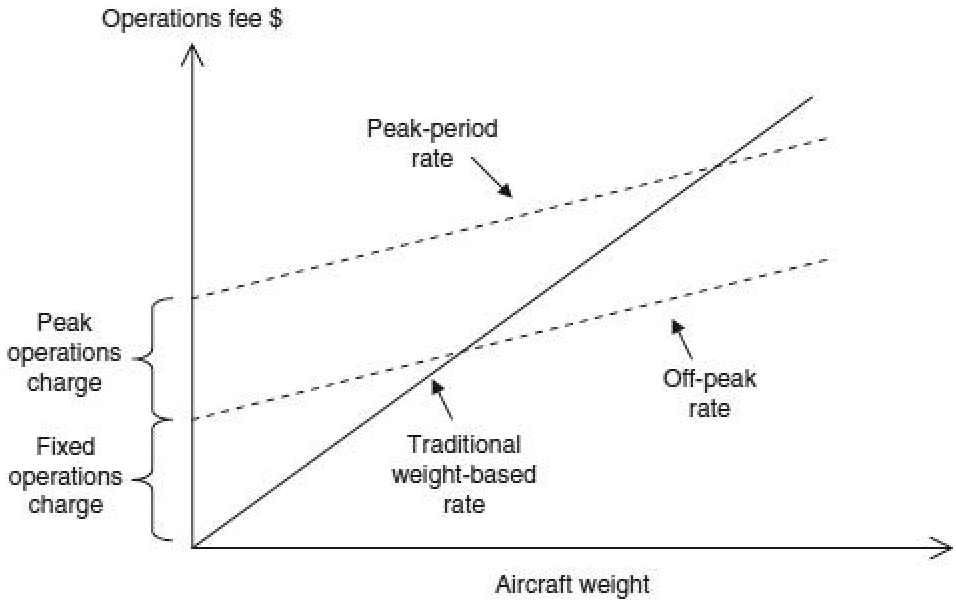


FIGURE 12.5 New landing fee structure versus traditional weight-based fee.

The landing fee would consist of three elements: a fixed charge per flight, a variable charge proportional to the weight of the aircraft, and a surcharge that would apply only to aircraft movements (landings and takeoffs) taking place during peak periods. The peak-period surcharge would be modest—of the order of $100 to $200. This range was selected so that the surcharge would have some noticeable impact on demand, while not making access to Boston/Logan prohibitively expensive for smaller aircraft. The actual marginal external delay costs per aircraft movement at Boston/Logan in the 1990s were much higher than $200 during peak traffic periods.

A peak period was defined as any period of 3 or more consecutive hours when the moving 3-hour average of expected demand exceeded the range of 100 to 110 movements per hour throughout the period. This threshold was selected because it is close to the average of the maximum throughput capacity of Logan Airport (about 115) when this average is computed over the entire year.[17] There are two reasons for requiring peak periods to last for at least 3 consecutive hours. First, serious delays occur only if high demand levels persist for an extended period of time. When a brief period of high demand is followed by a period of low activity, the airport has a chance to recover before delays become disruptive. Second, if peak periods are too short, some users may be able to avoid paying the surcharge by making only small changes to their schedules to "shift out" of the peak period. For example, if a peak period happened to be only 1 hour long, about half of the affected aircraft operations could avoid the surcharge by shifting their scheduled times, forward or backward, by at most 15 minutes.

The proposed landing fees were computed so that the pricing structure would be "revenue neutral"; that is, total revenue from runway fees would be the same under the new, congestion-pricing system as under the existing system that determined landing fees solely on the basis of aircraft weight. This would avoid any appearance that the proposed system might be a pretext for increasing airport revenues. The net result (see Fig. 12.5) would be that smaller aircraft would end up paying a significantly higher landing fee than under the existing system, especially if they operated during peak traffic periods. However, some of the largest aircraft would generally be charged a *lower* landing fee than under the existing system, even during peak periods when the surcharge were in effect.

A potential variation would exempt from the peak-period surcharge two flights per day to or from the 12 smallest markets served by Boston/Logan. This provision was designed to cushion small local communities from the impact of the proposed system and partly defuse political pressure and criticism.

Example 12.3 The British Airports Authority pioneered the use of marginal cost and congestion pricing at airports (Little and McLeod, 1972). It began using congestion-pricing schemes at the passenger terminals and the aprons of London/Heathrow in the mid-1970s. It later extended this practice to London/Gatwick and London/Stansted. For instance, Table 12.4 shows the 1998 schedule of passenger service charges (see Chap. 8) at the three airports. A Boeing 747 with 400 passengers on board would pay the substantial *additional* amount of £1240 (~ $2000 in 1998) in passenger service charges for a departure scheduled during a peak period at London/Heathrow. Note that the schedule did not differentiate between peak and off-peak *domestic* passengers, as the domestic terminal areas were not congested.

	Heathrow		Gatwick		Stansted	
Flight Time	**Peak***	**Off-peak**	**Peak***	**Off-peak**	**Peak†**	**Off-peak**
International‡	£8.50	£5.40	£6.95	£4.00	£6.75	£2.90
Domestic	£3.40	£3.40	£3.30	£3.30	£3.20	£3.20

*Aircraft departing between 09:00 and 15:29 GMT, April 1–October 31.

†Aircraft departing between 06:00 and 15:59 GMT, April 1–October 31.

‡Passengers from Gatwick and Stansted to Ireland pay a slightly different charge from the ones shown.

TABLE 12.4 Passenger Terminal Service Charge Payable per Departing Passenger at Heathrow, Gatwick, and Stansted Airports, 1998; a £1.50 Rebate per Passenger Is Provided in All Cases for Flights Departing from Remote Stands

On aprons, the standard charge in 1998 for parking an aircraft at London/Heathrow was based on the duration of the stand's occupancy and on the MTOW of the aircraft. It was £3.30 per quarter-hour or part thereof plus 5.4 p per ton. However, each minute's occupancy of a contact stand counted as 3 minutes for the period between 07:00 and 12:29 GMT. A similar arrangement was in effect at London/Gatwick. As of 2012, the practice of tripling charges for aircraft parking during peak periods was continuing at both airports.

Finally, two practical observations are particularly relevant to congestion pricing at airports. The first is that it will be most effective when applied to airports with non-homogeneous traffic. This is the case, for example, when, in addition to major carrier traffic, a significant fraction of the traffic mix consists of general aviation and/or regional airline flights, or when there is a reasonable mix of short-range flights by small aircraft and long-range flights by large ones. If such conditions do not exist, the ability of any landing fee to achieve price differentiation among users will be limited and a demand management system based solely on congestion pricing is likely to prove ineffective. Hybrid demand management systems, such as those to be discussed in the next section, would probably work better in such cases.

The second observation pertains to hub airports, where traffic may be dominated by a single airline. (An extreme example may be Memphis at nighttime, when essentially all the traffic consists of FedEx cargo aircraft.) In such cases, aircraft belonging to the dominant airline will absorb nearly all the external delay costs generated by any aircraft movement.

Thus, the dominant airline will internalize external delay costs anyway and will try to operate the number of flights that maximizes its total economic welfare.

In summary, congestion pricing is more appropriate at congested airports with a large number of competing carriers and with no dominant operator or operators. New York/LaGuardia (see Example 12.1) is an excellent case in point.

12.5 Hybrid Approaches to Demand Management

Hybrid demand management systems combine administrative and economic mechanisms. The starting point for all hybrid systems is the determination by some administrative authority of the number of slots to be made available at an airport. However, instead of (or, in addition to) schedule coordination, hybrid systems rely on economic devices such as congestion pricing, slot auctions, and/or slot markets to allocate slots among potential airport users. This section reviews these three approaches and concludes with suggestions for additional combinations of administrative and economic approaches.

Slots Plus Congestion Pricing

The *slots plus congestion pricing* approach may be widely adopted in the future as it offers a simple way for providing prospective airport users economic incentives that take into consideration not only congestion but also the environmental impacts of airport use. Its application involves four steps:

1. Specify the number of slots available in each time period of the day, based on the airport's declared capacity and, possibly, noise and other considerations.
2. Develop and announce a schedule of landing fees (and possibly other airport charges) that may vary by time of day and/or day-of-week and/or season.
3. Invite requests for slots from prospective users.
4. Use a schedule coordinator (or other administrative mechanism) to allocate slots, whenever the number of requests for a time period exceeds the number of available slots.

The main difference between this and the purely administrative schedule coordination approach is that prospective airport users must consider the cost of access to the airport at different times when preparing their requests for slots. The higher cost of access during peak periods may dissuade some prospective users from requesting slots for these times. A few European airports have already implemented variants of this approach. The fees charged for access during different parts of the day are typically also influenced by airport noise considerations. Examples 12.4 and 12.5 illustrate the possibilities.

Slot Auctions

Another hybrid approach to airport demand management would use auctions to allocate slots. Many economists have advocated this approach over the years and an extensive literature exists on the subject [see, e.g., Czerny et al. (2008) for several articles and numerous references]. The general idea is simple and attractive: After determining the number of slots available in each time period of the day, invite bids for all or some specified percent of these slots and award them to the highest bidders. The details of putting this approach to practice have, however, proved formidable. Although several airport operators or national governments have considered the adoption of slot auctions at airports such as Sydney, New York, and a number of European locations, there has been no actual implementation to date. The closest to a truly large-scale application was a detailed plan submitted by the FAA in 2008 that would gradually introduce auctions and slot trading as supplementary means for allocating slots at the three New York airports. An initial step would be the redistribution through an auction of about 10 percent of all the slots at the three airports: 89 at Kennedy, 113 at LaGuardia, and 81 at Newark. However, the airlines, through the Air Transport Association, and the local airport operator, the Port Authority of New York and New Jersey, strongly opposed the plan, each for different reasons, and mounted a legal challenge to it. After a setback in the courts, the FAA withdrew the plan in 2009. Given the widespread interest in slot auctions, it is likely that other proposals of this nature will appear in the future.

It is important to discuss briefly at this point some of the complexities of auctions in the airport context. First, there is the fundamental question of who should be the recipient of the proceeds from the initial auctioning of (some or all of) the slots at an airport. Grandfathered airlines, airport operators, and ANSPs have all made plausible claims to ownership rights and demanded a share of auction revenues. The monetary stakes can be very large as indicated later in this section.

Example 12.4 In 2010, Manchester Airport in the United Kingdom subdivided the day into six subperiods, three with peak charges and three with off-peak. Peak charges were imposed on movements during the periods 07:00–12:59, 16:00–19:59, and 23:00–5:29, a total of 15.5 hours of the day. The first two periods are associated with peak traffic demand, whereas the third is “protected” for environmental reasons. A (roughly) 300-ton Boeing 777 would pay a “runway and air traffic services charge” of £1423 if it departed during a peak period and only £637 in off-peak.[18]

Example 12.5 The 2005 landing fee schedule for the three BAA London airports is shown in Table 12.5. As the airports are fully coordinated, the combination of this schedule with the slot allocation procedure essentially amounted to a hybrid slots-plus-congestion-pricing system.

Aircraft Weight (tons)	Heathrow		Gatwick		Stansted	
	Peak	Off-peak	Peak	Off-peak	Peak	Off-peak
MTOW ≤ 16	£590	£250	£385	£110	£95	£85
16 < MTOW ≤ 50	£590	£250	£385	£110	£142	£105
50 < MTOW ≤ 250	£590	£250	£385	£125	£231	£131
250 < MTOW	£590	£250	£385	£125	£400	£400

TABLE 12.5 Landing Fee Schedules at Heathrow, Gatwick, and Stansted Airports, 2005

Note that the landing fees at London/Heathrow and London/Gatwick were essentially independent of aircraft weight during both peak and off-peak periods, but the landing fees in the peak period were more than twice as large as in off-peak. At the less congested London/Stansted, the landing fee varied with aircraft weight with a somewhat higher fee during peak periods. The definition of "peak period" changed across airports (e.g., it was 07:00–09:59 and 17:00–18:59 GMT, April 1–Oct. 31 for Heathrow) and, in fact, it was different for aircraft traffic and for passenger traffic (cf. Table 12.4).

In 2010, following BAA's sale of London/Gatwick to Global Infrastructure Partners (GIP), the landing fee structure at London/Gatwick and London/Stansted stayed roughly as in Table 12.5. At London/Gatwick, the landing fee for all aircraft with MTOW greater than 16 tons became £979 (roughly $1500) at peak periods and £321 at off-peak,[19] a 3:1 ratio, whereas at London/Stansted it continued to vary with aircraft weight and with peak and off-peak use. However, the landing fee structure at London/Heathrow was further simplified and a single charge of £776 at *all times other than late nighttime* was imposed on all aircraft with MTOW greater than 16 tons in the "Base Chapter 3" noise category. Note the sharp contrast between this policy ("all aircraft movements are charged an equal amount") and the traditional approach of setting landing fees in direct proportion to aircraft weight. The fee at London/Heathrow varied only with the noise rating of the aircraft, while a very large surcharge was imposed for late night (00:00–03:29) operations.

Second, it is necessary to specify precisely the rights and obligations associated with a slot, such as the length of time for which a slot is "valid" or the conditions under which a national government can withdraw or cancel a slot. Many divergent views have been expressed on such questions and the amounts that auction participants will bid strongly depend on the answers.

Third, the combinatorial complexity of slot auctions (Ball et al., 2006) gives rise to a slew of technical issues that require extensive analysis with many questions still unresolved. The complexity stems from the strong interdependence of slots, both at the local level and across airports. Consider, for example, an airline that wishes to obtain an arrival slot for some particular flight at Airport A. Assume that the preferred time for the arrival of that flight is between 09:00 and 09:30 local time, but a slot between 08:30 and 09:00 may be acceptable as well. The airline may then decide to bid a certain amount for a slot between 09:00 and 09:30 and a smaller amount for a slot between 08:30 and 09:00, for "insurance." Note now that, if the airline wins both slots, the 08:30–09:00 slot is essentially of no value to that airline. At the same time, if the turnaround time of the aircraft involved is about 60 minutes, the airline must also make sure to obtain a departure slot between 10:00 and 10:30, so that the arriving aircraft can operate efficiently. Otherwise, the value of the 09:00–09:30 slot will be greatly diminished. Moreover, once the airline obtains a 09:00–09:30 arrival slot at Airport A, it must obtain a corresponding earlier departure slot at Airport B from which the flight will originate.

Because of these strong interdependencies, the true value of the slots acquired will not be clear to an airline until all the slots are allocated. At that point, the airline will probably wish to dispose of some of the slots it has been awarded, revise the price it has offered to pay for others, and possibly acquire some additional slots. To make such postauction adjustments possible, a follow-up slot market is needed. This follow-up market is, in fact, an indispensable part of any demand management system based on auctions. Note, as well, that a major airline, which will bid for many slots at a particular airport, probably enjoys an inherent advantage over a smaller airline, because the major airline can choose among many possible combinations of slot usage once it obtains its slots.

In summary, an airline will face a most difficult task in preparing a bid for a set of slots at a busy airport—and so will the auction administrator who will have to select from the set of submitted bids, the optimal subset of bids that will be accepted. A postauction secondary slot market (see below) must also be available, in any event.

Secondary Trading: Buying and Selling Slots

Secondary trading is another hybrid approach to demand management. It can be used in combination with an administrative allocation of slots, such as IATA's schedule coordination process, or as a follow-up to a slot auction as described previously. Once the slots have somehow been allocated, their holders may treat them as assets—subject to any applicable limitations in each case—and may sell them or lease them for an agreed period.[20] An essential part of secondary trading and auctions is a clear articulation of the rights and conditions of use associated with a slot. An important side-benefit of both secondary trading and slot auctions is that they provide market mechanisms for determining directly the value of slots.

Quite surprisingly, the rules that govern secondary slot trading are still in a state of flux throughout the world, with the exception of the United States where they have been reasonably well-established since the late 1980s. The IATA schedule coordination process does not include any guidelines for the trading or exchange of slots with monetary compensation.[21] Secondary trading has therefore been taking place internationally in a "gray zone," with the most active market being apparently the one for slots at London/Heathrow (de Wit and Burghouwt, 2007). In 1999, a landmark U.K. court decision ruled that slot trading is not illegal, thus allowing open trading of slots in the United Kingdom, on an essentially one-by-one basis. The European Commission arrived at a similar viewpoint in 2008, leading to its 2011 proposal (see Sec. 12.3) to establish a framework within which airlines may engage in secondary trading.

In contrast to the situation elsewhere, a specific set of rules for secondary trading ("buy-and-sell") were established in the United States in connection with slots at the HDR airports. As noted in Sec. 12.3, the U.S. DOT abolished schedule coordination at these airports in 1985 and authorized the airlines that held slots as of that date to continue utilizing them under the following set of rules:[22]

- Slots are subdivided into three categories: air carrier, commuter/regional, and other (general aviation, military, and charter). Air carrier and commuter slots are further subdivided into domestic, Essential Air Service[23] (EAS), and international.
- Slots are authorized by the DOT and are not technically the property of their current holders. However, domestic slots can be bought, sold, leased, or used as collateral. Air carrier domestic slots can be used by aircraft of any size, but regional domestic slots can only be used by non-jet aircraft with a maximum seating capacity of less than 75 or turbojet aircraft with seating capacity less than 56.
- Any slots that are not utilized at least 80 percent of the time during any 2-month period are taken away from their current holders and transferred to the pool of "available" domestic slots.
- The U.S. DOT retains the right to withdraw some domestic slots for designation as EAS or international slots. It assigns each of these slots a *priority withdrawal number*. This number defines the order in which the Department can claim back slots, and thus indicates to a holder or to a potential buyer the likelihood that a slot may be recalled. However, slots held by users who possess eight or fewer slots at an airport are exempt from slot withdrawals at that airport.
- New slots, that is, those made available either through an increase in airport capacity or because of underutilization of an existing slot, are distributed periodically, with the first 15 percent of the slots reserved exclusively for new entrants.

The HDR environment has given rise to an active market for slots, initially at the four HDR airports and currently at New York/LaGuardia and Washington/Reagan.

Gillen (2006) estimated that the average price of a slot at London/Heathrow, before 2006, was of the order of £4 to 6 million (~ $10 million), while the value of a slot at Washington/Reagan was about $1 million and at New York/LaGuardia about $500,000. It is clear, however, that slot values change rapidly over time. For example, Continental Airlines purchased four pairs of slots at London/Heathrow in March 2008 for a price widely reported as $209 million—about $26 million per slot!

12.6 Policy Considerations

It is now possible to list the most desirable attributes of an airport demand management system. Such a system would

- Promote the economically efficient use of scarce airport capacity by discouraging access by aircraft operators who will impose more costs on others than the benefits they will derive for themselves
- Maintain access to the congested airport for all users willing to pay the full economic cost
- Be perceived as fair and non-discriminatory by all classes of prospective airport users
- Not serve as another way to regulate air transport
- Not provide opportunities for collusion among airlines and anticompetitive practices
- Not result in any “windfall” revenues to airport or aircraft operators; any additional revenues accruing through demand management would be used to increase airport capacity and relieve congestion
- Be transparent in its methodology and easy to administer and modify, if necessary

None of the approaches described in this chapter fully satisfies all these criteria. However, given any specific set of circumstances, certain approaches will be superior to others. For policy-setting purposes, a number of important points to bear in mind are summarized as follows.

1. Traditional weight-based landing fees do not take into consideration the costs associated with airport congestion. If anything, they contribute to congestion, by lowering the cost of airport access as demand grows and by encouraging users with low direct operating costs and low value of time to use busy airports.
2. Schedule coordination relying solely on administrative procedures can be effective at mildly congested airports and in environments where change is slow and gradual due to heavy regulation. However, in a dynamic, deregulated environment and at airports facing severe congestion, purely administrative procedures almost unavoidably inhibit competition and cause significant market distortion in the long run.
3. All purely economic approaches to demand management involve some form of congestion pricing. That is, they are based on the principle that, to optimize use of a congested facility, users should be forced to internalize, preferably fully, the

external costs they cause. Congestion pricing can be particularly effective at airports where traffic is nonhomogeneous and not dominated by one or two carriers. Although the theoretical principles of congestion pricing are well understood, it is difficult to apply them in practice, both for technical and for political reasons. The congestion-related landing fee schedules that have been implemented to date impose relatively low tolls on peak-period operations and are greatly simplified in structure.

4. Hybrid demand management systems combine elements of administrative and economic approaches. Their common characteristic is their use of administrative procedures to specify the number of slots available at an airport and their reliance on economic devices such as congestion pricing, slot markets, and slot auctions to arrive at the final allocation of slots. Several major airport authorities already use a combination of schedule coordination and (limited) congestion pricing to manage demand. Secondary slot trading ("buy-and-sell") is a well-established approach in the United States and the United Kingdom and is gaining ground elsewhere. However, it requires resolution of the difficult issue of who the original owner or "provider" of airport slots is. Auctioning of airport slots is an idea that has generated great international interest, but is largely unexplored in practice. Its application entails many practical complications. If slot auctions are implemented, they clearly should be supplemented by a secondary trading market.

More generally, the implementation of any demand management system requires attention to the details of the proposed approach, a large amount of analysis, consultation with all parties involved, and resolution of numerous difficult questions. Examples of additional issues that were only implicitly raised in this chapter include the number of slots to be offered for distribution, the duration of peak and off-peak periods, the use of scheduled or actual time of operation for application of the charge, possible exemptions for certain airport users, rights and obligations of slot holders, and duration of slot ownership.

Interest in airport demand management is high throughout the world. In general, access to many of the busiest and most congested airports seems to be seriously underpriced at present. The use of demand management measures will probably expand in the future, with hybrid systems likely to be widely adopted. Movement in this direction is gradual, however, as the strengthening of the economic components of hybrid systems encounters serious resistance from several segments of the air transportation community.

Exercises

12.1. Example 12.1 indicates that the limited number of AIR-21 slots at New York/LaGuardia were allocated among eligible users through a lottery whose results took effect at the end of January 2001. The use of a lottery was selected because of the pressure to deal quickly with delays that had reached crisis levels. It was acknowledged to be only a temporary solution to the problem. What is wrong with using a lottery to allocate scarce airport capacity? What are the advantages, if any, of using this approach?

12.2. As Sec. 12.5 noted, in May 2008 the FAA proposed a demand management system at New York/Kennedy and New York/Newark that would include the auctioning of some slots at these two airports. The full proposal can be found at: http://www.gpo.gov/fdsys/pkg/FR-2008-05-21/html/08-1271.htm. Review this complex scheme and summarize its main elements in simple terms. Discuss the strengths and weaknesses of the proposal, which was strongly opposed by the airlines. Several relevant comments (pro and con) have been posted.

12.3. Consider an airport experiencing serious congestion during peak traffic hours. Under peak conditions, the arrivals of airplanes at the vicinity of the airport can be assumed to be approximately Poisson with a rate of 55 aircraft per hour. Of these airplanes, 40 on average are commercial jets and 15 are small general aviation and commuter airplanes. The probability density function for the duration of the service time, t, to a random landing aircraft is uniformly distributed between 48 and 72 seconds. Peak traffic conditions occur during 1000 hours per year, and the average cost of 1 minute of airborne waiting time (i.e., of time spent in the air while waiting to land) is \$60 for commercial jets. (This accounts for additional fuel burn, extra flight crew time, and other variable operating costs.) Estimate the yearly costs to the airlines of peak traffic conditions. Assume the model described by Eq. (20.10) of Chap. 20 for estimating waiting time is valid for this case.

12.4. Continuation of Exercise 12.3. To alleviate congestion under peak traffic conditions, the airport's managers are considering an increase in the landing fees. They have concluded that demand by commercial jets is completely insensitive to moderate increases in the landing fee (i.e., demand will continue at the level of 40 per hour). However, demand by smaller aircraft is expected to drop drastically as the landing fee increases, as several small airports nearby are good alternatives to the main airport. A study has indicated that the relationship between demand by small aircraft and the increase in the landing fee is given by the relationship

$$Y = 15 - (X/32) \quad \text{for} \quad 0 \leq X \leq 480$$

where X is the amount (in \$) added to the landing fee and Y is the number of small aircraft per hour demanding access to the airport. (Note that when $X = \$0$, $Y = 15$, and when $X = \$480$, $Y = 0$.) What is the most desirable amount of increase in the landing fee from the point of view of the airlines? (Remember that the airlines will also be paying the higher fees.) [*Note:* The variance of the service times in this problem is equal to $(72 - 48)^2/12 = 48$ seconds^2.]

12.5. Repeat Exercise 12.4 for the case in which the cost per minute of waiting for commercial jets is \$100. (This would mean a virtually all wide-body aircraft mix on the airlines' side.) Compare the results of Exercises 12.4 and 12.5. Did your answer change in the right direction? Is the difference as large as you expected? Why?

References

Ball, M., Donohue, G., and Hoffman, K. (2006) "Auctions for the Safe, Efficient and Equitable Allocation of Airspace System Resources," in P. Cramton, Y. Shoham, and R. Steinberg (eds.), *Combinatorial Auctions*, MIT Press, Cambridge, MA.

Barrett, C., Drazen, M., Hoffman, W., Lewis, S., Murphy, R., Odoni, A., and Pearson, L. (1994) "Peak Period Airport Pricing as It Might Apply to Boston-Logan International Airport," *Transportation Research Record, 1461*, pp. 15–23.

Carlin, A. and Park, R. (1970) "Marginal Cost Pricing of Airport Runway Capacity," *American Economic Review, 60*, pp. 310–318.

Czerny, A. I., Forsyth, P., Gillen, D., and Niemeier, H-M. (eds.) (2008) *Airport Slots: International Experiences and Options for Reform*, Ashgate Publishing, Hampshire, U.K.

de Wit, J. and Burghouwt G. (2007) *The Impact of Secondary Slot Trading at Amsterdam Airport Schiphol*, Report for The Netherlands Ministry of Transport, SEO Economic Research, Amsterdam, The Netherlands.

European Commission (2011) *Proposal for a Regulation of the European Parliament and of the Council on Common Rules for the Allocation of Slots at European Union Airports*, COM (2011) 827 final, 2011/0391 (COD), Brussels, December 1.

European Commission (2004) *Regulation (EC) No 793/2004 of the European Parliament and of the Council of 21 April 2004 Amending Council Regulation (EEC) No 95/93 on Common Rules for the Allocation of Slots at Community Airports*, Brussels, April 21.

FAA, Federal Aviation Administration (2001) "Notice of Alternative Policy Options for Managing Capacity at LaGuardia Airport and Proposed Extension of the Lottery Allocation," *Federal Register, 66* (113), June 12, pp. 31731–31748.

Fan, T. and Odoni, A. (2001) "Demand Management as a Means of Relieving Airport Congestion," Working Paper, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA.

Gillen, D. (2006) *Slot Trading in North America*, presentation prepared for the EUACA Seminar on Secondary Trading, 28 June 2006, Amsterdam, http://www.fhkd.org/cms/uploads/media/David_Gillen_01.pdf.

Hotelling, H. (1939) "The Relation of Prices to Marginal Costs in an Optimum System," *Econometrica, 7*, pp. 151–155.

IATA, International Air Transport Association (2000) *Worldwide Scheduling Guidelines*, Montreal, Canada.

IATA, International Air Transport Association (2012) *Worldwide Slot Guidelines*, Montreal, Canada.

Jacquillat, A. and Odoni, A. (2012) "A Study of Airport Congestion at JFK and EWR," *Proceedings 5th International Conference on Research in Air Transportation—ICRAT*, Berkeley, CA.

Kilian, M. (2008) "The Development of the Regulatory Regime of Slot Allocation in the EU," chapter in *Airport Slots: International Experiences and Options for Reform*, A. I. Czerny, P. Forsyth, D. Gillen, and H-M. Niemier (eds.), Ashgate Publishing, Hampshire, UK.

Levine, M. (1969) "Landing Fees and the Airport Congestion Problem," *Journal of Law and Economics, 12*, pp. 79–108.

Little, I. M. D. and McLeod, K. (1972) "New Pricing Policy for British Airports," *Journal of Transport Economics and Policy, 6*(2), May, pp. 101–115.

Morisset, T. (2010) "Comparing Capacities and Delays at Major European and American Airports," SM Thesis, Technology and Policy Program, Massachusetts Institute of Technology, Cambridge, MA.

Morrison, S. (1983) "Estimation of Long-Run Prices and Investment Levels for Airport Runways," *Research in Transportation Economics, 1*, pp. 103–130.

Morrison, S. (1987) "The Equity and Efficiency of Runway Pricing," *Journal of Public Economics, 34*, pp. 45–60.

OIG, Office of Inspector General, U.S. Department of Transportation (2010) *New York Flight Delays Have Three Main Causes, but More Work Is Needed to Understand Their Nationwide Effect*, Tech. Rep. AV-2011-007, October.

Oum, T. and Zhang, Y. (1990) "Airport Pricing: Congestion Tolls, Lumpy Investment and Cost Recovery," *Journal of Public Economics, 43*, December, pp. 353–374.

U.S. Department of Transportation (DOT) (1995) "Report to Congress: A Study of the High Density Rule," DOT, Washington, DC.

Vickrey, W. (1969) "Congestion Theory and Transport Investment," *American Economic Review, 59*, pp. 251–260.

[1]The maximum takeoff weight (MTOW) of the aircraft is typically used as the basis for computing the landing fee. Other types of user charges (passenger service fees, aircraft parking fees, hangar fees, etc.) may also be imposed, depending on the type of flight involved. Chapter 8 discusses airport user charges in detail.

[2]International flights must be authorized in every case under the applicable bilateral or multilateral agreements.

[3]This is only partly true at airports where traffic is dominated by a single airline—see Sec. 12.3.

[4]This frequently updated document was formerly known as the *Worldwide Scheduling Guidelines*.

[5]The November SCC is concerned with schedule coordination for the upcoming summer season (in the Northern Hemisphere) and the June SCC with the upcoming winter season.

[6]This is the so-called "use-it-or-lose-it" rule.

[7]In the 2012 version the objective of the WSG is restated as "ensuring the most efficient use of airport infrastructure in order to maximize benefits to the greatest number of airport users."

[8]Note that at a busy airport with, for instance, 1000 movements per day, this means 50 movements, a large increase over the IATA limit of 4.

[9]Misbehavior is classified as "frequent," "significant," or "intentional." An example of intentional misbehavior is the consistent operation of a flight at a time other than the allocated slot.

[10]New York/Newark was exempted from slot limits in October 1970.

[11]One of these changes was a requirement for a minimum 70 percent utilization of slots by current slot holders.

[12]Delays at New York/Kennedy and New York/Newark declined gradually by nearly 40 to 50 percent between 2007 and 2010. This, however, had little to do with the scheduling limits and can be largely explained by the reduction in the number of flights at the two airports during the years in question (Jacquillat and Odoni, 2012).

[13]Briefly, average-cost pricing (Chap. 8) consists of three basic steps when it comes to determining landing fees: (1) a target amount of revenue, X, to be collected from the fees, is specified at the beginning of the airport's fiscal year (typically, X is equal to the annual cost of the airfield, including a reasonable return on the airport's investment); (2) a forecast is made of the total number of units of weight, Y, of all the aircraft that will utilize the airport during that year; and (3) the landing fee per unit of weight, Z, is set equal to the ratio X/Y.

[14]For example, revenues from automobile parking now exceed revenues from landing fees at many major airports in the United States!

[15]Note that curve *I* for the cost of delay has the shape of the nonlinearly increasing expected delay functions seen in Chap. 11.

[16]On the opposite side, airport congestion pricing is generally supported politically by environmentalists, as well as by neighbors of major airports, who see in this approach a means of postponing airport expansion through access control.

[17]The average of the maximum throughput capacity is computed from the capacity coverage curve (see Chap. 10) for Boston/Logan.

[18]Depending on their noise rating, certain types of aircraft would also pay a penalty during some off-peak times (e.g., between 05:30 and 05:59).

[19]The landing fees cited are for aircraft in the "Base Chapter 3" noise category (see Chap. 6); they vary considerably with the noise rating of the aircraft.

[20]Under the HDR in the United States, slots were even used as collateral for securing loans.

[21]However, recent versions of the IATA WSG recognize the possibility of such trading by stating that "slot exchanges for compensation or consideration may only take place where they are not prohibited by the laws of the relevant country."

[22]The rules have undergone several changes since 1985.

[23]EAS refers to government-subsidized flights serving small remote communities.

CHAPTER 13

Air Traffic Management

This chapter provides an introduction to air traffic management (ATM), with emphasis on terminal airspace and airport operations. ATM in developed countries has evolved into a complex, large-scale system that depends heavily on the smooth interaction of a highly skilled labor force with increasingly advanced technologies, and on close collaboration with a diverse community of users. ATM planning, investments, and operations must constantly make tradeoffs among objectives involving safety, efficiency, and cost.

ATM systems can be classified into several "generations" depending on their technological characteristics and level of automation. As of 2012, most developed countries operated systems that begin to take advantage of satellite-based technologies, collaborative decision-making (CDM), advanced automation, and decision-support tools. Some less developed countries, however, still operate in parts of their territory with quite primitive systems, little different, in some ways, from the earliest air traffic control systems. This chapter briefly describes the aspects of more advanced systems that are relevant to terminal airspace and airport operations. These include the classes of airspace, the operation of airport traffic control towers and of terminal airspace control centers, principal types of surveillance equipment, and instrument landing systems.

Air traffic flow management (ATFM) increasingly plays an important role in airport operations. It is a means for avoiding traffic overloads and excessive congestion and delay costs. ATFM systems in Europe and the United States have evolved greatly in complexity and sophistication since their beginnings in the mid1980s. In particular, the adoption by the U.S. Federal Aviation Administration (FAA) of a CDM approach to ATFM is one of the most significant events in the history of ATM. It marked a major change from the traditional philosophy for operating air traffic control. The chapter describes some early steps in the application of CDM to ATFM.

The chapter concludes with a summary of expected near-term and medium-term innovations in surveillance, navigation, and automation as they apply to terminal airspace and airports. It is difficult to forecast the impact these innovations will have at major airports in the long run. The most significant benefits will be in the areas of safety (fewer runway and taxiway incursions) and of environmental impacts, especially noxious emissions. But delays in implementing some of the planned enhancements, as well as the slow pace of adoption of the required avionics by aircraft operators, suggest that only modest increases in runway capacity can be anticipated by 2025.

13.1 Introduction

ATM is essential to the operation of airports. Any airport planner or manager must therefore be familiar with at least its most fundamental aspects. This chapter provides the requisite basic background, with special emphasis on ATM in terminal area airspace. A few references that focus solely on ATM in general (Nolan, 2011; Cook, 2007) or ATM in terminal airspace (Mundra, 1989) or ongoing innovation (EUROCONTROL, 2010; FAA, 2012) offer more detailed treatment.

The ATM system provides a set of services aimed at ensuring the safety and efficiency of air traffic flows. Advanced ATM systems are becoming increasingly complex. They must:

- Accommodate growing numbers of users with different capabilities and requirements
- Achieve exceptional levels of safety under close scrutiny from the public and the mass media
- Mesh seamlessly a large labor pool of skilled human operators (the air traffic controllers and other technical staff) with a network of computers and other sophisticated communications, surveillance, and navigation equipment
- Take advantage of technological developments, while evolving gradually to allow users to keep pace with the rate of change
- Accomplish all this at reasonable cost to service providers and users

Viewed from a long-term perspective, ATM systems have been reasonably successful on all these scores. Undoubtedly, their greatest achievement is the extraordinary level of safety attained by commercial air travel in developed countries throughout the world and the resultant benefits to more than 2.5 billion passengers per year (ACI, 2011; Barnett, 2009).

Any ATM system is comprised of six components (Braff et al., 1994). In fact, ATM systems can best be compared with one another—and their evolution over time can be traced—by referring to the state and quality of each of these critical components:

- *Procedures and regulations* according to which the ATM system operates and the organization of airspace around airports and en route
- *Human air traffic controllers*, who are responsible for providing ATM services

- *Automation systems* (e.g., computers, displays, and special-purpose software) that provide information to the controllers on the status, location, and separation of aircraft in the system, and assist them ("decision support") in processing safely and expediting the flow of traffic
- *Communications systems* that enable air–ground, ground–ground, and air–air voice communications and data exchange and sharing
- *Surveillance systems* (e.g., radar) that provide real-time positional information to air traffic controllers—and, possibly, to the cockpit—for tracking aircraft and hazardous weather
- *Navigation systems* that provide real-time information to individual aircraft on their own position and assist them in navigating through airspace and on the airport surface

Section 13.2 summarizes the evolution of ATM systems and describes aspects of more advanced systems that are particularly relevant to ATM in terminal airspace and airports. It also describes the three types of control centers and the various air traffic control positions that monitor and serve a typical IFR (instrument flight rules) flight. Finally, it briefly discusses the principal types of surveillance equipment in terminal areas, along with the instrument landing systems (ILS), which are still the principal navigation aid supporting precision approaches to runways at major airports. ATFM is the subject of Secs. 13.3 and 13.4. ATFM systems in Europe and the United States have become essential to airport operations. In particular, Sec. 13.4 describes ATFM as conducted under CDM in the United States, using a simple example to explain how CDM motivates the timely exchange of information between the FAA and the airlines and thus improves the performance of ATM. Finally, Sec. 13.5 surveys near-term and medium-term innovations in surveillance, navigation, and automation in terminal airspace and airports, and identifies the types of impacts they may have.

13.2 Evolution and Main Characteristics of ATM Systems

Enormous differences exist among ATM systems around the world with respect to technology and level of sophistication. In fact, ATM systems are often said to transition from one "generation" to the next, as their characteristics evolve. This evolution is slow—each transition period may take 10 years or more. Thus, the boundaries between successive generations are blurred.

The "first-generation" ATM systems that operated in North America and Western Europe before and during World War II are notable for the definition, for the first time, of a system of airways and for their reliance on traffic reports from pilots. Air traffic controllers could not observe airborne aircraft, but voice communications from pilots kept the controllers informed of their current positions and altitudes along the airways. Controllers updated this information manually, sometimes by moving plastic strips ("shrimp boats") representing each aircraft on a map that depicted the geographic area for which they were responsible (Gilbert, 1973). Air traffic in parts of the en route airspace over Africa, Asia, and South America is still controlled in approximately the same way.

The transition from the first to the second generation of ATM systems was marked by the introduction of radar after World War II. Primary radar systems were developed, consisting of medium-range airport surveillance radar (ASR) for terminal airspace, and longer-range air route surveillance radar (ARSR) for en route airspace. These provided surveillance by relying on the "skin effect," that is, the reflection of the transmitted radar signal from the aircraft's metallic skin. Air traffic controllers were thus able to observe the horizontal position of aircraft, although the quality of the information was not particularly good. This generation of ATM systems evolved from the late 1940s to the early 1970s, in the United States and in Western Europe. A large number of terminal areas in less developed countries and even parts of the airspace in some developed countries still feature essentially second-generation ATM systems.

The adoption of digital technology for acquiring, processing, and distributing information is the main distinguishing characteristic of the "third generation" of ATM systems. This was introduced during the late 1960s and especially the 1970s in a number of countries in North America, Western Europe, East Asia, and the South Pacific. Its important components include the following:

- *Secondary surveillance radar* (or ATC Radar Beacon System, ATCRBS, in the United States), which interrogates aircraft every few seconds and receives back digitized messages that report each aircraft's identification and altitude, among other data
- Automatic aircraft tracking and alphanumeric displays using digitized radar data at both terminal and en route airspace control centers
- Extensive data processing by networks of computers
- Automation of some routine ATM tasks, such as the distribution and updating of information about each flight at all controller positions that will handle that flight

Third-generation systems also feature enhanced voice communication systems, primary radar systems (for tracking en route, terminal airspace, and airport surface traffic), and airborne navigation systems.

Finally, most developed countries have entered by now the era of "fourth-generation" ATM systems, characterized by progress in three general areas:

- Development of automation tools, both on the ground and in the aircraft, that aid controllers and pilots in ways that go beyond the routine processing and updating of data
- Use of advanced technologies such as satellite-based communications, navigation, and surveillance (CNS), precision instrument approaches supported by global positioning systems (GPS), digital data links, and advanced weather systems
- Partial decentralization of ATM decision-making, primarily through real-time collaboration and exchange of information among ATM service providers, airlines, and flight crews

While this chapter focuses on the more advanced ATM systems, primarily as they pertain to terminal airspace and airports, one should not lose track of the fact that systems resembling the first- and second-generation ones still operate in a number of countries and geographic regions. For example, it is remarkable that until 1999, when Greece transitioned to a state-of-the-art ATM system for en route and terminal airspace traffic control, a 3.5-hour flight from London to Athens traversed the technological history of ATM, starting with some of the most advanced ATM systems in the world (in the United Kingdom and elsewhere in Northwestern Europe) and ending with a first-generation en route system in Greece and an essentially second-generation system in the terminal airspace of Athens.

Airspace Structure

Flights are conducted under either visual flight rules (VFR) or instrument flight rules (IFR). When flying under VFR, pilots must try to stay outside of clouds at all times and are responsible for maintaining safe separation from all other aircraft by visually scanning their surroundings (Braff et al., 1994). By contrast, air traffic controllers on the ground are responsible for maintaining safe separation between any two aircraft flying under IFR in controlled airspace (discussed as follows). While IFR flight was initially associated with instrument meteorological conditions (IMC), essentially all airline jet flights, as well as many other commercial and general aviation flights, now operate under IFR even in good weather (visual meteorological conditions, VMC). In the presence of increased traffic density, this ensures the availability of surveillance and of separation assistance by air traffic controllers during all phases of flight. Although pilots flying under IFR are technically responsible for separation between their aircraft and VFR traffic when outside clouds, in practice they receive extensive assistance from air traffic controllers in this respect. Flights operating under IFR must file a flight plan with the air navigation service provider (ANSP) responsible for each part of the airspace they will traverse and must receive clearance of the flight plan from an ATM facility. ANSP is the official term for agencies that provide ATM services. For example, the FAA is the ANSP in the United States.

To facilitate the air traffic control process, airspace is subdivided into three types: positive controlled airspace, controlled airspace, and uncontrolled airspace. In uncontrolled airspace, the ATM system does not provide any aircraft separation services. Responsibility for maintaining safe separation rests with pilots and, consequently, uncontrolled airspace is normally populated solely by VFR flights. The volume of uncontrolled airspace is gradually diminishing internationally. In the United States, for example, the only uncontrolled airspace left is essentially Class G airspace[1] (FAA, 1992a), that is, the airspace below 1200 ft above ground level and away from any busy airports.[2]

Certain parts of the airspace that are heavily populated by IFR flights are designated as *positive controlled airspace*. Access to such airspace by VFR flights is either prohibited altogether or is limited by a number of restrictions. For example, in the United States, only IFR flights can operate at an altitude between 18,000 ft mean sea level and FL 600.[3] This is called *Class A* airspace, and it is the part of the airspace that airline jet flights almost always utilize during their en route phase of flight (Fig. 13.1). The second class of positive controlled airspace, *Class B*, is particularly relevant to this text, as it comprises the terminal area airspace around the busiest airports. Access to Class B airspace is not limited to IFR flights, but is restricted: to be admitted to this airspace, aircraft operating under VFR must be equipped with appropriate communications and navigation equipment, must obtain clearance from air traffic control to enter the airspace, and must be operated by pilots holding a private certificate or a student certificate with appropriate instructor endorsements (FAA, 1992b). Class B airspace (also referred to as *terminal control airspace*, TCA) is shaped like an inverse wedding cake, extends to an altitude of at least 10,000 ft above ground level (AGL) and is usually centered on a major airport (see Fig. 13.1). VFR aircraft which are accepted in Class B airspace must comply with instructions issued by air traffic controllers, who are responsible for ensuring standard separations between every pair of aircraft in that airspace, whether operating under visual or instrument flight rules. As a rule, Class B airspace is fully contained within the jurisdiction of a Terminal Radar Approach Control (TRACON) facility (discussed as follows) in the United States.

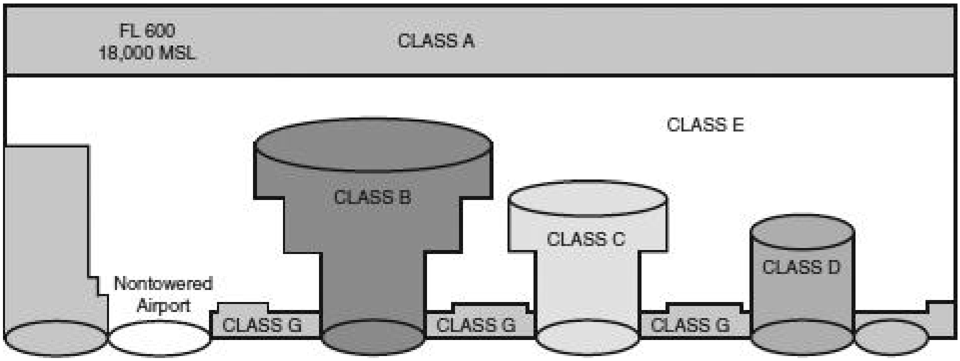


FIGURE 13.1 Classes of airspace, United States. (*Source*: Modified from Nolan, 2011.)

In the third type of airspace, controlled airspace, the air traffic controller's responsibility for maintaining standard separations is limited to pairs of IFR aircraft only, as well as to runway operations, when applicable. IFR aircraft are still responsible for separating themselves from VFR aircraft, while VFR aircraft are responsible for separating themselves from all other aircraft. Controlled airspace includes Classes C, D, and E. Airspace around medium-sized airports, which are not close to major airports, is usually in *Class C*. In Class C airspace, air traffic controllers will issue traffic advisories and conflict resolution advisories for IFR/VFR aircraft pairs and traffic advisories for VFR/VFR pairs. *Class D* and *Class E* respectively refer to airspace around smaller airports with air traffic control towers, and to all other controlled airspace and control zones around airports without air traffic control towers. The volume of Class E airspace is vast.

Handling of a Typical Airline Flight

In countries and regions where the ATM system is well-developed, three types of control facilities play a critical role during the successive phases of a typical airline IFR flight between two sizable airports. In generic terms, these are the *airport traffic control tower*, the *terminal airspace control center*, and the *en route control center*. Different countries may use different names and acronyms for these facilities. In the United States, the major terminal airspace control centers are called *approach control facilities* or TRACON facilities—names that do not fully reflect their function, as these facilities handle departures in addition to arrivals. Similarly, en route control centers are known as *Air Route Traffic Control Centers* (ARTCC). Figure 13.2 summarizes the role that each of these three types of facilities plays in controlling a typical flight. While the figure is drawn for the ATM system in the United States, ATM facilities elsewhere have an analogous allocation of responsibilities.

Type of Facility	Terminal Area Facilities		En Route Facilities	
Controlling Facility	Airport Traffic Control Towers	Approach Control Facilities	Air Route Traffic Control Centers (ARTCCs)	
Type of Control	Ground Traffic Control Takeoff and Landing Control	Approach and Departure Control	ATC during Transition and Cruise	
Airspace	Airport Traffic Area Typically 5 nmi and 3000 ft AGL	Approach Control (Tracon Area) Typically extending up to 40 nmi + 10,000 ft from the airport	En Route Airspace Transitional Phase Typically 50–150 nmi from airport	Cruise Phase Up to 60,000 ft
Typical Flight Time	Typical Ground Time 5–10 min	Typical Flight Time 10–20 min	Typical Flight Time 10–20 min	Typical Flight Time 20 min to several hrs
Flight Profile				

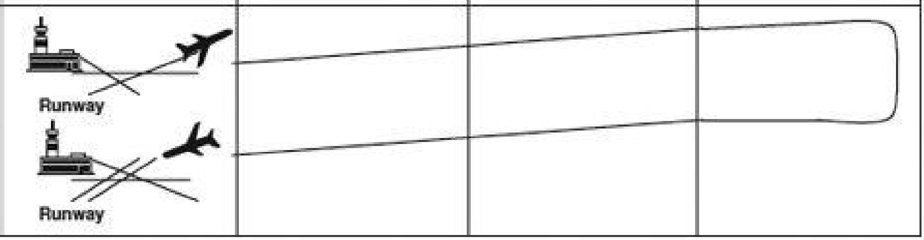


FIGURE 13.2 Role of the three principal types of ATM facilities in a typical flight. (*Source:* Mundra, 1989.)

The operator (airline or other) of an IFR flight is required to file an IFR flight plan with the ANSP (the FAA, in the United States). This is usually done one or more hours before the expected time of departure. Flight plans for regularly scheduled airline flights are typically stored in a computer and activated automatically some time before a flight. They may be updated as the time of departure approaches and even modified while the flight is airborne. The flight plan contains a detailed description of the route to be flown and must receive clearance by the ATM service provider responsible for the airport of departure. In the United States, the ARTCC with jurisdiction over the departure airport typically gives clearance, possibly after suggested modifications have been made to the flight plan. The operator of the flight is expected to adhere as precisely as possible to the flight plan and to notify the FAA of any significant changes before or during the flight. The approved flight plan is also entered into the FAA's computer system and may be updated several times prior to and during the actual flight. The flight plan is used by the computer system to notify the various air traffic control jurisdictions and positions that will handle a particular aircraft of the imminent entry of that aircraft in the airspace ("sector") or airport area of their responsibility. These notices give air traffic controllers advance information about traffic loads and typically appear at each position 20 to 30 minutes before the aircraft's arrival in the form of either a printed message ("flight strip") or a message on an electronic display.

Airport Traffic Control Tower

The airport traffic control tower provides a good vantage point for observing most of the airfield under good-visibility conditions. The following traffic control positions are located in the tower: clearance delivery; gate hold (only at some of the busiest airports); ground control; and local control. Depending on the size and complexity of the airport, these positions may be staffed by more than one controller. For example, two local controllers are active most of the time at airports that operate with two independent runways. In addition to these positions, each of which can make voice contact with pilots, a tower supervisor oversees operations, while one or more other controllers provide support services, primarily in the form of entering, updating, processing, or distributing flight data.

The first contact between an aircraft and the traffic control tower prior to a flight's departure takes place when the pilot requests predeparture clearance for the flight from the clearance delivery position. If a gate-hold position exists in the tower, the pilot is then transferred to the frequency of that position. The pilot may be informed at that time of a "gate holding" (or "ground holding") delay due to traffic flow management restrictions or other reasons (see Sec. 13.4). When the aircraft is finally ready to leave its apron stand, the clearance delivery position—or the gate-hold position, if one exists—is contacted again for permission to push back from the gate or move out of a noncontact stand, as the case may be. Once permission to leave the apron area is given, the aircraft is handed off to a *ground controller*, under whose instructions and supervision it proceeds through the taxiway system to its assigned departure runway.[4] Just before it reaches the departure runway (or the queue of aircraft awaiting takeoff), the aircraft is handed off to the *local controller*, who supervises the aircraft's takeoff. Soon after the aircraft is clear of the runway, the local controller hands it off to the appropriate departure control position in the TRACON.

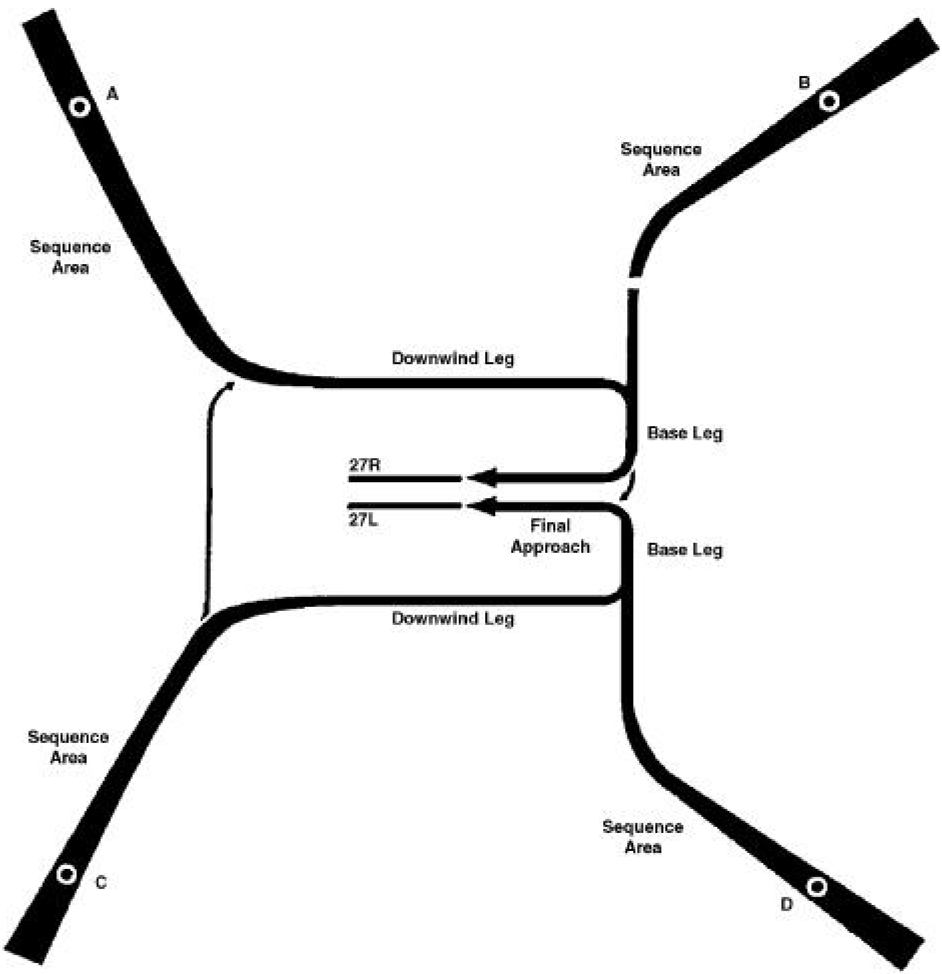
Conversely, during the arrival phase of a flight, responsibility for a landing aircraft is handed off by one of the final control positions in the TRACON to a local controller in the tower. This happens when the aircraft is on its final approach segment to the arrival runway. After the aircraft is safely off the runway, usually on an exit taxiway, it is handed off to a ground controller, who guides the aircraft to its apron area.

The airport tower is equipped with several visual displays, the most important of which are the radar display for the terminal airspace, weather-related displays, and a display of airport surface traffic, which in the United States is increasingly the ASDE-X display. The latter is particularly important when visibility from the tower is poor—for example, when it is surrounded by fog—and visual contact with the airfield is fully or partly lost. Electronic displays of conditions in apron areas (e.g., showing occupied stands) may also be important, especially at airports where physical obstructions may impede visual contact between the tower and such areas. Other controller aids may include an air traffic situation display tied to the ATFM system (see Sec. 13.4), showing traffic headed toward and away from the airport in question; flight-strip trays or electronic displays showing imminent aircraft movements; and displays indicating equipment outages, pending runway closures, etc. Under various modernization programs, airport towers are increasingly being equipped with various decision-support ("automation") tools intended to assist controllers in increasing the efficiency of airport operations, as well as safety-related monitoring and alert systems for preventing the occurrence of hazardous events, such as runway and taxiway incursions by taxiing aircraft or ground vehicles (Sec. 13.6).

Terminal Airspace Control Center

The physical location of the terminal airspace control center varies. In many cases it is in the building below the airport control tower. However, it can also be at a different location, especially in the case of multi-airport systems controlled by a single terminal airspace control center (New York, London, Paris, Milan, etc.). Terminal airspace control centers near major metropolitan regions are also responsible for coordinating traffic headed to/from that region's major airport(s) with traffic to/from secondary, mostly general aviation, airports in these regions.

Figure 13.3 shows an idealized two-dimensional depiction of arrival patterns in a terminal airspace "feeding" two parallel arrival runways, 27R and 27L. It helps explain the way operations are carried out in major terminal areas in the United States and the allocation of responsibilities in a TRACON (Mundra, 1989). Despite local differences, this description is typical of advanced ATM systems elsewhere. As Fig. 13.3 shows, aircraft enter the airspace that the TRACON controls over one of four navigational *arrival fixes* or *approach feeder fixes*, A through D, located at a distance of 30 to 40 mi (50–65 km) from the airport. Arriving aircraft are handed off at these fixes, usually by an en route control center, to one of the *arrival control* positions in the TRACON. The symmetric configuration of fixes shown in Fig. 13.3 is called a *four-corner* or *four-post arrangement* and is encountered at locations where airspace is unrestricted by major physical obstacles or other constraints. Two arrival control positions will typically be active with an arrival configuration such as this. One arrival position ("Arrival North") will be responsible for aircraft entering the TRACON's airspace from the north (fixes A and B) and the other ("Arrival South") for those entering from the south (fixes C and D). It is the arrival position's responsibility to develop a desirable sequence of landing aircraft while funneling them through the *sequencing area* and onto either the *downwind leg* (e.g., for aircraft entering from C) or the *base leg* (e.g., for aircraft entering from D), where they are handed off to a final control position in the TRACON. For the configuration shown, there will typically be two final control positions ("Final North" and "Final South"), with each position being handed aircraft primarily by the corresponding arrival control positions (“Arrival North” and “Arrival South,” respectively). The final control positions are responsible for directing aircraft through the *pattern area*, which consists of the downwind leg and the base leg and onto the final approach. At or near the beginning of the final approach, control over landing aircraft is transferred from the TRACON’s final control position to the airport traffic control tower’s local controller. The final control positions in the TRACON have a critical role in spacing aircraft for the final approach and in merging streams of aircraft at the base leg (e.g., aircraft entering from arrival fix A with those from arrival fix B). Merging may also occur at the beginning of the final approach, especially when only one exists, as in the case of airports with a single runway.



Note: A, B, C, D are arrival fixes, and are typically located 30–40 mi from the runway.

FIGURE 13.3 Idealized planar view of terminal area arrivals pattern. (*Source*: Mundra, 1989.)

It is also possible that aircraft entering the terminal airspace from the south (fixes C and D) will be routed onto the north downwind leg or final approach (or vice versa) in order to better balance the traffic load on final approaches and runways.[5] This is indicated by the narrow arrows on the left side of Fig. 13.3 and at the end of the two base legs.

The geometric configuration shown in Fig. 13.3 is by no means standard. The actual configuration depends on several factors, such as the precise location of the primary airport or airports served by the TRACON, the presence of important physical obstacles, the traffic composition (high-altitude vs. low-altitude traffic), and the runway configurations available. Three-corner arrangements of the arrival fixes are used at several locations. The number of controller positions also varies. The Boston TRACON, for instance, has two arrival positions for traffic into Boston/Logan, but only a single final position.

The TRACON must also handle departures following takeoff. The local controller in the tower hands off aircraft to one of the *departure control* positions in the TRACON. For a configuration such as the one shown in Fig. 13.3 there will usually be two such positions. Departure controllers oversee the ascent of an aircraft through the TRACON's airspace and eventually hand it off to another facility, usually an en route center, for the transitional and cruise phases of the flight.

The volume of airspace associated with a terminal control facility, which in the United States may extend as high as 17,000 ft above ground level, is structured so that departing traffic is separated in altitude from arriving traffic. In every horizontal section of the airspace, certain altitude bands are reserved for arrivals and others for departures. In cases where the TRACON serves more than one major airport, the overall three-dimensional structure of the airspace may be complex. For instance, the structure of the New York airspace changes dynamically, with certain portions of it allocated to either LaGuardia or Kennedy airports, depending on the runway configuration in use at each airport. Prominent examples of highly complex terminal airspace and procedures include those of London, New York, and San Francisco Bay—due to the presence of multiple major airports—and of Hong Kong and Zürich—due to a combination of local physical obstacles and airspace constraints imposed by different local or national jurisdictions.

In addition to the arrival, final, and departure control positions in the TRACON, which are dedicated to controlling traffic to and from the primary airport(s), a varying number of other arrival and departure positions may have responsibility for control of traffic into and out of secondary airports in the same area. Terminal airspace control centers in the United States and elsewhere are also staffed by a supervisor, a varying number of personnel performing flight data processing and other support functions, and possibly by ATFM specialists, who provide coordination with the national (or international, in the case of European countries) ATFM system (see Sec. 13.4).

In general, centers controlling the terminal airspace of hubs and of major international airports often perform the most complex tasks in the ATM system. These centers are critical to the efficient operation of the entire air transportation system. For this reason they are staffed by some of the most experienced and skilled air traffic controllers. Not surprisingly, these centers are also the focus of many programs aimed at introducing advanced ATM automation aids and decision-support tools (see Sec. 13.6).

Central to the operation of a terminal airspace control center are the information-processing systems and associated displays that serve as the air traffic controllers' primary source of information for managing and controlling traffic. Many countries in Europe, North America, and Asia and the Pacific rim are in the process of installing much improved systems and displays in terminal airspace control centers, often as replacements of antiquated ones. For example, the FAA has finally replaced ARTS, the Automated Radar Terminal System, first installed in the 1970s, with STARS, the Standard Terminal Automation Replacement System.[6] STARS receives and processes traffic and weather data from the primary and secondary traffic and weather radars and presents this information to air traffic controllers in high-quality, color displays. By displaying six distinct levels of weather "intensity" (identified by different colors), as defined by the National Weather Service, and by superimposing traffic and weather data, STARS assists controllers direct air traffic around bad weather. It can track up to 1350 airborne aircraft simultaneously within a terminal area and can interface with up to 16 short- and long-range radars, 128 controller positions, and 20 remote airport towers in a 400-by-400-mi region. Equally important, it has been designed with an open architecture that facilitates integration with advanced decision-support tools, such as CTAS (see Sec. 13.6). STARS also includes a built-in backup in case of failure of the primary system.

Surveillance

Surveillance and navigation are two of the fundamental functions of ATM systems. Some of the most important types of navigation and surveillance equipment used in terminal airspace operations are briefly described next.

Surveillance is the function that provides the current location of an aircraft to air traffic controllers. This can be accomplished in three different ways: (1) pilot reporting of the aircraft's position and altitude via voice communication; (2) returns from primary surveillance radar; and (3) automatic responses to secondary surveillance radar. Modern terminal airspace control centers—and, more generally, all advanced ATM systems—rely heavily on the second and third approaches. Interestingly, an automated form of the first approach, *automatic dependent surveillance* (ADS), is likely to become increasingly important as a means of surveillance in the near future (see discussion that follows).

Primary surveillance radar relies on "skin effect" or "skin tracking" to obtain information about an aircraft's position. A rotating antenna on the ground emits pulses, which are reflected by the metallic exterior of the aircraft and returned to the antenna. This process generates the information needed to determine the polar coordinates—distance and angle (or *azimuth*)—of each aircraft relative to the antenna. By measuring the time it takes for a round trip of the pulse, the distance of the aircraft from the antenna is computed, while the azimuth is derived from the corresponding angular position of the antenna. Note that no information is obtained about the altitude of the aircraft in this way. The latter must be determined either through pilot reporting or, now routinely, through the secondary surveillance radar (discussed as follows). This type of surveillance is also referred to as *independent*, because it requires no avionics equipment on the aircraft.

The primary radar used to track traffic in terminal airspace is known as *airport surveillance radar* (ASR). ASR technology has advanced greatly, providing good-quality target resolution and high levels of reliability. Its most modern versions are digital—the ASR-9, installed at the busiest airports, and the ASR-11. ASR is sometimes referred to as "short range," to distinguish it from long-range radar, which is used for surveillance of en route airspace (*air route surveillance radar*, ARSR). ASR typically performs 10 to 15 revolutions per minute, and thus updates information about each aircraft's position every 4 to 6 seconds. It has a principal range of 30 to 60 nmi, which may extend as far as 120 nmi in some of the most recent versions.[7]

Another type of primary radar, known as *airport surface surveillance radar* (ASSR), is used for surveillance of traffic, including ground vehicles, on the airport's surface. Despite considerable progress, ASSR still presents difficult technical challenges, as it is highly sensitive, by design, and must cope with the many reflecting surfaces and physical obstacles at or around airports that may provide false or distorted signal returns. The most advanced versions of ASSR, known as ASDE or *airport surface detection equipment*, combine radar with multilateration techniques to overcome most of these problems. The ASDE-X (ASDE-Model X) radar, despite its high cost, is now being installed at all major airports in the United States—and at many airports internationally—and has been instrumental in reducing runway and taxiway incursions and the frequency of collisions on the airport's surface. ASDE plays a particularly important role during periods of poor visibility, including nighttime.

The *secondary surveillance radar* (SSR) is a rotating antenna on the ground, which emits interrogation messages at a frequency of 1030 MHz that trigger automatic responses from a transponder on the aircraft in the form of a digitized message on a different frequency[8] (1090 MHz). SSR-based surveillance is called *cooperative* because it requires that aircraft carry a transponder. The transponders are distinguished into "modes," depending on the format of their response messages. For all practical purposes, only two such modes are currently in use, Mode C and Mode S.

Mode C transponders automatically report aircraft identification (four-digit code) and altitude at 100-ft increments in 13-bit responses. They cause significant congestion of the 1090-MHz frequency because they generate multiple responses to a single interrogation. Mode C responses from different aircraft on the same angle from the radar beacon also often interfere with one another ("garbling"), resulting in loss of information. In the United States, a Mode C transponder is the minimum requirement for all aircraft flying above 10,000 ft or within 30 mi of a major airport.

Mode S transponders offer improved capability for aircraft-specific interrogation and provide for response messages with greatly increased information content. Each aircraft is assigned a 24-bit identification number, and the ground antenna can selectively interrogate any specific aircraft through that number. Moreover, the interrogation message can request a reply in any one of 256 message formats. Reply messages are 56 bits long. Far more information than just aircraft identification and altitude can thus be transmitted. Mode S transponders respond only once to each interrogation. This reduces frequency congestion and the possibility of message garbling. Finally, in addition to responses to interrogation messages, Mode S transponders automatically broadcast about every second a short message (called a "squitter") with their identification. This message is broadcast throughout a flight (whether or not the aircraft is within range of an SSR) and provides part of the basis for the operation of traffic alert and collision avoidance systems (TCAS). Due to their essential role in TCAS, the United States requires Mode S transponders on all jet aircraft with more than 10 seats and on all commercial aircraft with more than 30 seats.

Automatic Dependent Surveillance systems are based on the notion of relying on aircraft to report their own position. In ADS, the aircraft determines its position, through GPS or other means, and automatically reports that position to the ANSP and to other aircraft. One can differentiate among different ADS systems depending on who receives the position reports and for what purpose. ADS-A (for "addressable") allows the aircraft to exchange information with the ANSP upon request, while with ADS-C (for "contract") the aircraft transmits position information continuously or upon the occurrence of specific events, such as crossing specified waypoints. Both of these systems are particularly useful in oceanic operations or when flying over remote areas. For terminal airspace and airport operations, ADS-B (for "broadcast") may end up playing a most important role in the near future. Aircraft equipped with ADS-B broadcast information about their position to ATM facilities and to all other equipped aircraft in their vicinity. That information can then be processed as needed by the ATM system, as well as displayed directly in the cockpit of aircraft equipped with a CDTI (Cockpit Display of Traffic Information). Because of the high quality of the information transmitted in this way, the high update rates, and the low cost of the equipment, ADS-B has emerged as an attractive future alternative to radar-based surveillance.

Navigation for Precision Instrument Approaches

The ability to continue operating at near-normal levels in instrument meteorological conditions is essential to major airports. The *instrument landing system* (ILS) is by far the most widely used navigation aid for conducting precision approaches to an airport under low visibility conditions. Practically every major airport in the world is equipped with at least one ILS. The ILS provides landing aircraft with a single, straight-line path that they can follow on final approach to the runway. The straight-line path may extend as far as 25 to 30 km (~ 15–20 mi) from the near end of the arrival runway. This straight line is the intersection of two planes that are defined by two different types of transmitters (Fig. 13.4):

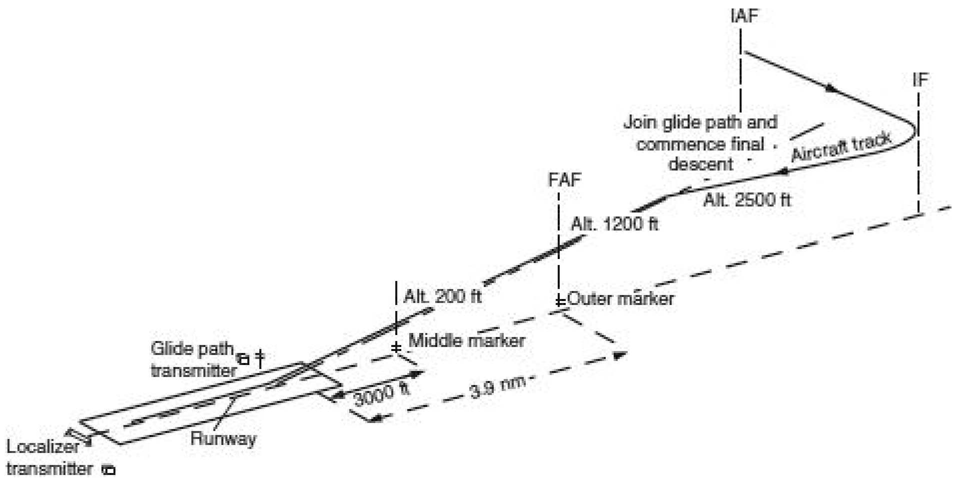


FIGURE 13.4 Schematic configuration of an Instrument Landing System. (*Source:* Mundra, 1989.)

1. *The localizer*, which defines a plane (sometimes referred to as the runway centerline plane) that extends vertically up from the runway centerline and the extension of that centerline.
2. *The glide slope*, which defines a "ramp," typically inclined at an angle of 2.5 to 3° to the horizontal plane.[9]

The localizer operates at a VHF frequency in the range of 108.10 to 111.95 MHz and provides lateral guidance by indicating whether the aircraft is to the left or to the right of the extended runway centerline. The system consists of a localizer transmitter, located roughly 1000 ft from the far end (or “departure end”) of the runway and a localizer transmitter building, located about 300 ft to one side of the localizer antenna.

The glide slope operates at a UHF frequency of 329.3 to 335.0 MHz and provides vertical guidance by indicating whether the aircraft is above or below the correct line of descent to the runway. The glide slope transmitter building and antenna are located to one side of the runway centerline at 250 to 600 ft from the centerline and at a typical distance of 1000 ft from the near end (or “approach end”) of the runway.

Two or three *marker beacons*, transmitters that emit a cone-shaped local signal, supplement the localizer and glide slope. When this signal is received by aircraft flying over the beacons, it “marks” the position of that aircraft along the final approach course. The two standard marker beacons on all ILS are the *outer marker*, at roughly 4–5 mi (6.5–8 km) from the near end of the arrival runway on the extension of the runway centerline, which marks the final approach fix (FAF) to the runway, and the *middle marker* at about 3000 ft (900 m) from the runway. A Category II ILS (discussed as follows) is also equipped with an *inner marker* located approximately 1000 ft (300 m) from the end of the runway.

Figure 13.5 indicates the sequence of events that take place in connection with an ILS approach. The aircraft “captures” the ILS-defined final approach path, typically at a distance of 5 to 10 nmi from the runway threshold, and follows that path until it reaches the appropriate decision height (discussed as follows). A missed approach is executed if visual contact with the runway has not been made by the time the decision height is reached.

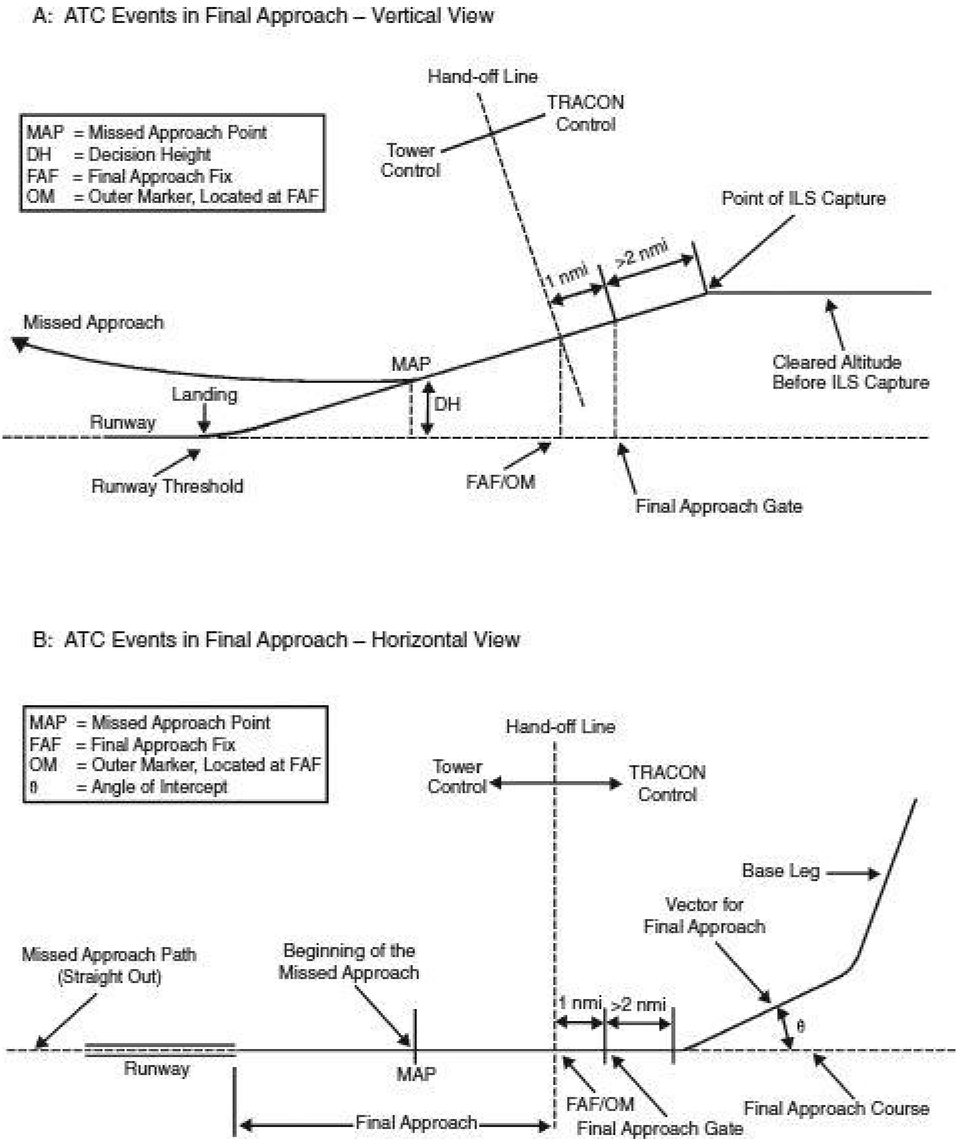


FIGURE 13.5 Side and planar view of a precision approach. (*Source*: Mundra, 1989.)

As described in Table 13.1, instrument landing systems are classified into three categories, depending on decision height and on visibility or runway visual range (RVR). Equipment along the runway measures this RVR distance, defined as the range over which the pilot of an aircraft on the centerline of a runway can see the runway surface markings or the lights delineating the runway or identifying its centerline (ICAO, 2009).

	Decision Height	Visibility or Runway Visual Range (RVR)
Category I	60 m (200 ft)	Visibility: 800 m (0.5 mile) *or* RVR: 550 m (1800 ft)
Category II	30 m (100 ft)	RVR: 350 m (1200 ft)
Category III-A	0 m	RVR: 200 m (700 ft)
Category III-B	0 m	RVR: 50 m (150 ft)
Category III-C	0 m	RVR: 0 m

Table 13.1 Precision Instrument Approach Categories (ICAO, 2009)

The standard ILS is a Category I system. The Category II system differs in relatively minor ways from Category I; it requires an inner marker and some additional approach lighting. However, Category III ILS is significantly different (requiring specially designed localizer and glide slope systems), can be deployed only at sites satisfying stringent landscaping requirements, and is far more expensive. Only specifically certified pilots and aircraft may perform Category II and Category III approaches. Runways approved for Category I, II, and III approaches should also satisfy various requirements with regard to obstacle clearances and parallel taxiways (see also Chap. 9), paved runway length, runway markings, holding position signs and markings, runway threshold location, runway edge lights, and approach lighting systems. These requirements (e.g., for approach lighting systems) may differ for each of the three categories. FAA (2012) gives details and additional guidance to relevant documents.

Instrument landing systems have a number of disadvantages. First, the quality of their signals can be affected seriously by distortions and reflections caused by both stationary and moving objects and vehicles. This is particularly true of the localizer signal. In addition, the presence of several inches of snow or of high sea waves, in the case of coastal airports, may affect significantly the accuracy of the glide slope transmission. These systems therefore require careful calibration and potentially difficult adjustments to adapt to local topography, as well as continuous monitoring of performance. Other precautions include the designation of a localizer critical area and of a glide slope critical area, where the presence of aircraft and vehicle traffic or of objects and obstacles (e.g., piled snow) when aircraft are performing ILS approaches must be either completely prohibited or strictly controlled (see Chap. 9).

A second disadvantage is the fact that the ILS provides only a single, straight-line path to the runway for instrument approaches. This may force aircraft to fly persistently over areas that are sensitive environmentally. In addition, having aircraft fly in a single file toward the runway will lower operating efficiency when aircraft with different approach speeds use the same runway. In such cases, the spacing between a "fast" arriving aircraft and a "slow" one that trails immediately behind will increase as the two aircraft fly down the final approach path,[10] thus reducing airport capacity (see Chap. 10).

Finally, a scarcity of frequencies exists for installation of new ILS in major terminal areas where several of them are already in operation. The resulting inability to develop additional ILS approaches to existing runways may restrict airport capacity.

The global positioning system (GPS)—and other existing or future global navigation satellite systems (GNSS)—can also be used for airport approaches. As of 2012, the standard civilian GPS system already provides sufficient accuracy by itself for nonprecision approaches. GPS-based approaches are thus becoming increasingly common at smaller airports everywhere. However, the positioning accuracy of the GPS (as well as its availability and integrity) does not meet requirements for precision approaches. For this reason, augmentation of the GPS positional information is needed to attain the required performance. Two different technical approaches can provide such augmentation. A *local area augmentation system* (LAAS) essentially consists of a reference station located at or near an airport and a monitor station that together make it possible to measure precisely any GPS errors at that airport and transmit corrective information ("differential data") to aircraft. LAAS is also known as *ground-based augmentation system* (GBAS) because of its reliance on these ground stations. An LAAS can support even Category II and III approaches by providing sufficient navigational accuracy to distances of 20 mi or more from an airport. It may also provide the means to navigate on the airport's surface in low-visibility conditions. Eventual installation of LAAS at as many as 150 airports in the United States is expected.

The alternative to LAAS is the *satellite-based augmentation system* (SBAS), known in the United States as the *wide area augmentation system* (WAAS). It is designed to provide corrections to the GPS signal regionally or nationally, for example, over an area comparable to that of the United States. It consists of a network of ground reference stations, master stations, and a geosynchronous communications satellite that broadcasts corrections to the GPS signal to aircraft. A WAAS can support approaches only to roughly Category I approach limits. This capability is already available in the United States, while analogous SBAS systems are being developed or are already operational in many other nations and regions.

En Route Control Centers

En route control centers handle IFR traffic outside terminal airspace and thus are responsible for controlling practically all airline en route traffic. Only a brief description of the operation of en route control centers will be presented here. MIT Lincoln Laboratory (1998) provides far more detailed and still up-to-date information.

En route control centers have jurisdiction for parts of the airspace of a country or a region, with the number of centers varying with the size of the area to be covered. Figure

13.6 shows 20 such centers (ARTCC), which control en route airspace over the continental United States, whereas a single center may control the en route airspace of a small country. Due to multiple national jurisdictions, 80 centers operate in Europe (not including Russia) in a land area of size similar to that of the United States. The airspace controlled by an en route center is, in turn, subdivided into sectors that constitute the fundamental "unit" of airspace volume from the ATM point of view. In the United States, en route sectors are referred to as super-low, low, high, and super-high, according to the flight levels they control, with varying floors and/or ceilings for each of these types. For example, "high" sectors often have a floor at Flight Level FL 240, but their ceiling may vary from FL 310 to FL 370. "Super-high" sectors typically have floors at FL 350 or higher.

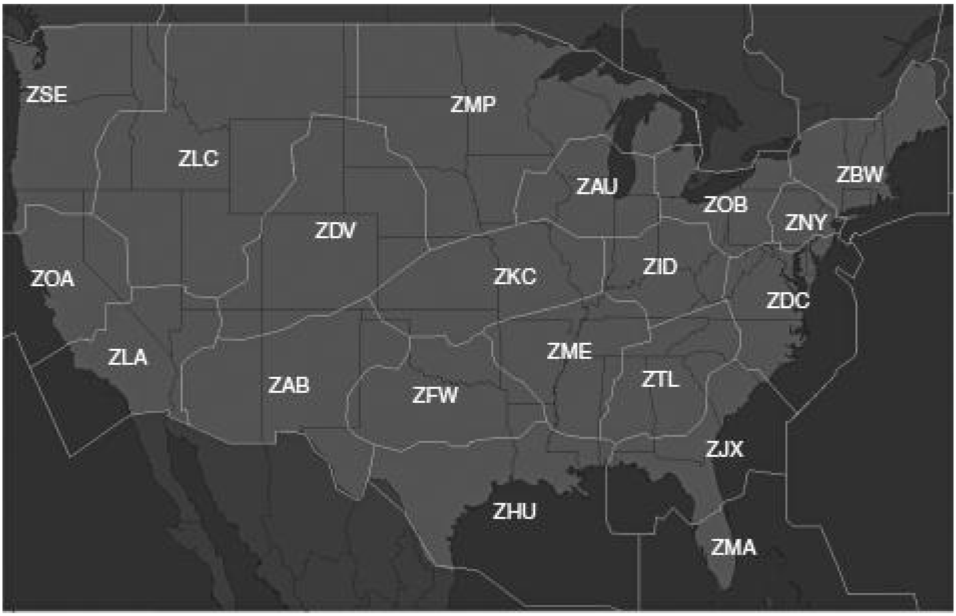


FIGURE 13.6 En Route Air Traffic Control Centers, United States. (*Source*: MIT Lincoln Laboratory, 1998.)

Navigation en route still depends largely on systems of airways, essentially networks of "highways in the sky." A set of ground-based navigation aids, the *VHF omnidirectional range* (VOR) *finders*, generally define these airways. In the United States, the network of Victor airways extends up to (but not including) 18,000 ft above mean sea level, and the network of jet routes from 18,000 to 45,000 ft. The centerline of each airway coincides with a radial (i.e., a straight line at a given angle) projecting out of the antenna of a VOR station. Each VOR station has its own frequency. By tuning a navigational radio to the proper frequency, a pilot can fly along straight radial paths from one VOR station to another. To fly between two points A and B on the earth's surface, an aircraft may, for instance, fly from A to X, from X to Y, and from Y to B, where X and Y are the locations of the antennas of VOR stations. This accounts for the "dogleg" paths that most aircraft still fly in traveling between two points on the earth's surface.

Practically all VOR stations are also equipped with some type of *distance-measuring equipment* (DME), which an aircraft can use to determine the distance between itself and the VOR station. Stations that combine VOR and distance-measuring capabilities (known either as VOR/DME or as VORTAC facilities) thus provide the means of navigation by making it possible for each aircraft to determine its polar coordinates (distance and angle) relative to the (known) locations of these facilities.

Area navigation (RNAV) refers to the capability to navigate directly between any two defined points without having to adhere to the system of airways. RNAV allows airspace users to specify, through a set of waypoints, an origin-to-destination path—a *user-preferred route*—which is optimal according to some combination of performance criteria such as minimizing travel time or minimizing fuel burn. RNAV can be performed by utilizing various navigation aids, including the network of VOR/DME or VORTAC facilities, the long-range (civilian) navigation system (LORAN-C), and inertial navigation systems (INS). By far the most important of these aids is GPS and other global navigation satellite systems. Since GPS and INS do not require the support of ground stations, such as VOR/DMEs, they are also useful for navigation over oceans and remote tracts of land—and have, in fact, revolutionized air travel over such areas.

While all commercial jet aircraft now have RNAV capabilities, the ATM system in many countries is often unable to accommodate routinely requests for user-preferred, point-to-point routes, because of the additional complexity that such routes imply for the prediction and resolution of conflicts. The National Route Program (NRP), in effect in the United States since the early 1990s, authorized user-preferred routes in phases, beginning at the highest flight levels and lowering progressively over time the altitude above which NRP flights are routinely authorized.

13.3 Air Traffic Flow Management

The development of advanced air traffic flow management (ATFM) systems in the United States and in Europe during the 1980s and 1990s has had an enormous impact on ATM and airport operations. ATFM has become essential to keeping the delays of airborne aircraft within manageable levels, thus reducing the cost of delays to airlines and other airspace users, and achieving better utilization of airport and ATM resources. ATFM undoubtedly produces important safety benefits as well, by controlling the flows of aircraft into crowded portions of airspace and by reducing the probability that aircraft will be subjected to excessive airborne delays. On the negative side, ATFM has been criticized at times for deficiencies that occasionally contribute to slowing down air traffic operations and exacerbating, rather than reducing, delays. These flaws, however, are gradually being corrected as expertise, technological support, and decision-making processes improve.

The adoption of a CDM approach to ATFM has been one of the most important ongoing developments in ATM. It has greatly accelerated progress toward correcting some of the earlier deficiencies of ATFM. While ATFM and the CDM approach have an impact on every part of the airspace, their effects on airport and terminal airspace operations are particularly critical.

Objectives and Limitations of ATFM

The objectives of ATFM can be summarized as (1) preventing any overloading of airports and ATM facilities and services that might affect safety and (2) minimizing the economic and other penalties imposed on aircraft operators by air traffic congestion. This is accomplished by adjusting the flow of aircraft dynamically so that demand matches the available capacity at airports, in terminal airspace, and in en route airspace.

The extensive use of ATFM is a relatively recent development. While essentially *ad hoc* ATFM systems operated both in the United States and in Europe during the 1970s, the event most responsible for stimulating interest in advanced ATFM was the 1981 strike of air traffic controllers in the United States. To reduce pressures on the ATM system, it was decided that airborne delays should be avoided as much as possible during the strike. This was accomplished by holding aircraft on the ground prior to takeoff. An aircraft would not be allowed to take off unless there was reasonable assurance that, after departure, it would be able to proceed to its destination with a minimum amount of delay in the air. This marked the first extensive application of the strategy of ground holding and provided a good illustration of the meaning of "matching demand to available capacity." After the strike ended and the ATM system returned to normalcy, use of ground holding was maintained as an option for dealing with the most serious instances of air traffic congestion. Equally important, the 1981 experience resulted in better appreciation of the potential of ATFM.

While ATFM can ease the effects of congestion and overloading, its contributions have some inherent limitations. A helpful distinction in this respect is between ATFM interventions caused by "bottlenecks" in en route sectors and those that are made in response to inadequate airport capacity, either at the runway system or in terminal airspace. In the former instance, the en route bottlenecks can often be bypassed at modest cost to aircraft operators through countermeasures such as rerouting of aircraft and restructuring the flows of traffic in the airspace. In such cases, ATFM essentially generates some additional capacity by mobilizing airspace resources that would not have otherwise been utilized. In this way, ATFM can reduce the total delay that aircraft operators would otherwise experience collectively. However, when the bottleneck is at a flight's airport of destination, the locus of the congestion cannot be bypassed. Delay is then unavoidable and ATFM can do little about reducing the total delay to all aircraft operators, as measured in units of lost time. ATFM can, nonetheless, accomplish two things in such situations:

1. Reduce the cost of unavoidable delay—for instance, by delaying aircraft on the ground and thus providing savings on fuel consumption
2. Modify, if desirable, the way in which unavoidable delay is distributed among aircraft operators

At the same time, it is important to note that flow management may, in practice, increase total delay if not performed properly or applied excessively. For example, ATFM actions may sometimes lead to underutilization of the available airport capacity, as explained in the discussion that follows.

ATFM Operations

The ATFM systems of Europe and of the United States and Canada are operated hierarchically, as might be expected of systems that attempt to coordinate traffic flows over vast geographic regions, with local units acting under the direction and control of a centralized unit. In the United States, the Air Traffic Control System Command Center (ATCSCC), an impressive FAA facility located near Washington, DC, has the role of national coordinator, while Traffic Management Units (TMU) operate at each of the regional en route control centers (ARTCC) and at the major terminal airspace (TRACON) facilities. In addition to implementing directives of national scope issued by the ATCSCC, the regional and local TMU may also take actions of more limited scope on their own, in order to relieve local problems.

The ATFM system is more centralized in Europe, where local actions at all times must be coordinated with and, in principle, approved by the EUROCONTROL's Central Flow Management Unit (CFMU), an impressive facility in Brussels.[11] Since 1996, the CFMU has been charged with providing flow management services for all country members of the European Civil Aviation Conference (ECAC). As of 2012, a total of 44 states participated, spanning practically the entire airspace of Europe, with the notable exception of Russia. Flow management positions (FMP) operating at each of Europe's Area Control Centers (ACC) and major terminal airspace control facilities implement the directives ("regulations") of the CFMU.

ATFM systems perform three principal functions:

1. Prediction of the location of potential overloads

2. Development of strategies for relieving these overloads
3. Overseeing implementation of these strategies, often revised in "real time"

The prediction of overloads in Europe tends to be more "proactive" and more focused on en route airspace compared to the United States. EUROCONTROL's CFMU has a strategic planning phase that begins with the submission of airline flight schedules for the following (winter or summer) season, that is, about 6 months in advance. The CFMU reviews these schedules and associated probable flight routings and projects typical air traffic flows through en route sectors and at airports. It then consults with the airlines, trying to relieve anticipated habitual overloading and to balance en route sector workloads by suggesting alternative flight routings and even some modifications to flight schedules. No similar process takes place in the United States, where advance flow planning is limited to special events that may attract heavy volumes of air traffic (e.g., the Super Bowl football game) or may significantly affect air route capacity or airport capacity at some locations (e.g., the closing of a runway for repairs).

Once airline and other flight schedules are in place, ATFM must predict overloads on a daily and hourly basis. In Europe, the CFMU routinely performs most such predictions 24 to 48 hours in advance. In the United States, this is done on a shorter term basis, at the beginning of each day (typically around 6 a.m., Eastern time). In both cases, the initial predictions are constantly updated. In the United States, the principal concern is with weather conditions that may affect the capacity of key airports. Such conditions are the primary and most common cause of serious delays, although portions of en route airspace may also become problematic in the presence of weather fronts. In contrast, for European ATFM, en route sector capacity[12] is as much or more of a concern as airport capacity. This is partly due to the extensive use of airport schedule coordination in Europe, which limits *a priori* the potential for demand overloading at most of the major airports (see Chap. 12).

The strategies that ATFM deploys to deal with overloads employ three principal types of interventions:

1. *Ground holding*: intentionally delaying an aircraft's takeoff
2. *Rerouting*: changing or restructuring some flight routes to modify the distribution of traffic flows
3. *Metering*: controlling the rate at which traffic crosses some specified spatial boundaries by adjusting the spacing between aircraft or their speed

Of these, the first is the most drastic, as it controls the number of aircraft moving through the ATM system, while metering is the most tactical in nature.

Whereas the ATFM strategies employed on the two sides of the Atlantic are similar, utilizing various combinations of these three types of interventions, the European and American views differ when it comes to the question of how active ATFM should be. In the United States, the role of ATFM is seen as mostly reactive: ATFM intervenes only as called for by weather conditions or other circumstances. In Europe, by contrast, the CFMU must clear all aircraft for which it receives a flight plan, that is, give them each a "departure slot"—typically a 20-minute window—before they can leave their parking stand. Moreover, if a flight fails for any reason (e.g., slow boarding of passengers) to meet its departure slot, the flight is reassigned to a (later) departure slot, which typically means a significant delay. As is the case with overload predictions, the CFMU typically develops its overload-resolution strategies earlier than the FAA (24–48 hours ahead of the event, vs. 4–6 hours, respectively). In both instances, strategies are revised, if necessary, in response to developments in the field.

The acquisition, processing, and display of accurate and timely information are the most important prerequisites for a successful ATFM system. A noteworthy technical achievement of the FAA's ATFM system has been the development of the Enhanced Traffic Management System (ETMS) and the associated aircraft situation display (ASD). ETMS has amassed an enormous and constantly expanding information base, both historical and real-time, comprising geographic, air traffic, weather, and traffic management data.[13] Most of this information can be readily displayed through a menu-driven interface at any location equipped with the ASD.[14]

ATFM operates in a difficult decision-making environment, which can be described as information intensive, inherently stochastic (i.e., subject to uncertainty), and highly dynamic. ATFM receives and processes large amounts of information every day and must decide what part of that information is relevant to support flow management and what specific data are needed at each level of decision making. Key parameters describing future operating conditions, such as the available runway capacity at an airport, are often subject to a high level of uncertainty, even on a time horizon of less than 1 hour, due to their dependence on unpredictable or partly predictable variables, such as the incidence and intensity of fog or of thunderstorms. Moreover, operating conditions change constantly during the course of each day. To operate well in this challenging environment, ATFM must adopt strategies that take into consideration the level of uncertainty associated with key parameters and are flexible so that they may be revised as new information becomes available. In such an environment, even the best-trained human operators must be assisted in their tasks by well-designed, computer-based decision-support systems (DSS).

Ground Delay Programs

Ground delay programs (GDPs) illustrate how ATFM works, as well as the complexity of the problems that must be addressed. A GDP is initiated in the United States whenever a serious and possibly persistent overload is predicted for the next several hours at an airport. Each GDP is specific to an airport. Thus, several such programs may be run on any single day, often simultaneously.[15] Boston/Logan, Chicago/O'Hare, New York/LaGuardia, New York/Newark, and San Francisco/International are some of the airports with the highest incidence of GDPs. These are also airports where the difference between the VMC and IMC capacities of the runway system is large (cf. Table 10.4). The corrective action taken by ATFM during a GDP requires delaying on the ground, prior to takeoff, aircraft bound for the airport in question. The rationale is that it is both safer and less expensive to absorb unavoidable delays on the ground rather than in the air. The duration of a GDP, that is, the length of time during which takeoffs of flights bound for the affected airport are constrained, may be as long as 12 hours or more, generally as a result of persistent adverse weather conditions. However, a more typical duration is a few hours. Example 13.1 provides a brief and simplified description of how a GDP worked in the United States prior to the initiation of CDM in 1998. It illustrates the rationale for GDPs and explains the motivation for the adoption of CDM (see Sec. 13.5).

Example 13.1 Consider airport XYZ, where 12 flights have been scheduled to arrive between 0700 and 0900, local time, of a particular day. A forecast of heavy fog that will begin at 0700 and end at 0900 has been issued for XYZ. It is estimated that, as a result, the arrival capacity of XYZ will be reduced to 6 per hour for the duration of this event. The capacity estimate of 6 per hour is called the *airport acceptance rate* (AAR) and plays a critical role in GDPs.[16] The original schedule of arriving flights at XYZ, beginning with 0700, is indicated in the first two columns of Table 13.2. The first column identifies the airline (A, B, or C) and flight number, and the second indicates the "estimated time of arrival" (ETA) of the flight to the nearest 5 minutes, absent any GDP restrictions or other disturbances.

Airline	ETA	CTA	Delay Minutes
A1	0700	0700	0
A2	0700	0710	10
B1	0705	0720	15
B2	0705	0730	25
B3	0710	0740	30
B4	0710	0750	40
A3	0720	0800	40
C1	0720	0810	50
B5	0740	0820	40
C2	0740	0830	50
A4	0820	0840	20
B6	0840	0850	10
Total A			70
Total B			160
Total C			100
Total			330

TABLE 13.2 The Original Time Schedule and the Initial GDP for Example 13.1

Given this situation, the FAA will "run a GDP" by assigning a new time of arrival, the controlled time of arrival (CTA), to each flight, as shown in column 3. The strategy is very simple. When airport capacity is limited, that capacity is rationed among airlines in accordance with the original schedule of flights. Since in this example the capacity is down to 6 per hour, the arrival of flights will be scheduled for the "slots" of 0700, 0710, 0720, etc., that is, they will be evenly spaced during the hour according to the rate indicated by the AAR. Column 4 indicates the resulting delay to each flight, the difference between the flight's CTA and ETA. The bottom four rows of Table 13.2 show the total delay suffered by each airline separately and by all the airlines together.

The FAA will next delay the departure for XYZ of each of the flights A1 to B6 in Table 13.2 by the amount of time shown in column 4. A controlled time of departure (CTD)—also known as expected departure clearance time (EDCT)—is assigned to each flight at its airport of origin. For example, if flight C1 were originally scheduled to leave Boston at 0540 to make the 0720 ETA at XYZ, it will now be assigned a CTD of 0630, for a ground delay ("ground hold" or "gate hold") of 50 minutes.

Note that one of the objectives of the GDP is to follow a procedure and develop an arrival sequence that is considered fair to all users. This is accomplished in two ways: by allocating slots according to the original schedule of flights, as described (this is called *rationing by schedule*, RBS), and by scheduling the CTAs in the same order as the ETAs, that is, by implementing a "first scheduled, first served" policy.

The revised schedule of arrivals, as shown in column 3 of Table 13.2, usually provides only a preliminary indication of what will eventually happen. In practice, some flights may be cancelled, either for reasons independent of the GDP or because the delay assigned to them is so long that it makes little sense to perform the flight. It is also possible that a flight may be unable to meet the CTA assigned to it because of delays caused by mechanical problems or other reasons. The CTAs would be revised dynamically, if information about such flight cancellations or delays were made available to the ATFM system in a timely manner.

Consider, for example, the case in which airline A decides to cancel flight A2 for some reason. If the airline makes this known to ATFM soon enough, Table 13.2 would be revised as shown in Table 13.3. The CTA of every flight from B1 through B6 has now been changed to 10 minutes earlier, taking advantage of the gap created by the cancellation of A2.

Airline	ETA	CTA	Delay Minutes
A1	0700	0700	0
B1	0705	0710	5
B2	0705	0720	15
B3	0710	0730	20
B4	0710	0740	30
A3	0720	0750	30
C1	0720	0800	40
B5	0740	0810	30
C2	0740	0820	40
A4	0820	0830	10
B6	0840	0840	0
Total A			40
Total B			100
Total C			80
Total			220

TABLE 13.3 The GDP Obtained if Airline A Cancels Flight A2 and So Informs ATFM in a Timely Fashion

On second thought, however, the GDP shown in Table 13.3 might never come to pass. By informing ATFM of the cancellation of A2, airline A would reduce its own total delay by 30 minutes, but the delay of its competitors by a total of 80, compared to Table 13.2. Moreover, airline A would disclose to B and C the fact that flight A2 has been cancelled, information that might be valuable to these competitors in ways other than just saving delay minutes.[17] It is therefore entirely possible that airline A will choose to simply not inform the FAA of the cancellation of A2 until it is too late for the ATFM system to take advantage of the gap in the schedule. The final GDP schedule may then end up being the one shown in Table 13.4, instead of Table 13.3. Note that the 10-min slot between 0710 and 0720 has now been wasted. As this instance suggests, it may be necessary to offer incentives to the airlines in order to get them to share information with the FAA and their competitors during a GDP. They will need to know that sharing information will work to their benefit or, at the very least, not put them at a disadvantage.

Airline	ETA	CTA	Delay Minutes
A1	0700	0700	0
VOID	–	0710	–
B1	0705	0720	15
B2	0705	0730	25
B3	0710	0740	30
B4	0710	0750	40
A3	0720	0800	40
C1	0720	0810	50
B5	0740	0820	40
C2	0740	0830	50
A4	0820	0840	20
B6	0840	0850	10
Total A			60
Total B			160
Total C			100
Total			320

TABLE 13.4 Revision of the GDP of Table 13.2, if Airline A Fails to Inform the ATFM System of the Cancellation of Flight A2 in a Timely Fashion

Another important practical consideration when planning GDPs is that airlines may desire to change the order of the CTAs that were assigned to their own flights. Consider, for instance, flights B1, B2, B3, and B4 in Table 13.2, all scheduled originally to arrive at XYZ within the 5-minute interval from 0705 to 0710. Because of the sequencing of the respective ETAs, ATFM has assigned to flight B1 a delay of 15 minutes and to flight B4 a delay of 40 minutes. It may, however, be far more important to airline B to have flight B4 arrive at XYZ closer to its scheduled ETA than flight B1. For example, flight B4 may be bringing to XYZ many passengers who will be connecting to other flights of airline B or pilot crews who will operate subsequent flights from XYZ. Note that, because B4's ETA of 0710 is earlier than B1's CTA of 0720, it is perfectly feasible to swap the assigned CTAs of flights B1 and B4 in Table 13.2, so that B4 will suffer a delay of only 10 minutes and B1 of 45. This may be much more palatable to airline B, under the circumstances described, and will result in the same total number of delay minutes as before for airline B and for all the other airlines.

13.4 Collaborative Decision-Making

Example 13.1 points to a number of ways in which the GDP planning process, as practiced until the late 1990s, could be improved. At the most obvious, a very fast, two-way communications environment was required so that the ATFM system and aircraft operators could exchange information efficiently. Second, and more subtly, a partial decentralization of decision-making authority might be beneficial. The FAA (or, more generally, any ANSP operating an ATFM system) does not have the information necessary to make certain decisions on behalf of the aircraft operators. The example of exchanging the CTAs of flights B1 and B4 is a case in point. Airline B, not the FAA, is the entity qualified to decide whether this feasible exchange is worth making. Third, if decision-making were to be partially decentralized in this way, it would be essential to have all participants in the decision-making process operate from a shared knowledge base: they should all have a "common picture" of the current situation at all times, so they can take into consideration everyone else's actions when making their own decisions. Fourth, the process of developing a GDP schedule must be concluded quickly, despite some of the complicated decisions and extensive information exchanges that must take place. This calls for the availability of (preferably common) computer-based decision-support tools to facilitate the interventions of the participants. Finally, there are many instances in GDPs where, for competitive reasons, an airline may prefer to withhold certain information from the FAA and other airlines for as long as possible. The GDP planning process should therefore offer aircraft operators incentives for sharing information.

The CDM approach aims at addressing all these requirements and marks a fundamental change in ATFM's operating philosophy. The basic premise is that "shared information and collaboration in planning and executing ATFM initiatives benefits all ATM users as well as the ATM service provider" (Metron, 2000). CDM's specific stated goals (Metron, 2000) are to:

- Provide the FAA and the airlines with a common picture of current and predicted air traffic conditions by having them look at the same data
- Allow each decision to be made by the person or organization in the best position to make it
- Make these decisions openly so that all know what is happening and can contribute as necessary or desired

"Prototype" GDPs using CDM began in January 1998 on an experimental basis at San Francisco/International and New York/Newark. The experiments were judged so successful that by the end of that year all GDPs at all airports in the United States were being conducted via CDM. An integrated environment for collaborative ATFM with the active participation of every airline of significant size in the United States is now in place. Its scope goes well beyond GDPs. The communications infrastructure for this collaborative system is an Internet-like network called the *CDMnet*, which gives CDM participants the capability for two-way exchanges of real-time information.

Example 13.2 Table 13.2 is the starting point for illustrating some features of GDPs in the CDM environment. It will be assumed again that airline A has decided to cancel flight A2 and that airline B has assigned priorities among flights B1, B2, B3, and B4 in the order of B4, B2, B3, B1, top to bottom.

Under CDM, the proposed CTA schedule shown in Table 13.2 is sent to the airlines via the CDMnet in the form of a "GDP advisory." Airlines and other ATM users may reschedule, substitute, cancel, or delay flights and send back to ATFM their proposed revisions before a known cutoff time. Demand may be reduced sufficiently in this way to eliminate the need for a GDP or make it possible to delay the start of the GDP. This happens quite often and is one of the benefits that CDM offers.

With regard to substitutions and cancellations, the airlines and the FAA have adopted, in connection to CDM, two GDP operating rules that are particularly relevant to this example:[18]

1. Each airline may freely substitute flights within the set of its own flight slots and may move any flight to any one of its slots, as long as that slot is not earlier than that flight's ETA.
2. An airline that cancels a flight has the right to advance its later flights to the first feasible slot that becomes available as a result of the cancellation; this is known as the "slot credit substitution" rule.

Under the first of the rules, airline B is free to assign its six slots (at 0720, 0730, 0740, 0750, 0820, and 0850) to its flights B1 to B6 in any way it wishes, as long as no flight is assigned to a slot earlier than that flight's ETA. This means that airline B may now assign flight B4 to B's first slot at 0720 (since B4's ETA is at 0710) consistently with the flight priorities it has determined. Similarly, airline B will assign flights B2, B3, and B1 to the 0730, 0740, and 0750 slots, respectively, reflecting its preferences.

Consider now the second of the operating rules above. If airline A cancels flight A2 and if B rearranges the CTA of B1 to B4 in the manner just described, the situation shown in Table 13.5 will be obtained. Under the second rule, airline A has priority for utilizing the 0710 slot vacated by the cancellation of A2. In this case, however, the ETA of A3, the first flight of airline A after 0710, is 0720. This means that A3 cannot be moved to the 0710 slot. Under the CDM procedures, flight B4, the next eligible flight, will then be moved to the 0710 slot. (Note that B4's ETA is 0710.) This vacates the 0720 slot, which is a feasible one for A3 whose ETA is also 0720. Thus, A3 will now "leapfrog" over B2, B3, and B1 to occupy the 0720 slot. By being given the first slot that becomes available at or after A3's ETA, airline A is, in essence, "rewarded" for returning to the pool the slot from the cancellation of A2. These rearrangements will lead to the situation shown in Table 13.6. Note that the empty slot has now moved to 0800, the spot vacated by flight A3.

Airline	ETA	CTA	Delay Minutes
A1	0700	0700	0
VOID	–	0710	–
B4	0710	0720	10
B2	0705	0730	25
B3	0710	0740	30
B1	0705	0750	45
A3	0720	0800	40
C1	0720	0810	50
B5	0740	0820	40
C2	0740	0830	50
A4	0820	0840	20
B6	0840	0850	10
Total A			60
Total B			160
Total C			100
Total			320

TABLE 13.5 Revision of the GDP of Table 13.2, if Airline A Cancels Flight A2 and B Rearranges the Order of B1, B2, B3, and B4

Airline	ETA	CTA	Delay Minutes
A1	0700	0700	0
B4	0710	0710	0
A3	0720	0720	0
B2	0705	0730	25
B3	0710	0740	30
B1	0705	0750	45
VOID	–	0800	–
C1	0720	0810	50
B5	0740	0820	40
C2	0740	0830	50
A4	0820	0840	20
B6	0840	0850	10
Total A			20
Total B			150
Total C			100
Total			270

TABLE 13.6 Revision of the GDP of Table 13.2, After Airline A Cancels Flight A2 and Some Flights Are Moved to Fill the Vacant Slot

Following a similar line of reasoning, the GDP shown in Table 13.7 is finally obtained. C1 will now be moved up to occupy the 0800 slot, B5 will occupy the 0810 slot, and A4 will leapfrog over C2 to occupy the 0820 slot. This is the earliest slot that A4 can occupy without violating its ETA. Airline A (and flight A4) became eligible for preferential treatment when A3 vacated the 0800 slot and moved up to occupy the one at 0720.

Airline	ETA	CTA	Delay Minutes
A1	0700	0700	0
B4	0710	0710	0
A3	0720	0720	0
B2	0705	0730	25
B3	0710	0740	30
B1	0705	0750	45
C1	0720	0800	40
B5	0740	0810	30
A4	0820	0820	0
C2	0740	0830	50
B6	0840	0840	0
Total A			0
Total B			130
Total C			90
Total			220

TABLE 13.7 The Final GDP for Example 13.2

Compare now Table 13.7 with Table 13.3, which was obtained under the assumption that airline A informed ATFM of the cancellation of flight A2 in a timely fashion. Both GDPs have resulted in exactly the same total number of delay minutes, 220 (saving 110 minutes from the original GDP of Table 13.1), but with very different distributions of the resulting benefits. The main beneficiary, by far, in Table 13.7, is airline A, whose three remaining flights now suffer no delay. Airline A is now motivated to report promptly the cancellation of A2 so it can take advantage of the slot credit substitution rule. Ball et al. (1998) have reported that before CDM, airlines informed the ATCSCC of the cancellation of flights about 50 minutes, on average, *after* the scheduled time of departure of the cancelled flights! Under CDM, this time became 45 minutes before the scheduled departure time, a difference of more than 1.5 hours. This means a huge improvement in the quality of information that CDM participants work with when planning GDPs and other ATFM interventions.

Note, as well, that in the final GDP schedule of Table 13.7, airline B receives benefits, which cannot be measured in terms of delay savings, by having its flights B4 and B1 exchange positions, per its priorities. This illustrates the point that some of the (very real) economic benefits of CDM can be quantified only by the airlines themselves.

Airline Operations Centers (AOC) make available via CDMnet information regarding changes to flight arrival and departure schedules, assignments of flights to airport arrival slots, flight cancellations, and newly created flights. This information is used by the ATCSCC to revise ongoing GDPs and to determine whether any capacity/demand imbalances exist that warrant additional ATFM intervention. The demand information is consolidated approximately every 5 minutes and returned to aircraft operators as an “aggregate demand list" (ADL). The ADL allows airlines to see where their flights fit in the traffic flow and to plan accordingly.

The *flight schedule monitor* (FSM) is the software that provides the shared knowledge base of current and predicted conditions for all CDM participants and makes possible user collaboration in GDP decision-making. The FAA and the AOCs use it to implement and manage all GDPs. FSM provides a graphical presentation of airport demand and capacity information, as received through CDMnet, displaying flight-specific information, airport arrival and departure rates, open arrival slots, and other traffic flow information. FSM also provides many of the requisite decision-support tools. It contains a set of computer algorithms and utility programs to support GDP management and analysis so that users can react quickly to airport and airspace capacity constraints. FSM users can also test alternative ATFM scenarios involving flight cancellations, delays, or substitutions, and observe the results before taking any action on their flights.

Overall, the CDM approach to designing GDPs can be seen from Example 13.2 as consisting of a succession of substitution and compression steps. The latter refers to the process of filling any gaps in the GDP schedule that are created by the cancellation of flights. The steps involved in a GDP under CDM can now be summarized as follows:

Step 1. The ATCSCC obtains daily estimates of the airport acceptance rate (AAR) for each of the airports where capacity may be reduced due to unfavorable weather conditions or other reasons.
Step 2. If delays are projected to be severe at any airport, the ATCSCC prepares to run a GDP for that airport by assigning slots to airlines on a first-scheduled, first-served basis ("ration by schedule," RBS) using the predicted AAR.
Step 3. A GDP advisory is sent to the airlines and other users via CDMnet that includes the planned CTA for all arrivals that will be affected by the planned GDP.
Step 4. Each airline informs the ATCSCC by a cutoff time on how it plans to use its slots, including any substitutions and flight cancellations.
Step 5. After receiving user responses, the ATCSCC performs compression to take advantage of any empty slots and finalize slot assignment to flights. (If flight delays have been reduced sufficiently due to cancellations, the GDP may be cancelled altogether at this point.) The ATCSCC thus computes the final CTA assigned to each flight.
Step 6. By working backward from the CTA, the ATCSCC estimates the controlled time of departure (CTD) at which each flight affected by the GDP will depart from its airport of origin for the GDP airport. It then communicates the CTA and CTD for each flight to the airlines.

The ATCSCC monitors continuously the execution of each GDP and may revise it by repeating the above process in light of any significant changes in the expected AAR, new flight cancellations, etc. It may decide at some point to cancel the GDP earlier than planned, for example, if the weather improves earlier than expected or if air traffic demand is lower than expected; the GDP may also be extended in time, if conditions worsen.

Additional Technical Issues and Extensions

Numerous other issues arise in practice in connection with the management of air traffic flows into airports. As an example of the kind of detailed question one must contend with, consider the example of the "double-penalty problem." Suppose an airline is forced to delay the departure of a flight due to a mechanical problem. Under earlier GDP rules, if the airline informed the FAA of the new departure time, the FAA would compute a revised ETA for that flight. If a GDP were then initiated for the airport of destination, the flight in question would receive a CTA based on the revised ETA, that is, suffer additional delay on top of the delay due to the mechanical problem. Faced with the prospect of this double penalty, airlines refrained from reporting such mechanical delays to the FAA. GDP slots were wasted as a result, as late-departing aircraft failed to meet the CTA assigned to them. Under CDM, the rationing of slots is now based on the original ETA, not a revised one. Thus, an airline can truthfully report a mechanical delay, knowing that its flight will be assigned the slot it was originally entitled to. Through slot credit substitution, it will be able to use the earliest slot that the flight can make after departing late due to the mechanical delay. Similar special-purpose policies have been developed for several other problems, such as accommodating, without further delay, flights that were previously diverted to other airports due to weather.

However, some ATFM problems of a more technical nature are particularly difficult to solve. Certainly the most fundamental is the setting of the AAR, which is typically estimated several hours in advance of the starting time of the GDP, based primarily on weather forecasts that are highly uncertain in many cases. Think, for example, of the difficulty in predicting several hours in advance when exactly (i.e., with a tolerance of about 15 minutes) heavy fog will roll in or burn off at an airport, or when a line of thunderstorms will arrive near an airport, when it will move away, and how severely it will affect the airport's operations. Since CTA schedules and ground delay assignments are all based on the AAR, an incorrect prediction of the AAR will lead to one of two types of errors. If the estimate is too high, that is, if it turns out that the airport is not able to accept as many arriving aircraft per hour as predicted, then these aircraft will suffer additional, possibly long, airborne delays, on top of the ground delays already assigned to them. If, on the other hand, the estimate is too low, aircraft will be held for an excessive amount of time on the ground and suffer unnecessary delay. Note that in this second case valuable airport capacity will be wasted, as the rate of arrivals at the airport (restricted to 6 per hour in Example 10.1) will be lower than what the airport could accommodate. This is sometimes referred to as "starving the runways." Airlines often complain that this type of error is all too common.

They argue that ATFM tends to adopt worst-case scenarios regarding the AAR, that is, is by nature biased toward low-side estimates of capacity, because its primary concern is avoiding overloads.

One possible response to this type of problem is the managed arrival reservoir (MAR). Under this approach, the ATFM system plans for some amount of airborne delay when determining the controlled time of departure (CTD) of an aircraft during a GDP. For example, with reference to Table 13.7, flight C1 could be given a CTD from its airport of origin, which is only 25 minutes later than its scheduled departure time, not 40 minutes. If everything goes according to schedule, 25 of the planned 40 minutes of delay will then be taken on the ground prior to takeoff and 15 minutes in the air.[19] The intent in this case would be to create an airborne queue near the GDP airport with arriving aircraft having to wait roughly 15 minutes before landing, if the predicted AAR proved to be correct. If, however, the true AAR turned out to be higher than the predicted one, the 15-minute queue would provide a "reservoir" of aircraft, which would be available to utilize the additional available slots and avoid wasting airport capacity. Note that if the true AAR proved lower than predicted, the MAR would add further to the resulting unplanned airborne delays. A more genuine and better-performing solution than MAR to the critical problem of the uncertainty associated with the AAR can come only from a combination of improved weather forecasting technology and more advanced methodologies for setting the AAR based on stochastic optimization (Ball et al., 2007; Vossen et al., 2012).

The CDM environment offers the potential for developing other important extensions or new applications of ATFM. Two examples can be described briefly. One is the concept of collaborative routing (CR), which involves coordinating through CDM the rerouting of aircraft in en route airspace, whenever it appears likely that an en route sector will be overloaded with traffic or whenever a weather front might necessitate such action. CR requires (1) availability of timely data and of reliable methods for predicting delays en route, so an airline can decide which alternative routes between two points are, *a priori*, the most attractive, as well as (2) a procedure and set of rules for allocating in a fair and efficient manner the alternative routes requested by the airlines. Indeed, an airspace flow program (AFP) has been operational since 2006 extending GDP CDM procedures to the en route environment. A major concern in CR is the balancing of resulting flows, so that rerouting does not create overloads and congestion on the alternate routes.

A second important example is an extension that still has to be fully addressed. It is motivated by the observation that GDPs are currently limited to allocating arrival capacity only.[20] If the allocation also included departures, while taking into consideration the tradeoff between the arrival capacity and the departure capacity of an airport (see Chap. 10) this would provide a great deal more flexibility and could be most helpful in increasing the reliability of flight connections at airline hubs (Hall, 1999).

13.5 Near- and Medium-Term Enhancements

Efforts to modernize ATM systems are under way all over the world. By far the two largest and most ambitious among them are NextGen in the United States (FAA, 2011) and SESAR (Single European Sky ATM Research) in Europe. While both of these programs have been under way for several years, they are still evolving, with frequent changes in their content and boundaries. This is not surprising in view of their complexity and broad scope. Some of the elements that the two programs comprise (e.g., ADS-B and ASDE-X, see Sec. 13.2) have existed since the 1990s, while others are entirely new and still in the conceptual stage.

The central themes of NextGen and SESAR are similar, despite some differences in their approaches and selection of technologies. They emphasize four fundamental types of changes:

1. Much greater reliance on satellite-based technologies for navigation, communications, and (through position broadcasting) surveillance
2. Planning of flights based on four-dimensional (4D) trajectories
3. System-wide information management (SWIM)
4. Increased automation of routine ATC tasks, allowing humans to concentrate on "high-value" interventions

Progress toward these goals has been uneven due to many challenges. First, both NextGen and SESAR are enormously complex undertakings. For example, SESAR, which is operated as a public-private partnership led by EUROCONTROL, had several hundreds of parallel projects in progress as of 2012, with each project interacting, on average, with five or six others. Second, both programs are expensive and must rely heavily on public funding. In the case of NextGen, such funding becomes available on essentially a year-to-year basis, as part of the annual FAA budget authorization cycle. Funding is therefore subject to delays and uncertainty. Third, some of the modernization initiatives have environmental implications, such as shifting noise impacts from one community to another. This may require lengthy environmental impact studies and approval processes. Finally, and perhaps most important, airlines and aircraft operators have been slow or reluctant to invest in updating relevant avionics and other equipment. Before doing so, they want to be certain that the various proposed changes will (a) be implemented and (b) produce the promised benefits.

It is useful to examine briefly what the proposed NextGen and SESAR plans mean for major airports. Significant benefits will undoubtedly be obtained in the area of safety, especially when it comes to runway and taxiway incursions, a problem that has become particularly acute during the past 20 years as a consequence of the rapid growth in airport surface traffic. These safety benefits will be driven by the increased use of the ASDE-X surveillance radar, especially when coupled with AMASS (Airport Movement Area Safety System), which alerts tower controllers to potentially hazardous situations on the airports surface, and RWSL, the Runway Status Lights system, that provides visual warnings of potential conflicts to pilots preparing to cross active runways and busy taxiways while taxiing.

The ADS-B system, when used in terminal airspace, may also provide safety benefits by increasing the awareness of pilots equipped with a CDTI of the position of other (ADS-B-capable) aircraft in their vicinity. Similarly, ADS-B could be helpful in surface traffic surveillance in conditions of poor visibility. By making it possible to monitor the positions of aircraft and ground vehicles, ADS could reduce the likelihood of runway and taxiway incursions, as well as assist in reducing taxiing times. This would lessen dependence on the expensive ASDE radars.

Ever-improving weather systems can also contribute substantially to airport safety. For example, the integrated terminal weather system (ITWS) in the United States is an automated weather system that provides short-term (0–60 minutes) predictions of significant terminal-area weather at major airports. ITWS provides predictions or information about windshear and microbursts, storm-cell hazards, lightning, and strong winds.

ATM modernization may also alleviate some of the environmental impacts of airport traffic. Automation and decision-support systems are increasingly being applied to optimize the handling of departures and surface traffic. An important observation in this respect is that, for each airport and for each runway configuration, one can determine an optimal number of departing aircraft to be allowed onto the taxiway system, such that, on one hand, the runways can operate at their full capacity, while at the same time congestion, fuel burn and engine emissions from taxiing aircraft are kept to a minimum (Pujet, 1999). Two independent experiments performed at Boston/Logan and New York/Kennedy in 2010 and 2011 demonstrated the benefits this approach can generate (Simaiakis, 2012). A parallel ambitious effort with regard to airport surface traffic is the Advanced Surface Movement Guidance and Control System (A-SMGCS) in Europe. This provides a real-time capability for planning and implementing traffic flows on every part of an airport's surface.

Another technology-based procedure that may significantly affect noise impacts at airports, as well as reduce fuel consumption and emissions is the Continuous Descent Approach (CDA)—also known as Optimized Profile Descent (OPD). In this procedure aircraft approach runways at idle thrust while maintaining a constant descent angle, typically of 3°, until they intercept the instrument landing system (ILS). Ideally, aircraft can descend in this way from cruising altitude to the runway, without having to fly horizontal segments and apply throttling to meet the ILS (see Sec. 13.2). Several airports in Europe have already standardized such OPDs.

Finally, turning to the question of increasing capacity and reducing airport delays, most near-term benefits will be modest and driven as much by decision support tools as by improved technologies. A prominent example in this respect is a set of software tools that support the spacing and sequencing of arrivals and departures on runways. The best known among those is a comprehensive system called the Center TRACON Automation System (CTAS), which NASA developed and is operating at a growing number of airports in the United States, beginning with Denver/International in the 1990s. CTAS consists of three key modules that assist controllers in performing most of the tasks essential to (approximately) the last 40 minutes of flight (Erzberger, 1995). The Traffic Management Advisor (TMA) facilitates the planning of the sequence of arrivals and their allocation to active runways while aircraft are still in the en route phase of flight. The Descent Advisor (DA) is used in developing near-optimal four-dimensional paths from start of descent in an en route sector to delivery at the approach feeder fixes (see Sec. 13.2) in the terminal area. The DA provides recommendations concerning the point where descent should be initiated and the descent rate, path, and speed. Its objective is to deliver each aircraft at the feeder fix at a specified time, while minimizing fuel consumption. Finally, the Final Approach Spacing Tool (FAST)[21] aims at maximizing runway acceptance rates and balancing traffic on multiple runways. It assists arrival controllers in the TRACON (see Sec. 13.2) in determining loads on the runways, generating a desirable sequence of landings on each runway, and achieving tight spacing between them. The latter is done through speed, heading, and turn advisories to aircraft.

One of the features of CTAS is the ability to deviate from first come, first served (FCFS) sequencing of arriving aircraft in order to increase capacity and reduce delay. This is achieved by moving an aircraft by up to a specified maximum number of positions from its FCFS order. For example, if a particular airplane is 12th in line for landing according to FCFS and if the maximum number of position shifts is two, then that airplane can be the 10th, 11th, 12th, 13th, or 14th to land. This is called *constrained position shifting* (CPS). It has been shown that, with a maximum position shift of two or three, CPS can reduce significantly delays by avoiding the most undesirable landing sequences, such as a “heavy” aircraft followed by a “small” one and requiring 6 nmi of longitudinal separation. This issue was also discussed in Sec. 10.5. Note that when the maximum position shift is as small as one, two, or three, CPS also maintains a sense of fairness by guaranteeing that no aircraft will be given a position in the landing sequence that differs significantly from the FCFS order (Dear and Sherif, 1991; Balakrishnan and Chandran, 2010).

The CDM approach (Sec. 11.4) marks a turning point in the way ATM systems have been operated. It introduces an entirely new philosophy to the allocation of decision-making responsibilities between ANSPs and aircraft operators. Such an approach can contribute to increasing airport capacity or, at the very least, reducing the cost, if not the size, of airport delays. An example is the development of "airport CDM" procedures at a number of major European airports (e.g., Paris/de Gaulle, Munich/International) that have successfully brought together users, airport operators, and ANSPs to increase the efficiency and predictability of airport operations.

Technology-based innovations may also yield some capacity gains. With ADS-B, for example, pilots can obtain, even in instrument meteorological conditions, visual information through the CDTI about the position of other aircraft in their vicinity. More accurate spacing between aircraft near the airport and on final approach might be achieved, in much the same way as when pilots separate themselves visually in VMC. As a second example, improved weather systems may provide significant capacity benefits by helping predict more accurately the time when near-airport thunderstorms or other weather events will end. Allan et al. (2001) present an interesting case study in this respect, based on an ITWS demonstration program at the New York City airports.

While promising and useful, the capacity enhancements and delay reductions that were outlined above will initially be limited. Major gains can only come through such more dramatic changes as, for example: reduced separation requirements for landings and/or takeoffs on the same runway; better wake-vortex detection and avoidance on final approach; and reduced separation requirements for independent operations on different runways. Both NextGen and SESAR include steps of this kind among their objectives. For example, one of the objectives of NextGen is to eventually permit independent parallel approaches in IMC to parallel runways separated by as little as 1200 ft. It is safe, however, to expect that such targets will not be reached in the immediate future.

Exercises

13.1. One of the most interesting ideas under consideration in terminal area air traffic control involves the "sequencing" of aircraft on arrival in order to achieve certain benefits. In this problem we examine the effects of various sequencing schemes in a terminal area. We consider for this purpose the terminal airspace around an airport at which a single runway is used exclusively for landings. Aircraft arrive at this terminal area at random times, are sequenced by air traffic controllers, and land at the runway. We assume the following somewhat simplified conditions:

(i) All aircraft fly a 5-nmi final approach.

(ii) The minimum airborne longitudinal separation between successive landing aircraft is 3 nmi for all possible pairs of aircraft, except behind aircraft approaching at

150 knots. In this latter case, 4 nmi are required, regardless of the type of the second aircraft in the sequence.

(iii) No buffers are added to the minimum separations between aircraft. (Note that in computing the minimum separations between successive landing aircraft, the runway capacity model presented in Sec. 10.5 is used.)

Assume now that, just before t = 0, there are no aircraft in the terminal area or on the runway and that, beginning at t = 0, a sequence of six aircraft enter the terminal area according to the data shown in Table 13.8. (Note that the "nominal arrival time at the runway" indicates when an aircraft would reach the runway in the absence of any other traffic.)

Aircraft Identification Number	Terminal Area Entrance Time* (seconds)	Approach Speed (knots)	Terminal Area Transit Time† (seconds)	Nominal Arrival Time at Runway (seconds)
1	0	120	990	990
2	20	135	930	950
3	55	150	900	955
4	110	120	990	1100
5	180	135	930	1110
6	350	150	900	1250

*Time when the aircraft enters the terminal area.

†Time it would take an aircraft to travel through the terminal area and reach the runway threshold in the absence of any other traffic.

TABLE 13.8 Data for Exercise 1

Consider that aircraft are sequenced according to one of the following four different strategies:

- *Strategy 1*. Their time of entrance into the terminal area, that is, first come, first served according to entrance time. (This means that aircraft would land at the runway in the sequence 1–2–3–4–5–6, spaced apart by (at least) the required ATC separations between them.)
- *Strategy 2*. Their nominal arrival time at the runway, that is, first come, first served according to nominal arrival time at the runway. (This means a 2–3–1–4–5–6 order.)

- *Strategy 3*. The sequence that minimizes total delay, that is, the sum of the differences between the actual time when each aircraft is assigned to reach the runway and the "nominal arrival time at the runway" for that aircraft.
- *Strategy 4*. The sequence that maximizes "throughput," that is, so as to land the last aircraft to reach the runway as soon as possible (this is equivalent to maximizing the "flow rate" of aircraft onto the runway).

Part a: For each of the four strategies described previously, compute:

(i) The sequence in which the six aircraft in our example should land (You have already been given the answer for strategies 1 and 2.)
(ii) The total delay corresponding to this sequence
(iii) The time when the last of the six aircraft in the sequence would reach the runway

Part b: Suppose now that an aircraft cannot be re-sequenced by more than one position from the order in which it entered the terminal area. For example, aircraft 3 can land second, third, or fourth but not first, fifth, or sixth. Repeat part *a* under this restriction for strategies 3 and 4. [The restriction in this part is often referred to as constrained position shifting (CPS).]

Part c: Comment briefly on the advantages and disadvantages of each of the four strategies described previously, as well as on CPS. Why might CPS (with the objectives of strategies 3 or 4) be an attractive idea? Please write a thoughtful few sentences. Think of such items as delay costs, controller workload, airline perceptions, predictability (to the pilot) of when an aircraft will actually land, etc.

Part d: Suppose that, at some particular time, nine aircraft are queued up waiting to land at this runway. How many possible sequences would have to be examined, in the worst case, under strategies 3 and 4, and how many under CPS with a maximum shift of one position? (Think systematically in answering the CPS part.)

13.2. You may be familiar with the traveling salesman problem (TSP), famous in operations research and applied mathematics. One form of the TSP can be stated as follows: Suppose one is given a set of n points on a plane and the Euclidean distances between every pair of these points. The TSP is the problem of finding the shortest-distance tour that begins at one of the points, visits all the other points exactly once, and returns to the starting point. A variation of the TSP is the Hamiltonian path problem, which is identical to the TSP as described but does not require a return to the starting point (i.e., visit all points exactly once, but end at the nth point visited). Suppose that n airborne aircraft are waiting to land on a runway and that air traffic controllers can sequence them in any way they wish. Suppose also that the sequence will be determined so as to implement strategy 4 of Exercise 1, that is, so as to land the last aircraft to reach the runway as soon as possible. Argue that this problem is equivalent to solving the Hamiltonian path problem, but with asymmetric distances between the points (i.e., the distance from *i* to *j* may be different from the distance from *j* to *i*).

13.3. Consider ground delay programs (GDPs). Current practice exempts long-range flights from ground holding. For example, a flight from Frankfurt/International to Boston/Logan, which takes roughly 8 hours, will be allowed to take off on time, even if it is likely (but not certain) that, at the time of the flight's arrival, Boston/Logan's capacity (or AAR—see Secs. 13.3 and 13.4) will be low. This policy means that the brunt of ground holding delays is borne by short- and medium-range flights. Does this policy make sense? What is the rationale for it? What are its advantages and disadvantages?

References

ACI, Airports Council International (2011) *World Airport Traffic Report 2010*, Montreal, Quebec, Canada.

Allan, S., Gaddy, S., and Evans, J. (2001) *Delay Causality and Reduction at the New York City Airports Using Terminal Weather Information Systems*, Project Report ATC-291, MIT Lincoln Laboratory, Lexington, MA.

Balakrishnan, H. and Chandran, B. (2010) "Algorithms for Scheduling Runway Operations under Constrained Position Shifting," *Operations Research, 58*, pp. 1650–1664.

Ball, M., Hall, W., Hoffman, R., and Rifkin, R. (1998) *Collaborative Decision Making in Air Traffic Management: A Preliminary Assessment*, National Center of Excellence in Aviation Operations Research (NEXTOR), University of Maryland, College Park, MD.

Ball, M., Barnhart, C., Nemhauser, G., and Odoni, A. (2007) "Air Transportation: Irregular Operations and Control," in *Handbook in Operations Research and Management Science*, C. Barnhart and G. Laporte (eds.), Vol. 14, Elsevier, Amsterdam, The Netherlands.

Barnett, A. (2009) "Aviation Safety and Security", in *The Global Airline Industry*, P. Belobaba, A. Odoni, and C. Barnhart (eds.), John Wiley & Sons, Chichester, UK.

Braff, R., Powell, J., and Dorfler, J. (1994) "Applications of the GPS to Air Traffic Control," in *Global Positioning System: Theory and Applications: Volume II*, B. W. Parkinson and J. J. Spilker, Jr. (eds.), American Institute of Aeronautics and Astronautics, Washington, DC.

Cook, A. (ed.) (2007) *European Air Traffic Management*, Ashgate Publishing, Hampshire, England.

Dear, R. and Sherif, Y. (1991) "An algorithm for computer-assisted sequencing and scheduling of terminal area operations," *Transportation Research Part A*, 25, pp. 129–139.

Erzberger, H. (1995) *Design Principles and Algorithms for Automated Air Traffic Management*, AGARD Lecture Series 200, www.aviationsystemsdivision.arc.nasa.gov/publications/more/index.shtml#ctas.

EUROCONTROL (2010) *SESAR Master Plan Update*, Brussels, Belgium.

FAA, Federal Aviation Administration (1992a) "Designation of Federal Airways, Area Low Routes, Controlled Airspace, Reporting Points, Jet Routes, and Area High Routes," *Federal Aviation Regulations*, Part 71, FAA, Washington, DC.

FAA, Federal Aviation Administration (1992b) "General Operating and Flight Rules," *Federal Aviation Regulations*, Part 91, FAA, Washington, DC.

FAA, Federal Aviation Administration (2011) *NextGen Implementation Plan 2011*, Washington, DC.

FAA, Federal Aviation Administration (2012) *Airport Design*, Advisory Circular 150/5300-13, U.S. Government Printing Office, Washington, DC.

Gilbert, G. (1973) "Historical Development of the Air Traffic Control System," *IEEE Transactions on Communications, 21*, pp. 364–375.

Hall, W. (1999) *Efficient Capacity Allocation in a Collaborative Air Transportation System*, Ph.D. dissertation, Operations Research Center, Massachusetts Institute of Technology, MIT, Cambridge, MA.

ICAO, International Civil Aviation Organization (2009) *Aerodromes: Annex 14 to the Convention on International Civil Aviation, Volume I: Aerodrome Design and Operations*, 3d ed., ICAO, Montreal, Canada.

Metron, Inc. (2000) *CDM Web site*, http://www.metsci.com/cdm/.

MIT Lincoln Laboratory (1998) *Air Traffic Control Overview: Kansas City ARTCC*, Group 41 Report, MIT Lincoln Laboratory, Lexington, MA.

Mundra, A. (1989) *A Description of Air Traffic Control in the Current Terminal Airspace Environment*, Report MTR-88W00167, The MITRE Corporation, McLean, VA.

Nolan, M. (2011) *Fundamentals of Air Traffic Control*, 5th ed., Cengage Learning, Delmar.

Pujet, N. (1999) *Modeling and Control of the Departure Process at Congested Airports*, Ph.D. dissertation, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA.

Simaiakis, I. (2012) *Analysis, Modeling and Control of the Airport Departure Process*, Ph.D. dissertation, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA.

Vossen, T., Hoffman, R., and Mukherjee, A. (2012) "Air Traffic Flow Management," in *Quantitative Problem Solving Methods in the Airline Industry*, C. Barnhart and B. C.

Smith (eds.), *International Series in Operations Research and Management Science*, 169, Springer New York.

---

[1]Airspace in the United States is classified into Classes A, B, C, D, E, and G. This is consistent with the ICAO classification of airspace, which, however, also includes Class F airspace that does not exist in the United States.

[2]Class G airspace may extend to higher altitudes in remote areas.

[3]"FL" refers to flight level as measured according to constant atmospheric pressure relative to a reference datum of 29.92 in of mercury. Flight levels are expressed in units of hundreds of feet, so FL 600 corresponds to 60,000 ft.

[4]At some airports, a taxiing aircraft may come under the control of more than one ground controller on its way from the apron to the runway or to the apron from the runway.

[5]This type of operation requires close coordination between the relevant control positions in the TRACON.

[6]The ARTS replacement program was delayed considerably by the need to adjust STARS displays and other features in response to extensive controller comments and requests. This is a typical problem with all ATM systems in which human factors and ergonomic considerations play a critical role.

[7]ARSR have a slower revolution rate (about 6 per minute; or 10-second updates) and a range of the order of 250 nmi.

[8]In the United States, SSR systems are often referred to as *air traffic control radar beacon systems* (ATCRBS).

[9]The glide slope angle may range from a minimum of 2° to a maximum of 7°.

[10]By contrast, the longitudinal distance along the final approach path between a slow leading aircraft and a fast trailing one cannot be reduced below a specified separation minimum.

[11]The CFMU replaced five old regional flow management centers operating in Europe.

[12]En route sector capacity in Europe is determined primarily by the day-to-day availability of personnel to staff sector positions, as well as by the directionality and overall configuration of traffic flows. Personnel shortages or other events may at times require merging or partial reconfiguration of sectors.

[13]Additional types of information are constantly being added to the ETMS database; in fact, one of the principal consequences and benefits of the application of CDM to ATFM (Sec. 13.4) has been the inclusion in the ETMS database of large amounts of information provided by the airlines and other CDM participants on a dynamic basis.

[14]Access to ETMS or to certain parts of it is available to airlines and to many other non-FAA organizations. Several of those use ETMS-derived information to support commercial activities. For example, some operators of limousine services at airports use ETMS and ASD to obtain accurate, real-time information on the time of arrival of the flights of their customers.

[15]As of 2012, about three GDP are initiated, on average, every day in the United States.

[16]The number of flights and airport capacity are both unrealistically small in this example, to facilitate the presentation. At the major airports where GDPs are most often applied, the typical number of scheduled arrivals per hour is 50 or more, while the AAR is of the order of 30 (i.e., a slot every 2 minutes) or greater. The example is otherwise typical of what may occur in practice.

[17]Information about the cancellation of A2 is valuable to A's competitors in the market in which A2 operates, as they may be able to attract some passengers originally booked on A2. Equally important, if these competitors were also considering canceling some of their own flights in that market that day, knowledge of the cancellation of A2 might persuade them not to do so.

[18]Although somewhat simplified, these descriptions capture the essence of the two rules.

[19]Note that, under this scheme, a flight that has been assigned a delay of less than 15 minutes by the GDP will not be required to take any ground delay at all.

[20]EUROCONTROL allocates slots to both arrivals and departures; however, it works with a fixed arrival capacity and a fixed departure capacity for each airport and does not consider potential tradeoffs between the two.

[21]The initial version of FAST was referred to as p-FAST (for “passive”) and a later version as a-FAST (for “active”).

PART IV

Landside

CHAPTER 14

Configuration of Passenger Buildings

The selection of the configuration of terminal buildings is a crucial design issue. Inappropriate choices have hurt several major airports and their communities. Bad designs create difficulties for passengers and airlines; reduce the competitiveness of the airport; drive away traffic; and hurt the regional economy. Major examples illustrate this point.

Airport planners and designers thus have great responsibility. They need to choose configurations of terminals that are right for their region. They should carefully consider the operational and economic implications of their designs. The purpose of this chapter is to guide this important process.

14.1 Overview

As a start, we should focus on "passenger buildings." This is the more general term for the range of facilities that serve both passengers and airlines. The more restrictive notion of "terminals" misleadingly suggests that these structures mainly serve travelers who are ending their trips. Complementarily, the concept that these buildings are "gateways" for a region also misleadingly stresses their role as portals for arriving and departing passengers. The "terminal" and "gateway" designations neglect the functions these buildings fulfill for passengers that transfer between flights, who may account for over half the traffic at major airports. This neglect of transfers has been an important source of poor choice of design for landside facilities.

This chapter shows how to evaluate the performance of alternative shapes of airport passenger buildings, at the aggregate level appropriate to overall planning. It also suggests which designs are preferable in which circumstances. A crucial element of this presentation is the recognition that passenger buildings serve a wide variety of different users and functions. The choice of the preferred configuration for passenger buildings should therefore balance the desires of these diverse clients.

Airport passenger buildings serve the many needs of different types of users. They process check-in and baggage for arriving and departing travelers, move transfer passengers between flights, accommodate aircraft on departure and arrival, and provide shopping malls that help finance the airport. They should perform efficiently and profitably for each of their distinct stakeholders. In addition to passengers, these include the airlines managing the aircraft, the owners who provide the capital, and the operators of airport services such as the security and border control agencies.

Airport passenger buildings come in the five basic configurations suggested by Fig. 14.1:

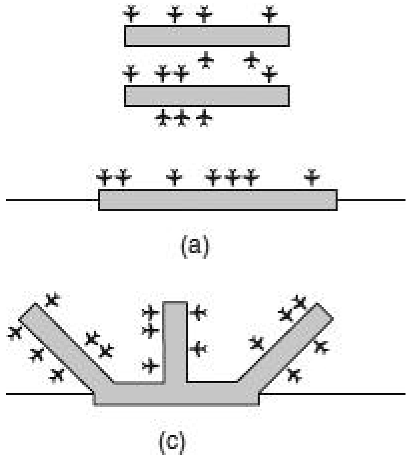


FIGURE 14.1 Sketches of basic configurations of passenger buildings: (*a*) midfield linear, (*b*) midfield X-shaped, (*c*) finger piers, and (*d*) linear.

1. Finger piers
2. Satellites with and without finger piers
3. Midfield, either linear or X-shaped
4. Linear with one airside
5. The transporter design that substitutes some form of bus for buildings as the means to connect passengers and aircraft

At the largest airports, the buildings may be centralized or dispersed into distinct blocks. Each configuration has advantages and disadvantages for different types of traffic. Some passenger buildings are or have evolved to be hybrid, insofar as they combine several of these elements. Indeed, as discussed in the following text, hybrid buildings are generally most desirable.

The desirability of any configuration of airport passenger buildings depends significantly on three factors:

1. The overall level of traffic—a shape that works for a smaller airport may be impractical for larger airports, where distances between aircraft gates may become a dominant consideration
2. The seasonality of the traffic—facilities that will be used only a few months of a year imply high costs per passenger, a factor that should influence their design
3. The percentage of traffic that transfers between aircraft

The different configurations have advantages for different types of traffic. As Sec. 14.5 describes, finger piers can provide good service for originating traffic; linear midfield concourses are efficient for high levels of transfer traffic; transporter solutions are reasonable element of design when the seasonal traffic peaks are more than twice that of the low season.

Which configuration is best depends on the importance of airport stakeholders at any location and over time. As the mix of airport stakeholders can change dramatically, it follows that the desirability of any configuration can vary enormously over time. For example, the original transporter design for Washington/Dulles was suitable for arriving and terminating long-distance passengers, but inadequate and inefficient for the connecting traffic associated with the transfer hub that United Airlines eventually established there. The airport has thus wisely transitioned to a midfield concourse design served by a people-mover system. The long-term performance of any choice of configuration depends on its flexibility to adapt to different types of traffic that may use the airport.

A hybrid configuration that combines different shapes is generally preferable to a design that exclusively selects a single configuration. The different elements of a hybrid design serve the distinct needs of the variety of traffic and airline alliances and provide flexibility for future expansion. In practice, airport operators find that they need to develop specialized facilities over time to serve their range of clients. It is a simple fact that major airports end up with hybrid designs, whatever their original plans may be.

The reality that airports need hybrid designs runs against common assumptions. Master plans for new airports typically depict symmetrical designs with a single configuration. These make pretty pictures, but such designs are unlikely to provide the best service to the range of future airport users. Because as a rule major airports end up with hybrid configurations, it is generally more economical and efficient to start with this approach.

14.2 Importance of Selection

Designers face a crucial issue when they select the configuration of airport passenger buildings. Their decisions about the shapes the buildings have important consequences for the performance and profitability of the airport enterprise. For example,

- Buildings in the form of a T, L, or X use space inefficiently. They waste space by making it impossible to park aircraft along significant portions of the building, in the right angles where the crosspieces meet.
- Arrangements that do not feature common central places are ineffective in generating commercial revenues.
- Buildings that park aircraft only along one side require about twice the length as facilities with aircraft on both sides, and are both uneconomical and inconvenient for transfer passengers.
- Designs requiring aircraft to make numerous turns and stops are expensive for the airlines.

Where designers failed to select appropriate configurations for airport passenger buildings, these bad choices have led to major financial and operational difficulties.

Some of the more obvious bad choices of landside configurations involve buildings that could not process transfer traffic effectively. For example,

- *Kansas City* built an airport with three separate buildings, each designed to park aircraft on one side and automobiles on the other. This might have been an attractive solution if all passengers could pass directly between their car and the aircraft. However, it is a poor configuration for transfers, who have to walk twice as far as they would in buildings designed to serve aircraft on two sides. The Kansas City arrangement was particularly bad because it made transfers walk between separate buildings, in a climate that can be hot and rainy. This bad choice of configuration was one of the reasons that TWA, a major airline that used to base its operations in Kansas City, moved its hub to St. Louis; Kansas City then lost significant traffic, jobs, and economic visibility. This was a regional economic disaster.
- *Frankfurt/International* built a billion-dollar international terminal, primarily to serve the German national airline Lufthansa. Unfortunately, this airline came to realize that it could not workably transfer passengers between this building and its facilities serving domestic and European cities. Lufthansa therefore declined to move to the new building shortly before its inauguration. Smaller airlines, with infrequent schedules and relatively few passengers, then moved into the new building. As a result, this huge capital investment was for a long time underused and largely wasted, while the national airline remained in the crowded old facilities.

Many other issues besides transfer passengers have led to poor choices of the configuration of the airport passenger buildings. Here are some examples:

- *Decentralized building*: Baltimore built a separate international terminal largely to serve US Airways. When that airline moved its international operations to Philadelphia, it left Baltimore with an underutilized building. Meanwhile, the airport needed space for Southwest's expanding domestic service. The logical solution would have been to dedicate the vacated space to Southwest. Yet this was not feasible. Because Baltimore had an inappropriate configuration, it had to spend around $100 million on duplicate facilities. (See ACRP, 2012.)
- *Midfield concourse*: The British Airports Authority originally built London/Stansted with only midfield concourses widely separated from the main building and check-in facilities. Whereas these are convenient for transfer passengers who remain confined to these buildings, they are inconvenient for passengers connecting with ground transport and expensive for the airport operator. They are thus now underused and represent a large waste of resources. Ryanair, now the dominant airline at the airport, insisted on the construction of a finger pier design for its traffic even though capacity was already available at the midfield concourses.

As these examples suggest, airport owners and designers have indeed sometimes made poor choices of the configuration of their passenger buildings. This largely results from systematic failures to recognize the diversity and unpredictability of the functions that passenger buildings should meet.

Standard practice unfortunately encourages designers to base their plans on specific scenarios and fixed forecasts such as the number of million annual passengers, or thousands of passengers per peak hour in major categories. (See, e.g., Ashford et al., 2011; Horonjeff et al., 2010; IATA, 2004; ICAO, 1987.) Standard practice does not consider the possibility, in fact almost the certainty, of changing scenarios and forecasts. The focus on specific futures leads designers to adopt inflexible plans. This in turn causes problems for the airport owners and operators.

14.3 Systems Requirements for Airport Passenger Buildings

This section presents important factors airport owners need to consider before deciding on the configuration of an airport passenger building. It also suggests a general process designers can use for carrying out detailed analyses of the variety of its needs.

A root problem in standard practice is the way architectural programs for passenger buildings focus on a forecast of the number of passengers and aircraft operations. These totals do not describe the variety of distinct types of traffic, the market segments, or the needs of the several stakeholders. Moreover, these forecasts do not consider the range of scenarios under which these proportions and needs may shift over the life of the buildings. The failures to make good choices of the configuration of the passenger buildings thus have often been due to systemic deficiencies in the master planning process as it has been practiced (see Chap. 4).

To avoid poor choices of configuration of airport passenger buildings, airport owners should do the following:

- Understand the current and possible future role of their airport in relation to others internationally, nationally, and in their metropolitan region (see Chap. 5) as applicable
- Anticipate the various requirements of the significant stakeholders in the airport
- Guide designers on how to weight these distinct priorities with respect to each other

In short, owners and designers should consider the long-term risks and variety of future needs, and select a flexible configuration suitable for the plausible evolutions of the airport. They need to take a systems approach to the specification of their requirements. They should plan on a phased development of projects, choosing the first so that it can enable the efficient implementation of subsequent phases adapted to the actual future conditions.

General Considerations

Airport passenger buildings serve the many needs of different users. The buildings serve not only passengers, but also the airlines managing the aircraft, the owners who provide the capital, and the operators of the many services. The configuration of airport buildings is successful to the extent that it reasonably serves the requirements of all these constituencies.

This statement is fundamentally important, although it may appear obvious. It is significant because it defines a starting point that is critically different from where designers typically begin in practice. The reality has been that the design process has all too often ignored major stakeholders in the airport. It inherently claims to take into account the needs of the several stakeholders, such as airlines and retail operators, by setting aside space for them according to generalized industry norms. In most cases, however, the design teams responsible for configuring the airport buildings neither bring the stakeholders into the process nor listen to them. The standard design process sets aside space but generally does not provide adequate functionality.

Three examples illustrate how the design process routinely ignores important stakeholders. They concern the airlines, the stores and other commercial operations in the buildings, and the owners. All are crucial to the success of the enterprise.

Airlines have a considerable stake in the choice of the configuration of the buildings. This is because the arrangement of the buildings affects the time it takes an aircraft to maneuver into position and can easily imply millions of dollars in annual costs. Yet designers typically do not involve airlines in this choice (except when the airlines effectively commission the facilities for their own use and can therefore control designs, as often happens in the United States, see Chap. 3). For example, an international design review for the T5 passenger building at London/Heathrow pointedly did not invite representatives of British Airways, which was to be the major occupant of the facility. As one of the designers said, "first we design the building for the owner, and then we tell the airlines about the space they may have."

Designers normally set aside areas for commercial space, according to the amounts specified in their design program. Not until several years later, shortly before the airport opens the building, does the airport lease these spaces to commercial activities, which then cope with available opportunities as best they can. In Terminal 2 at Tokyo/Narita, the principal shops are on a mezzanine out of the flow of passenger traffic—practically invisible to most passengers. Similarly, stores at San Francisco/International are almost invisible from the International check-in lobby. Experienced designers of shopping malls would never tolerate such poor arrangements. To avoid such problems, the master designers should consult in advance with retail experts and learn how to lay out profitable commercial spaces. Unfortunately, architects and designers rarely undertake such early consultations. Typically, the expert retailers arrive on site when the building is nearly completed.

Investors are concerned with getting good returns on their capital, yet the design process typically operates within a fixed budget limited to the amount available from government grants or bond issues. As cost increases inevitably occur, designers drop many items, even when their extra cost might generate significant returns. For example, cost-reduction schemes eliminated a whole floor of commercial space during the original design of Kuala Lumpur/International. A superior design process will make sure that it does not sacrifice long-term profits for short-term savings.

In fairness, it is not easy for the traditional design process to determine the needs of the important stakeholders in the airport. Getting useful information from them is difficult. For example, airlines rarely have airport-planning departments and are unlikely to provide competent institutional links with the design team. Nobody can speak directly for the operators of retail stores who are unknown during the design process. Government officials running airports may not have the authority to approve deviations from prescribed budgets.

An effective design process will make a special effort to determine and take into account the needs of the important airport stakeholders. The master designers can do this, even though the specific companies and institutions that will be operating at the airport may be unavailable to specify their concerns. They can hire experts on the different issues to help them take a comprehensive approach to the selection of the best configuration of the airport buildings. Although the airlines may not have airport-planning groups, for example, many consultants with airline and airport experience understand the needs of airlines at airports. Master designers can engage appropriately qualified consultants to speak for the interests of the stakeholders in the future facilities. For example, Toronto/Pearson Airport used project funds to pay "airline liaison officers" to represent the airline interests in the construction of its major passenger building.

The rest of this section identifies some major stakeholder issues that a systems approach to the selection of the configuration of airport passenger buildings should address. It focuses on four perspectives: those of the passengers, the airlines, the owners, the commercial services, and government agencies.

Passenger Perspective

The major categories of passengers that deserve special consideration in the design process are the following:

- Domestic or other travelers who are not subject to passport or customs controls
- International travelers requiring government controls
- Transfer travelers who are at the airport simply to transfer between flights
- Business and commercially important travelers, generally more accustomed to travel, often with less baggage, who expect special amenities such as luxury lounges
- Vacationers and personal travelers, often with families and much baggage and using cheap airlines requiring inexpensive facilities
- Disabled, elderly, or other passengers who need level or wheelchair access

Some locations may also serve traffic that systematically requires special treatment, such as pilgrimages. Except for transfers, passengers require ground transportation, check-in facilities, security and other clearances, easy passage between the landside and the aircraft, waiting lounges, commercial services, and baggage delivery. Chapter 16 discusses the detailed design of airport passenger buildings.

Transfer passengers deserve emphasis for two reasons. First, because their needs differ from those of the other travelers. Second, because designers often forget them and the standard references hardly mention them. The crucial difference is that they require fast, reliable, and easy-to-find connections between aircraft. Their connections should be

- *Fast*, because the airlines using the airport for transfer operations need to be competitive with carriers using other hubs. For example, British Airways, serving Boston to Athens through London/Heathrow, competes with Lufthansa operating through Frankfurt/International. Moreover, it is unproductive for airlines to hold their aircraft on the ground; they need to transfer the passengers and get the aircraft flying.
- *Reliable*, because the cost to the airlines of stranded passengers or delayed bags is very high, due both to the direct costs of empty seats and delivery of bags by taxi to the traveler, and to the bad reputation that unreliable service generates.
- *Easy to find*, because complicated routes confuse and delay passengers, and are thus unreliable. Simple, direct routes in a single building are best, as within the United Airlines midfield concourse at Denver/International, or at Amsterdam/Schiphol.

Additionally, transfer passengers obviously do not require check-in facilities, baggage delivery, or easy access to and from ground transportation. Overall, the needs of transfers are distinct from those of the other passengers.

Wherever transfers constitute half or more of the total passengers, the needs of operators and airlines should dominate the choice of the configuration of the airport passenger buildings. This situation occurs at many hub airports, as Table 14.1 indicates. It also occurs at parts of airports for specific dominant airlines. For example, Detroit/Metro built a new passenger building specifically to cater to the transfer activities of Northwest, its then dominant carrier.

Airport	Transfer Rate (%)	Hub Airline
Amsterdam	40	KLM/Air France
Atlanta	60	Delta
Chicago/O'Hare	50	American; United
Dallas/Ft. Worth	60	American
Frankfurt/International	45	Lufthansa
Houston/Bush	55	United
London/Heathrow	35	British
Washington/Dulles	30	United

TABLE 14.1 Approximate Transfer Rates at Major Hub Airports in 2012

Airline Perspective

Airlines care about the configuration of the airport passenger buildings because it affects their operating costs. Poor designs impose heavy costs on the airlines. Good designs give them a competitive edge.

Airlines are sensitive to the costs of maneuvering their aircraft on the ground. They recognize and can be willing to pay for airport configurations that save them time. Significant reductions in the time spent taxiing aircraft can justify hundreds of millions of dollars in new construction. This factor is worth emphasizing, as discussions of the choice of airport configurations typically ignore this important consideration.

Simple calculations illustrate the great value to the airlines of easy ground movements. Recognize first that the direct operating cost of a large commercial jet is of the order of about $100 per minute. At an airport serving 100,000 operations of such aircraft per year (perhaps one with 25 million passengers), a configuration that saves just 3 minutes per operation would give the airlines around $30 million a year in direct benefits alone—the equivalent of about $300 million in capital investment.[1] At larger airports, the savings would be greater. A comparable improvement at London/Heathrow, with over 450,000 operations a year, might be worth the equivalent of about $1 billion in capital cost. Such savings provide a strong rationale for tearing down inefficient configurations and starting all over again, the case of London/Heathrow for its Central Terminal Area.

In this vein, United Airlines benefited greatly from the design of Denver/International Airport; its midfield concourse and completely paved apron dramatically improved the efficiency of its aircraft operations compared to the Denver/Stapleton airport it replaced. The new layout, similar to that of Atlanta shown in Fig. 9.7, features dual taxiways between the midfield passenger buildings that permit straight-in and straight-out maneuvers and reliably reduced the average taxi time per operation. After moving to Denver/International, United was able to tighten its schedules by about 15 minutes on a round trip through Denver. According to their facility manager at the time, this midfield configuration not only reduced direct costs but indirectly boosted aircraft productivity by enabling United to get an extra round trip a day from short-haul flights. Small savings in time may appear insignificant, but, when cumulated over a day, they can have a major impact. As in many aspects of airport systems, small improvements applied to frequent operations can add up to great savings.

Airlines that operate transfer hubs benefit from configurations that facilitate these operations. Regarding transfers, the airline perspective aligns with that of the passengers. Both want fast, reliable, and easy-to-find transfers. Configurations that facilitate these objectives are worth considerable money to the airlines, and they are willing to pay for it. This is why airlines have backed the construction of midfield concourses at Atlanta (Delta), Chicago/O'Hare and Denver/International (United Airlines), and Detroit/Metro (Northwest and now Delta).

Owners' Perspective

Many owners want their airports to be glorious. Airports are major public facilities that can adorn the community. National leaders and governments often aspire to creating monumental gateways, as expressed by a chief architect for the Aéroports de Paris:

> First and foremost, airports are places of great symbolic force…airports are now the place where travelers first come into contact with their destinations: an age-old legacy reminiscent of the "entrance gate" of walled cities…As a gateway in its own right, an airport is almost inevitably destined to be a landmark of great symbolic force, embodying the ambitions of a nation in which it stands. (Andreu, 1997, p. 11)

To further this ambition, airport owners often hold international competitions to choose architects who will celebrate a grand vision. Their results are often spectacular. Examples include the following:

- *London/Stansted*: The British Airports Authority (BAA plc before it was privatized) commissioned Sir Norman Foster to create a translucent glass box unencumbered by air bridges (hence the separate satellites) and design it right down to details of the shape of the check-in counters
- *Osaka/Kansai*: The Japanese Government selected Renzo Piano to design "bird in flight," a continuously changing, 1-km-long glass roof, for its original passenger building on this $13 billion airport on a man-made island
- *Washington/Dulles*: The U.S. government had Eero Saarinen design its original passenger building as a national gateway, uniquely served exclusively by sophisticated transporters

The concept of the airport as a monument conflicts with economic efficiency. Magnificent curved structures (e.g., the Renzo Piano design for Osaka/Kansai) are difficult to construct, expensive to maintain, and nearly impossible to expand compatibly. Custom-tailored interior details (e.g., Sir Norman Foster's for London/Stansted) are correspondingly both expensive and difficult to maintain. These extra costs may be tolerable if the airport owners remain committed to maintaining a monumental concept. However, airport operators and their clients are typically more interested cost-efficient operations.

Many airport operators aim to run their facilities economically. They neither want nor can afford to maintain airports as monuments. Moreover, the owners actually running the airport are not those who commissioned the passenger buildings. The personalities change. The institutions also change and develop new responsibilities. Most remarkably, many of the major airports worldwide are now increasingly privatized and run on strict commercial principles. Thus

- A stockholder company operates London/Stansted and requires all projects to meet stringent financial objectives.
- In 2011 the government of Japan created the Kansai International Airport Corporation to run Osaka/Kansai and Osaka/Itami airports profitably.
- The Metropolitan Washington Airports Authority depends on private capital to construct its expansions and must operate economically to repay its loans.

Airport operators typically emphasize airport economics. In practice, they generally insist on controlling the costs of airport passenger buildings. Because they normally operate within specified budgets, they focus on overall costs. They routinely control costs by scaling back the scope of their projects or cutting out various functions (as Kuala Lumpur did by eliminating a shopping area).

Airport operators expecting to be in business for the long term must consider the entire life of the passenger buildings. They should take into account both the immediate cost of construction and the future costs of expansion. This perspective has two consequences for the evaluation of the configuration of the airport passenger buildings. Airport operators with vision will insist these facilities allow for economical

- Expansion, for example by allowing space for extending facilities, as Paris/de Gaulle did
- Adjustment to different operating conditions, such as a shift in the level of transfer or international traffic, as occurred at Baltimore/Washington

Perceptive airport owners and operators will insist that the configuration of airport passenger buildings enable flexible future use. They will correspondingly reject designs that, while great for specific types of traffic, are difficult to adapt for changed airline operations, types of passengers, or different political situations (that lead to new government regulations for security controls, the separation of arriving and departing passengers, etc.).

Retail Perspective

Retail operators want traffic, persons ready to shop, visibility, access, and a coherent environment. This is what they look for when choosing store locations; this is what good developers of shopping centers provide. When planners choose the configuration of passenger buildings, they can significantly affect the airport's ability to meet these criteria and deliver a successful retail area.

The number of people flowing by a storefront is an immediate measure of the potential attractiveness of a retail location. More people equal more potential customers. Configurations that concentrate traffic create attractive commercial areas. Arrangements that feature central areas are more attractive to stores than buildings that have many entrances and exits. A good example of this is at Amsterdam/Schiphol, where the airport created a busy shopping plaza at the main entrance to its building complex, right above the railroad station.

To be profitable, the traffic must also be ready to shop. Passengers in a hurry or anxious to get somewhere else are not good for retail operators. At Washington/Reagan, for example, the stores between the ground access and the security checks get little business. Travelers typically rush by to get through security and to their gates. People waiting, looking for things to do, and ready to respond to impulsive desires are best for stores. The most important retail areas are thus those that are "past security." Passenger buildings with centralized common waiting areas, such as those at Frankfurt/International or Singapore, provide attractive retail space.

Visibility is also crucial for retail operators. A store not seen is a store not used. The retail area must be able to announce itself. (Of course, airport operators should prevent store signs from hiding important route-finding displays from passengers.) Configurations that lead the pedestrian traffic through shopping streets or past the stores are most successful. The JetBlue building at New York/Kennedy is a good example of successful design in this regard. Conversely, facilities located out of sight, on a floor above or behind the pedestrian flow, will not get many customers.

Retail operators are also concerned with accessibility for their goods. To function economically, they must be able to deliver merchandise easily through security to their stores and customers. Access is more a matter of detailed design than of configuration, but poor arrangements can constrain retail operations. Planners need to verify access to shops when they choose the overall design.

The important point is that designers should incorporate the perspective of retail operators from the beginning. An airport needs a coherent overall plan for organizing its commercial space. Good retail operations, producing the best revenues for the airport, result from good locations. The choice of the configuration of the passenger buildings is an important factor determining the success of retail operations. Master planners should work with knowledgeable retail consultants when designing the configuration of airport passenger buildings.

Government Agencies

Government agencies constitute a particular set of stakeholders that designers need to consult carefully. They deserve special attention, both for their specific needs and because their procedures and modes of negotiation differ from those of other stakeholders.

Government agencies have a particularly strong stake in airports that are international ports of entry and provide border control and customs services. Border control agencies impose a broad range of stringent requirements on the airport. They typically require tight controls on the flows of international passengers and on the mingling of international flights and crews with local employees and services. These requirements greatly complicate the design of the building and may preclude efficient arrangements that airport operators can use for domestic services.

Negotiations with government agencies can be very different from dealing with airlines, travelers, retailers, and other stakeholders. Government security groups are typically inflexible. Officials rarely have the authority to alter established rules, even when the proposed changes might be in everyone's interests. Most especially, they are rarely amenable to arguments based on economic efficiency, in sharp contrast to stakeholders with commercial interests, who are ready to listen to proposals that may save them money or increase their efficiency. Listening to the interests of government stakeholders is most important for airport developers.

Balance

Airport operators need to balance their own strategy and the interests of their stakeholders. They first need to be clear about their own objectives. These differ from airport to airport, as Chap. 3 describes. Are the owners managing the airport as a public service or as a profit-making venture? Are they catering primarily to local customers, or have the vocation to provide a transfer hub? Do they see their primary business as serving travelers and shippers, or developing stores and businesses on and around the airport? Their strategic decisions along these lines should influence how they balance the many conflicting interests of the stakeholders in the airport.

To preserve their own interests, airport operators must also carefully weigh the demands of the several stakeholders. Major clients, such as a dominant airline at an airport, would gladly advance their own interests to the harm of other users. They might want to have preferential access and to discourage competition. Airport operators, however, need to maintain control of their properties, to mediate appropriately between the distinct needs and desires of their stakeholders. It is essential that they consult and listen to their range of stakeholders.

Because forecasts are "always wrong" (see Chap. 4), the design process should also consider the performance of passenger buildings under multiple scenarios. Modern design seeking the best overall performance over the long term thus needs to take a much broader perspective than has been traditional (Table 14.2).

	Criteria Considered	
Forecast	**Single**	**Multiple**
Single	Traditional approach Focus on "terminals"	
Broad		Best practice approach Focus on "airport passenger buildings"

Table 14.2 Best Practice Design Evaluates the Performance of Passenger Buildings Considering Multiple Criteria and Broad Forecasts

14.3 Five Basic Configurations

Designers of airport passenger buildings face a fundamental problem: They need the buildings to be both concentrated and spread out. They need to bring passengers into common areas to facilitate check-in procedures, retail opportunities, and access to public transport. They must also spread out the passengers so that they can board their aircraft. The large wingspan of aircraft imposes long separations between adjacent gates. The lateral distance between gates must be in the range of 50 to 85 m, allowing for clearance between aircraft (Table 14.3). All configurations of passenger buildings represent approaches to resolve this fundamental dilemma.

Airbus	Wingspan	Boeing	Wingspan
A380	80	B747	64
A340-600	63	B777-300	63
A330	59	B767	51
A321	33	B737	35

Table 14.3 Representative Aircraft Wingspans in Meters

The possibilities for resolving this conflict changed greatly with the development of cost-effective "people movers" (see Chap. 17). These devices are small trains or horizontal elevators. They speed people away from a central point, such as a check-in hall, to buildings spread out over the airport. They make it practical to locate passenger buildings over several kilometers and have led to the widespread implementation of midfield concourses at major airports. This technical innovation fundamentally changed the possibilities for effective design of airport passenger buildings.

This section describes the basic configurations of airport passenger buildings from a functional point of view. The subsequent sections show how to analyze some of their essential elements of performance, and then summarize the overall attractiveness of these buildings. The underlying questions throughout are whether, and to what extent, these facilities fulfill the functional requirements of the several stakeholders.

There are five basic configurations of airport passenger buildings suitable for a major airport. (For minor airports, needing only four gates for example, the passenger building can be a simple box.) Designers shape these possibilities in many ways. They also combine these forms into hybrid configurations incorporating two or more distinct forms. As Figs. 14.1 and 9.20 illustrate, the basic configurations are the following:

- Finger piers
- Satellites, with or without finger piers
- Midfield, either linear or X-shaped
- Linear, with only one side devoted to aircraft
- Transporters

At large airports, the buildings may be centralized or dispersed.

Finger Piers

A finger pier is a relatively narrow extension to a central passenger facility. In plan view as seen from the air, finger piers resemble fingers attached to the palm of a hand—hence the name. This design places aircraft gates on both sides of the building extending away from the central core. A finger pier has the advantage of placing some aircraft gates close to the central facility, thus making them more convenient for the passengers than the gates at the end.

An alternative arrangement, known as a "hammerhead," widens the end of the finger pier so that it looks like a T in plan view. The end of this pier serves a number of aircraft around a small central core (located in the crosspiece of the T). This concentration of passengers in a single space enables shared use of facilities and decreases the space required for lounges by 30 percent or more (see Chap. 16). The number of passengers using the range of gates also increases the attractiveness and profitability of retail opportunities. A principal disadvantage of this plan, however, is that it places many aircraft and passengers far from the central part of the main passenger building and forces passengers to walk farther.

Designers introduced the use of finger piers in the 1950s as the first response to the need to serve dozens of gates from a central check-in hall. For several decades, finger piers constituted the standard configuration. Airports worldwide built passenger buildings with finger piers that are still used—for example, Chicago/O'Hare, Frankfurt/International, New York/LaGuardia, San Francisco/International, and others.

The difficulty with finger pier configurations is that, at large airports with many gates, they lead to long walking distances for passengers. To avoid long hikes, airport designers no longer propose finger piers as extensively. Instead, they prefer when possible to replace the long fingers with people movers that serve independent buildings such as satellites or midfield concourses, discussed subsequently.

Many airports continue to implement finger piers in some fashion. However, designers now minimize walking distances either by designing short finger piers (as at Washington/Reagan) or by incorporating a people mover (as at Osaka/Kansai and Detroit/Metro). Figure 14.2 shows a version of this design for the Nagoya/Chubu airport.

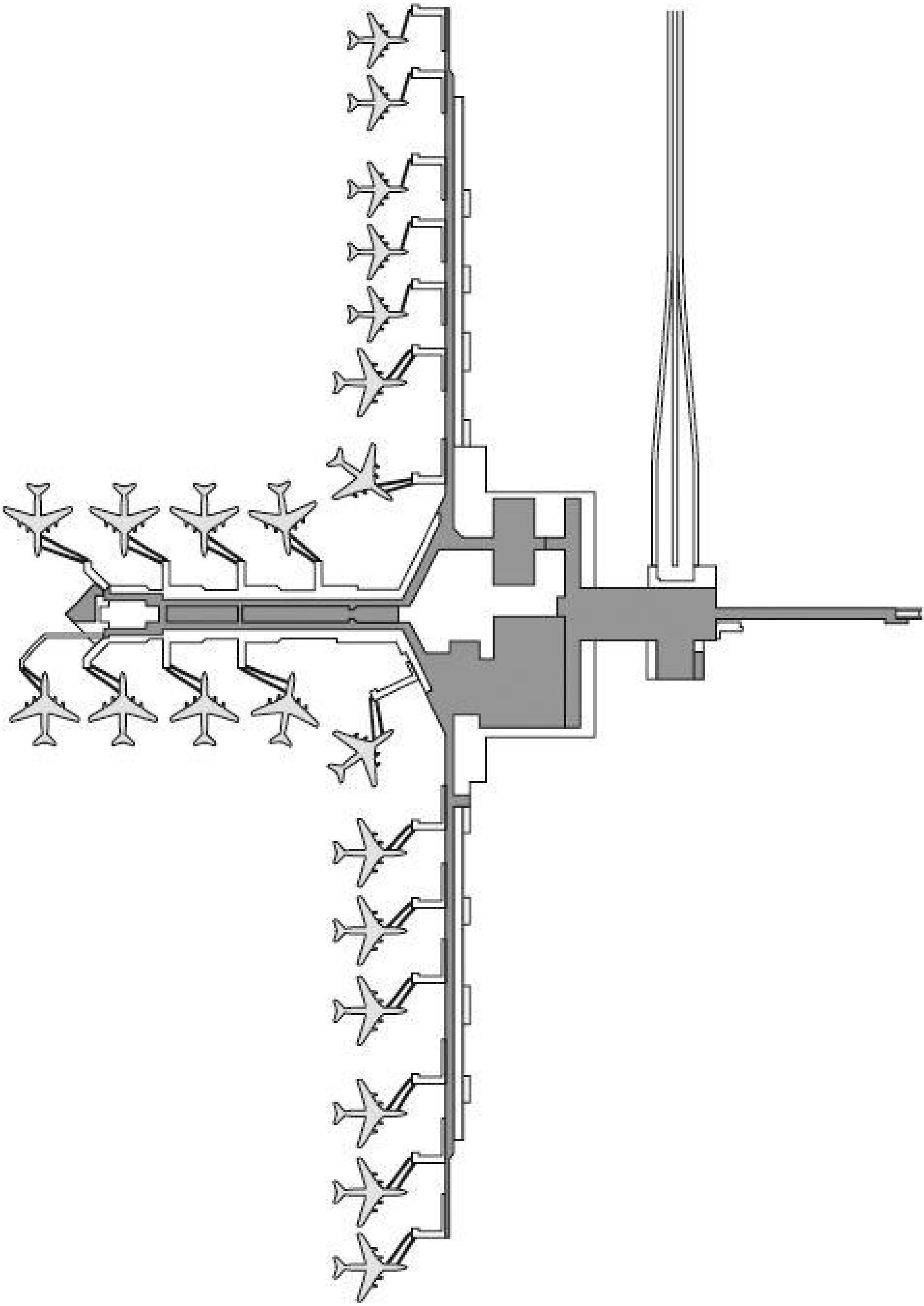


FIGURE 14.2 Figure pier design for Nagoya/Chubu airport in Japan. (*Source*: Pacific Consultants International.)

Satellites

Satellites are the logical extensions of T-shaped finger piers. They eliminate gates along the fingers and concentrate them at the end. Generally, the connection between the satellite and the central check-in area is above ground. Some designs place the finger underground so it is invisible. The satellite is sometimes connected to the central passenger building by a people mover, sometimes not. Table 14.4 and Fig. 14.3 indicate examples of the possibilities.

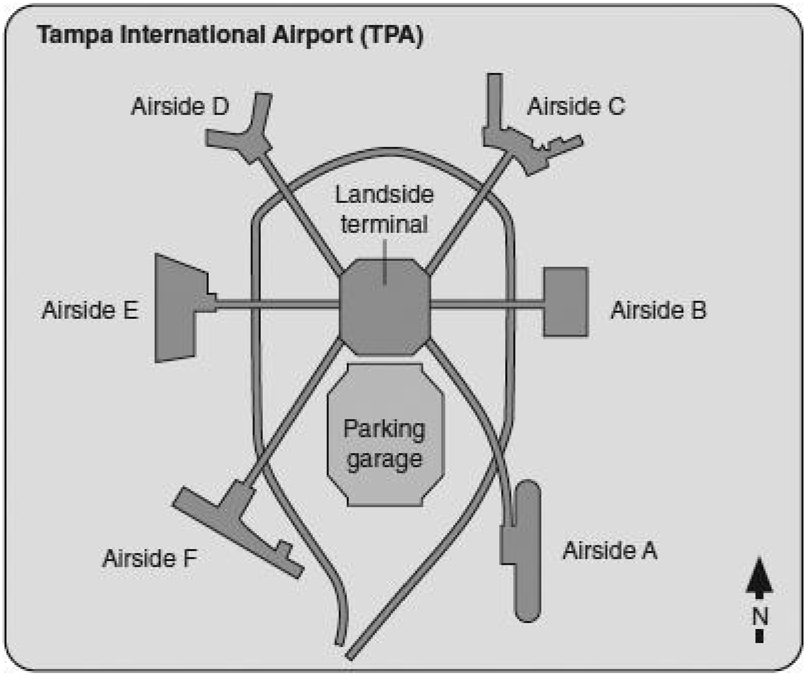


FIGURE 14.3 Layout of satellites at Tampa Airport. (*Source*: www.airportterminalmaps.com.)

Connection to Main Building Is	People Mover	
	No	**Yes**
Above ground	New York/Newark	Tokyo/Narita
Under ground	Paris/de Gaulle	Seattle/Tacoma

TABLE 14.4 Example Arrangements for Satellite Passenger Buildings

Satellites with underground connections to the main building have a singular advantage. They allow aircraft to maneuver freely around the satellite. This facilitates aircraft operations and saves the airlines time and money. This is the arrangement for Terminal 1 at Paris/de Gaulle and at Seattle/Tacoma. The detached facilities at Seattle/Tacoma connect to the main passenger building by a people mover and are functionally close to what we now call midfield concourses.

Midfield Concourses

Midfield concourses are major independent passenger buildings, often located far from the central passenger building that passengers access from the groundside. They may have around 50 gates and be about 1 km long. The linear Concourse B occupied by United Airlines at Denver/International is 990 m (3300 ft) long and has 46 aerobridge gates with a ground-level extension serving additional positions for small aircraft. The X-shaped midfield concourse at Kuala Lumpur offers 27 gates. Midfield concourses are typically between parallel runways and separated from the other passenger buildings by major taxiways. They can also be located on the edge of the runways as part of a complex of passenger buildings, as is the United Airlines midfield concourse at Chicago/O'Hare. Midfield concourses differ from satellites in their size and distance from the groundside, but this distinction is not firm.

Because of the distances and number of people involved, passengers usually access midfield concourses by self-propelled people movers. Reliable, economical people movers have transformed the possibilities for the design of airport passenger buildings and other landside facilities such as parking lots and car rental facilities.[2] They are indispensable for the development and operation of midfield concourses.

Midfield concourses come in two basic shapes: linear and X-shaped. Linear concourses are long buildings with aircraft gates on both sides (Fig. 14.4). They are frequently wider in the middle, to accommodate the people mover station, provide a central shopping area, and serve larger aircraft and their numerous passengers. They are typically flanked by dual parallel taxiways that allow aircraft to move between their gates and the runways with a minimum of turns and delays. Atlanta and Denver/International built their entire airport around linear midfield concourses. Chicago/O'Hare built a single midfield linear concourse for the transfer operations of United Airlines.

FIGURE 14.4 Midfield passenger buildings at London/Stansted. (*Source*: BAA plc.)

X-shaped midfield concourses feature four fingers intersecting at right angles. They are rare. The prime examples are at Kuala Lumpur/International and Pittsburgh. X-shaped configurations implemented so far have only used one midfield concourse, in contrast to Atlanta, Denver/International, and London/Stansted, which use several parallel concourses.

In the Pittsburgh version, the crosspieces are oriented at 45 and 135° from the parallel runways. This arrangement maximizes the number of gates that can fit between parallel runways. It was appropriate because Pittsburgh sits on difficult terrain that limits the space between parallel runways. In the Kuala Lumpur/International version, the crosspieces are perpendicular or parallel to the runways.

The motivation for cross-shaped midfield concourses was the idea that the maximum walking distance in this configuration is less than for a linear concourse, if both have the same number of gates. The intuition is that if you halve a line segment and cross the pieces, you also halve the maximum distance from end to end. Counter intuitively however, the X-shaped concourse actually increases the effective walking distances for most passengers!

An X-shaped concourse provides inferior service to passengers because

- Aircraft gates cannot be located at the center of the concourse (because of the corners), which means that the frontage of the building has to be greater than that of a linear concourse with the same gates.
- Furthermore, maneuvering is difficult in corners so that larger aircraft are usually positioned at the ends of the fingers, in contrast to the linear concourse in which the larger aircraft gates are at the center, convenient to the people mover station.
- So that, in the X-shaped concourse, the greater number of people has to walk the longest distances.

These features combine to raise the average walking distance experienced by most passengers. The X-shaped design also complicates airline operations, as compared to linear midfield concourses. This is because the X-shaped configuration involves more turns and delays. Section 14.4 explains these points further.

Linear Buildings

Linear buildings are long structures with one side devoted to aircraft and the other faced by roads and parking lots. This design was a response to the great walking distances associated with finger piers. Designers originally called it the "gate arrival" concept. The idea was that people could arrive at the airport right at their departure gate and walk to their flight through a thin building.

Several airports built gate arrival buildings in the 1970s and 1980s. These include Dallas/Ft. Worth, Kansas City, Paris/de Gaulle (Terminals 2A–D), and Munich (Fig. 14.5). Several of these "linear" buildings in fact curved around an interior landside parking area. This curved plan has the advantage of providing longer frontage on the aircraft side, where it is needed to accommodate aircraft wingspans (see Andreu, 1997). However, curved plans complicate both the initial construction and subsequent landside traffic flows.

FIGURE 14.5 Linear passenger building at Munich. (*Source*: Munich International Airport.)

"Gate arrival" buildings are no longer in favor. Their great disadvantage is that they are particularly expensive; they need to be twice as long as finger piers or midfield concourses for the same number of gates. They are also operationally inefficient. It is uneconomical to have passengers flow directly from curbside to their aircraft; airports and airlines cannot afford it to staff duplicate check-in and security facilities in front of each gate, instead of combining them in central services that cater to passengers for many gates simultaneously. It is also unproductive to spread passengers out along the building because it eliminates the possibility of significant retail areas, because single gates do not provide enough traffic to justify important stores.

As a practical matter, designers now allow only a few access points to linear passenger buildings. Passengers thus arrive broadly along the front of the building, depending on where space is available, flow to some central area for check-in, security, and shopping, and then proceed out to their gates. The linear concept in this configuration in effect amounts to a finger pier. As a result, it does not minimize walking distances as much as designers originally imagined. Ironically, the "gate arrival" concept, designed to minimize walking distances, evolved into a configuration in which these distances are significant.

Transporters

Transporters comprise the broad category of rubber-tired vehicles that move passengers between buildings and aircraft. Typically, these are specially designed buses with low platforms and wide aisles for easy access for passengers with bags. Bus systems require passengers to negotiate the stairs between the airport apron and the aircraft door. Major airports in Europe and Asia commonly use transporters, for example, Frankfurt/International, Munich, Paris/de Gaulle, Zürich, and Tokyo/Narita (Fig. 14.6).

FIGURE 14.6 Transporter in operation at Berlin/Schönefeld. (*Source*: Berlin International Airports.)

Complex transporters known as "lift lounges" have a cabin that drivers can raise and lower. They are similar to the range of catering vehicles that service aircraft, in which large hydraulic scissors or screw jacks move the cabin up or down. Lift lounges are much larger than catering vehicles. They carry about 80 to 100 passengers. In operation, the lift lounges let passengers board at the normal elevated level associated with departures through aerobridges, lower the cabin for travel to the aircraft parked at a remote stand, and then raise entire passenger compartment to the level of the aircraft door, enabling passengers to enter the aircraft horizontally. Lift lounges avoid the problems associated with getting passengers to use stairs to get to the aircraft. They keep the passengers inside at all times; avoid the climatic extremes of heat or cold, snow or rain; speed up the loading and unloading processes (many people find that getting up a long flight of stairs with carry-on baggage is an athletic challenge); and thus increase the productivity of the aircraft and crew. Lift lounges are particularly expensive to buy and operate, however, and are disappearing from airports.

Designers developed the transporter configuration as a way of avoiding the cost and long walks of construction of finger piers. The transporters take passengers to their aircraft, parked wherever convenient on the apron, directly from a central passenger building housing check-in, baggage claim, stores, and other facilities. This configuration has the obvious advantages of minimizing walking distances, eliminating significant construction costs, and freeing aircraft from the difficulties of docking at passenger buildings.

Transporters pose significant operational challenges, however. Most obviously, they are expensive. The manufacturers produce small numbers of these special vehicles and do not secure the economies of scale associated with the manufacture of cars. Transporters require fueling and maintenance (see the discussion in Example 14.2). Finally, transporters require specially trained drivers to navigate safely on the airfield and present a peculiar risk for airport operators; they place a small group of skilled workers in the position of being able to shut down airport operations.

Transporters also present difficulties for the airlines and their passengers. The use of transporters adds 10 to 15 minutes to a flight, because of the time it takes to load and unload these vehicles. These delays are particularly inconvenient on short-haul flights and in transfer operations. Additionally, the airside airport busses offer inferior levels of service because they force passengers to go out into the weather, cope with stairs, and stand while in motion.

Transporters are useful and effective in the special situations where their advantages outweigh their disadvantages (see Sec. 14.4). The practice of using transporters exclusively, which was the original design for Washington/Dulles airport, has now disappeared. Transporters are particularly useful as part of hybrid configurations, in particular to serve seasonal and low-fare operations, as Sec. 14.5 describes.

Centralized and Dispersed

Designers can centralize or decentralize any configuration. A centralized version provides a single point of access to the airport and is convenient for rail and other forms of public transport. Bangkok/Suvarnabhumi is a good example. A dispersed or decentralized concept substitutes smaller buildings for the single massive structure. This arrangement is sometimes called the "unit terminal" concept. The decentralized configuration can work well for airlines or airline alliances that have distinct operations. At New York/Kennedy, for example, American, Delta, and United Airlines all have their own complexes. Figure 14.7 shows a decentralized configuration of passenger buildings.

FIGURE 14.7 Decentralized passenger buildings at New York/Newark Liberty. (*Source*: Port Authority of New York and New Jersey.)

Decentralized configurations have several disadvantages. Separate buildings

- Complicate transfers between them, as the example of Kansas City cited in Sec. 14.1 indicates and as occurs for transfers between international and domestic terminals at Chicago/O'Hare, Los Angeles/International, and across Australia
- Inhibit the growth of airlines at an airport, as they find it difficult to operate in distinct buildings, as the experience of American Airlines at Dallas/Ft. Worth demonstrated
- Make it difficult to have a central rail station convenient for all passengers, as the cases of London/Heathrow and Paris/de Gaulle show

14.4 Evaluation of Configurations

Which configuration of airport passenger buildings is best? That is a basic question for airport owners and designers. The discussion and examples so far indicate that there are too many factors, and too many stakeholders with different concerns, to satisfy all stakeholders completely. There cannot be a universal answer best for all. Furthermore, the variety of designs recently implemented shows that there is no consensus among designers. How should owners and designers proceed?

The basic principle in deciding which configuration is best is to analyze the issues comprehensively—a systems approach in short. Any single-dimensional approach will leave out too many factors and risk operational disasters (as associated with the "gate arrival" design at Kansas City, which focused narrowly on walking distances of local passengers). Acting on intuition alone without analysis will frequently lead to error (e.g., assuming that cross-shaped midfield concourses minimize walking distances).

Systems designers recognize that, while it is impossible to define a solution that is best in all cases, it is possible to determine some fundamental characteristics of good design. Experience shows that the best design

- Depends on the specific circumstances, the site, the type of traffic, and the needs of the several stakeholders
- Includes features that cater to the specific needs of the variety of clients and stakeholders, and is thus unlikely to be described by a simple concept
- Is flexible, so that it can deal with the changing needs of the clients and stakeholders over the life of the project

Section 14.5 provides a guide of which configurations are best in which circumstances. This section presents four considerations that support those guidelines. These focus on the following:

- *Walking distances*, a factor that has motivated designers to search for better configurations—the analysis demonstrates that simplistic intuitions based on geometrical measures of distances are deceptive and frequently wrong.
- *Aircraft taxiing around the buildings*, which involves substantial costs and is fundamentally important to the major clients of an airport and, ultimately, to the competitiveness of an airport compared to other airports
- *Transporter economics*, to indicate when these vehicles provide a cost-effective complement to passenger buildings
- *Flexibility*, the ability to adapt to different types of traffic as they evolve, such as the development of transfer or international traffic into a significant share of traffic at an airport

Walking Distances

We can investigate the implication of the configuration of airport passenger buildings for walking distances, using spreadsheet programs such as Excel (de Neufville et al., 2002). These enable us to analyze rapidly the performance of different designs for any distribution of traffic, both between the landside entrances and the gates, and between gates for transfer traffic. Using the "data table" functions embedded in spreadsheets, we can also easily run parametric analyses for a wide range of conditions. These analyses require little time to create and run. Anyone familiar with spreadsheet programs can create their basic elements in a day or so for any specific airport.

The procedure for analyzing walking distances in airport buildings uses two origin-to-destination matrices: *impedance* and *flow matrices*. Each matrix captures a different aspect of the data on traffic within the airport passenger building.

The *impedance matrix* defines the level of difficulty in transiting between any gate (or access point to the passenger building) and any other. It describes the physical aspects, the geometry of the building. In a simple version, it defines the walking distances between these origins and destinations. It could also represent travel time or some modified measure of distance that accounts for either the benefits of moving sidewalks, people movers, or other devices, or the inconvenience and delays due to stairways, security checks, or other barriers to movement.

The *flow matrix* defines the number of passengers moving between each origin and destination represented in the impedance matrix. It embodies operational information about the airport passenger building. It accounts for two elements of the issue that purely geometric analyses ignore. Specifically, it reflects the following:

- *Transfer patterns*, that is, the passenger flows from one aircraft to another
- *Intelligent management of gate assignments*, whereby either airline managers or airport operators place flights with significant transfer traffic at gates close to each other

The impedance and flow matrices conveniently represent all the basic information on the passenger movements within airport passenger buildings. Their distinct functions facilitate analyses of different issues. Architects, for example, can investigate the effect of alternative configurations of the buildings for any airport with specific level of transfers. Airport managers can examine the implications of different operational strategies for assigning gates to aircraft.

It is easy to calculate all the interesting statistics on walking distances using these matrices. Multiplying them gives a *passenger-impedance matrix* in which each cell represents the impedance between each origin and destination, weighted by the number of passengers. For example, if we measure impedance in meters to be walked, each cell in the passenger-impedance matrix represents the passenger-meters walked by the traffic between the corresponding origin and destination in the building. Summing these results and dividing by the total traffic gives the average walking distance (see Example 14.1). Sorting the cells by distance, and summing the corresponding passenger impedances expressed in terms of percentage of the total, permits the analyst to develop cumulative passenger-impedance diagrams that show the proportion of passengers walking specified distances for any situation.

Example 14.1 This example illustrates the spreadsheet method for calculating walking distances. It also demonstrates that the average walking distance, when the airport operators allocate aircraft to gates intelligently so that connecting, originating, and terminating traffic are close to their gates, is considerably less than a purely geometric analysis would suggest.

Consider a finger pier with four gates and a point of access to the main body of the passenger building. It is 18 m (60 ft) wide. Its gates are 60 m (200 ft) apart and laid out for entry on the left-hand side of the aircraft (Fig. 14.8). Table 14.5 shows its impedance matrix in meters.

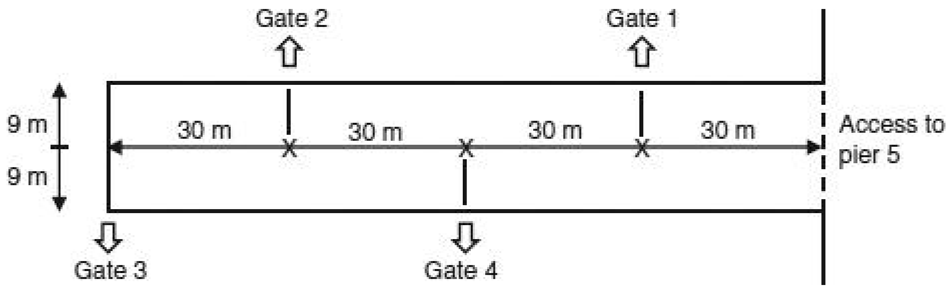


FIGURE 14.8 Sketch of finger pier for Example 14.1.

To Point	From Point 1	2	3	4	5
1	0	78	108	48	39
2	78	0	48	48	99
3	108	48	0	78	129
4	48	48	78	0	69
5	39	99	129	69	0

TABLE 14.5 Impedance Matrix (Meters) for Example Finger Pier

An aircraft with 100 passengers occupies each gate. The airport operator has intelligently placed the aircraft with the most transfers close together, at gates 4 and 2. Table 14.6 shows the assumed flow matrix. The summations at the right-hand side and the bottom indicate that each aircraft arrives and departs with 100 passengers and that 220 passengers enter and exit through the end of the finger pier. This means that 180 passengers transferred or stayed on board their aircraft.

	From Point					
To Point	1	2	3	4	5	Total
1	10	0	0	0	90	100
2	0	20	10	40	30	100
3	25	10	20	15	30	100
4	15	0	0	15	70	100
5	50	70	70	30	0	220
Total	100	100	100	100	220	620

TABLE 14.6 Flow Matrix (Passengers) for Example Finger Pier

The resulting passenger-impedance matrix is in Table 14.7. The totals on the right-hand side add up to a total of 42,630 passenger-meters walked in the finger pier. This implies an average distance per person of 68.76 m (42,630/620). This is far less than maximum possible distance of 129 m, and also less than the average of 74.67 m if there were no transfers. Looking carefully at what passengers actually experience, especially with intelligent management of the aircraft stands, substantially improves on a simplistic assessment of walking distances based only on the geometry of a facility.

	From Point					
To Point	1	2	3	4	5	Total
1	0	0	0	0	3,510	3,510
2	0	0	480	1,920	2,970	5,370
3	2,700	480	0	1,170	3,870	8,220
4	720	0	0	0	4,830	5,550
5	1,950	6,930	9,030	2,070	0	19,980
Total						42,630

TABLE 14.7 Passenger-Impedance Matrix (Passenger-Meters) for Example Finger Pier

Designers should note carefully how Fig. 14.9 demonstrates the importance of intelligent management of the gates on walking distances. Managers of gate assignments can dramatically reduce the walking distances passengers experience by placing connecting flights near each other, placing larger aircraft near the exits, etc. Thus, although they cannot affect the maximum distance passengers experience, they can reduce the average, sometimes by as much as a third.

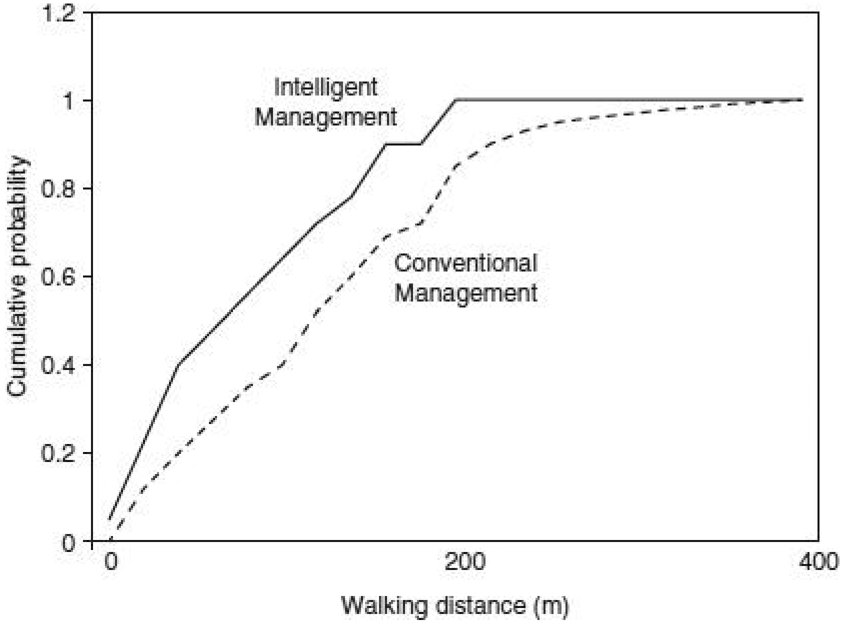


FIGURE 14.9 Passenger walking distances with and without intelligent gate assignment for an example linear midfield concourse (high transfer rate).

Table 14.8 summarizes the results of applying this analysis to example buildings (see de Neufville et al., 2002, for details). Designers should carefully notice that the results depend significantly on the level of transfer traffic. Moreover, the reader should appreciate that these results assume that the airport operators practice intelligent gate assignment; if connecting aircraft are far apart, the results degrade significantly. In any event, the desirability of the configuration depends on the current and anticipated level of transfer traffic. Overall, this analysis suggests the advantages of intelligently managed linear midfield concourses and the disadvantages of finger piers about walking distances.

		Average Walking Distance (meters per person)	
Overall Configuration	**Specific Form**	**60% Transfers**	**No Transfers**
Midfield concourse	Linear	90	109
	X-shaped	134	136
Finger piers		202	116
Linear building One airside	3 entrance points	109	98
	1 entrance point	144	157

*Comparison of walking distances in airport passenger buildings is based on the assumption of intelligent gate management for 20-gate building in example.

Source: de Neufville et al., 2002.

TABLE 14.8 Comparative Walking Distances in Airport Passenger Buildings*

Linear midfield concourses minimize average walking distances better than X-shaped configurations. Comparing buildings serving the same number of gates, the X-shaped buildings of course reduce the maximum walking distance, because they spread the gates in four directions rather than two. However, X-shaped buildings have higher average walking distances. This is because they have "dead zones" at the center that make it impossible to locate aircraft there, and force managers to place larger aircraft toward the ends of the piers.

The relative performance of linear midfield passenger buildings and finger piers depends on the effort required to transit between the midfield concourse and the landside, and the number of passengers who are not transferring and must cross this distance. For transfer passengers, either building performs well, to the extent that managers can cluster the aircraft along the pier or can use a shared lounge area at the end of the finger pier. For local passengers, the midfield linear concourse is superior regarding walking distances within the building itself. This is because it enables managers to position large aircraft at the entrance located conveniently in the middle of the building, and thereby to minimize walking between the entrance and the aircraft. Which of the two configurations is better regarding overall walking distance depends on the specific situation.

Linear buildings with one airside and one landside perform well for originating passengers but poorly for transfers. In principle, they minimize the walking distance between the curb and the plane. In practice, however, this advantage is lost because airport managers limit the number of access points and thus reduce the cost of security checkpoints—as they do at Dallas/Ft. Worth, for example. When the number of access points to a linear building is limited, the walking distances can be relatively long for any sizable building.

Transfer passengers in linear buildings with one airside find that, even with intelligent management of the gate assignments, their walking distances are necessarily relatively long. This is because a linear building with one airside is twice as long as one with gates on both sides of the building (as they are in a midfield concourse or a finger pier). The excessive walking distances for transfer passengers associated with the linear passenger buildings at Kansas City was one reason TWA transferred its base of operations to St. Louis when it set up a hub-and-spoke system to serve transfer traffic. However, the walking distances for local passengers in a linear building are reduced significantly if decentralized facilities are used, including the provision of several entrance points to the building.

Aircraft Delays

Planners can estimate aircraft delays due to terminal configurations in the same way as they can estimate walking delays, provided the area around the passenger building is not very busy and aircraft do not have to wait for each other. Such situations are likely to prevail around new buildings, because designers will be careful to provide sufficient taxiway capacity. Where congestion is likely or already exists, designers should use detailed simulations of airfield traffic.

A general result from the analysis of aircraft delays concerns midfield concourses. Equivalent linear midfield concourses incur less delay for aircraft than the X-shaped concourses at Kuala Lumpur/International. Because the linear concourse allows for direct access to the gate from the taxiway, whereas the alternative arrangement may require several turns around the X-shape, the linear concourse reduces the average taxi distance around the passenger building by 25 percent and halves the number of turns. This implies savings of about 1 minute or \$50 to \$100 per operation. Summed over several tens of thousands of operations a year, the advantage of the linear midfield concourse is of the order of millions of dollars a year.

Transporter Economics

Transporters provide a flexible, cost-effective alternative to investments in airport passenger buildings. Despite the unattractiveness of busing passengers, transporters are useful in special situations. They enable airports to deal with unavoidable airfield constraints. They provide a cost-effective way to deal with seasonal low-cost traffic, such as that at some vacation destinations. As this last point is not obvious, it is explained in detail.

Transporters are necessary at airports where it is impractical to extend airport passenger buildings to provide extra gates for aircraft. This is the situation at Milan/Linate, for example. Likewise, they also provide “gate capacity” until the construction of permanent gates catches up with needs, as they have done at Frankfurt/International, Lisbon, and Paris/de Gaulle.

Managers of airports with strong seasonal variations in passenger traffic should seriously consider using transporters to provide “gate” capacity in these peak periods. This is because they can minimize the cost of transporters during off-peak periods; they can park them and turn off their operating costs. Although transporters are in general an expensive way to connect passengers and aircraft, they are actually cheaper than building permanent gates for peak period demands.

To understand when and why transporters are more economical, we need to appreciate the difference in cost structure between transporters and permanent gates:

- Transporters are inexpensive to acquire but expensive to operate.
- Building gates at a modern passenger building can be very expensive, but the operating costs per passenger are low.

This means that the relative cost per passenger for each alternative depends on the numbers of passengers they serve. If many people use a gate in a building, the high capital cost spreads out and its average cost per passenger is low. Contrarily, if this structured gate is only needed to cater to traffic during a peak period, its average cost per passenger is high. (See Example 14.2.) By comparison, because the operating costs dominate the cost of transporters, their average cost per passenger is relatively constant. This difference in cost structure between the alternatives means that built gates are better for high volumes, transporters better for infrequent use.

The economic desirability of transporters depends on the variability in the number of passengers at an airport over the year. If the traffic is highly seasonable—as it might be for vacation destinations—managers can use transporters intensively during a few peak months and then not at all during the rest of the year. Although the transporters are full when used, their average use over the year may be low enough to make them economical and cost-effective overall.

Example 14.2 The cost per flight of capacity needed only for peak conditions can be very high. Suppose that an airport needs a gate only 30 days a year, for five operations a day. Suppose further that the cost of a gate at a significant airport is $10 million to build, that is, about $1,000,000 a year, including depreciation, maintenance, climate control, and so on. The marginal cost of this gate needed only for the peak period is then $6667 per operation ($1,000,000/150). This is the figure to compare to the cost of operating a transporter for these days.

Note that this marginal cost at the peak period is much higher than the average cost per operation for all the gates, which might be about $548 [= $1,000,000/(365)(5)]. The marginal cost of a gate in a building rises dramatically as its usage falls. This is the fact that can make transporters economical. The average cost of all the gates is not relevant to the analysis of what to do for the peak periods.

Example 14.3 illustrates the analysis. Note that the costs used in the example are plausible but not definitive. Building costs can easily differ by a factor of 2, depending on the standards adopted. Labor costs can differ by even more, depending on the salary levels, the benefits, and the number of hours worked. Readers should focus on the method and general results.

Example 14.3 Transporters can provide an economically efficient solution for providing peak capacity. Following from example 14.2, assume that

$$\text{Building gate \$/flight} = \$1{,}000{,}000/(\text{annual flights at gate})$$

For the transporter, the example assumes that its operating costs are $300/h when used, that it can serve two flights per hour, and that two transporters are needed for each flight. Assuming that the annual capital cost of a transporter is $100,000, then

Transporter \$/flight = \$100,000/(annual flights served) + \$300

If the airport needs the gate every day of the year, building a gate is cheaper. For example, at 4000 flights/year, the example building costs \$250/flight, less than the \$325/flight cost of the transporter. On the other hand, if the airport only requires this gate capacity in the high season for 400 flights, the marginal cost of the building gate is \$2500/flight whereas the transporter cost of \$550/flight is much less expensive.

Airport managers need to do their own analyses to determine when transporters are economical for them. No general rule is possible due to the large variation in transporters, building costs, and labor conditions worldwide. Much depends, for example, on the extent to which the airport can reduce the cost of drivers when transporters are not needed; seasonal or part-time employment is acceptable in some circumstances and not in others. Managers also need to consider the current and likely future degree of seasonal use, which differs significantly between airports. At some airports, gates provided for the peak months will be heavily used most of the year. This has been true for New York/LaGuardia as Fig. 14.10 indicates, and generally for airports focused on business travelers. However, at airports serving many leisure travelers for a few months of the year, such as New York/Kennedy, gates provided for these peak months may not be needed throughout the rest of the year. Transporters can provide an economically attractive solution at such airports.

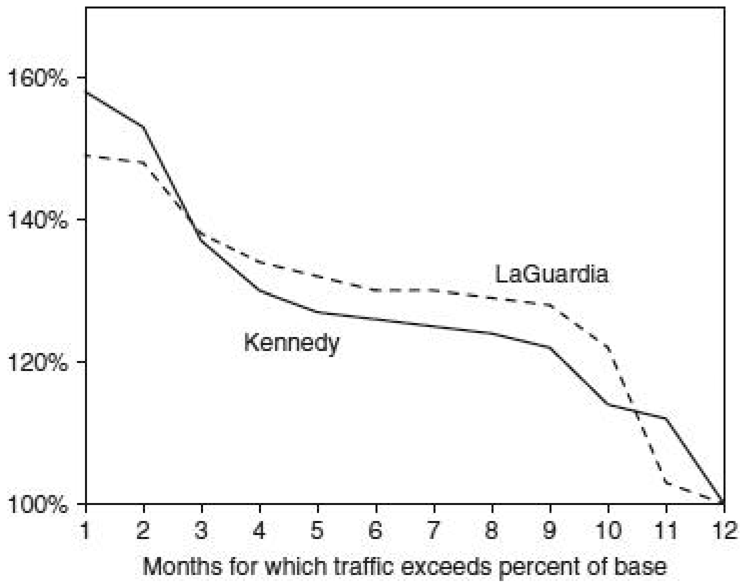


FIGURE 14.10 Example seasonal variation in passenger traffic at two airports. (*Source*: Port Authority of New York and New Jersey, 2009.)

Prospective users of transporters should note that while this solution may provide for a large fraction of the number of gates needed in peak periods, it serves only a small fraction of the total annual passengers. Consider the implications of the data on New York/ Kennedy airport in Fig. 14.10; the extra traffic in the peak 2 months (an average of about 155 percent of the traffic in the lowest month) is one-sixth or 15 percent higher than that of the next highest month. This extra traffic is the sum of the extra traffic over the peak months. (Technically, it is the integral under the curve down to the level of the third month and is only about 2.5 percent of the total traffic throughout the year, represented by the area under the entire curve.) If an airport in a similar situation provided 15 percent of the peak "gate" capacity using transporters, it would have them serve only about 2.5 percent of the total yearly traffic. Such disparities between the amount of peak capacity and the number of people needing it are common when airports have seasonal peaks. Therefore, the inconveniences associated with transporters used during peak periods affect only a small fraction of the traffic.

Flexibility

Whatever we build today may somehow be inappropriate sometime in the future, perhaps soon. Most obviously, traffic may be expected to grow, and larger facilities are likely to be needed. More subtly, and generally more importantly, the mix of traffic and its demands will somehow be different both from what they are today and from current expectations. For example,

- The mix of traffic may change to include a much higher or lower level of transfers. As Salt Lake City and Munich became transfer hubs, and Pittsburg and Cincinnati lost this traffic.
- The proportion of passengers requiring international clearances may change, as countries enter into common customs unions such as the European Community, Mercosur, Nafta, or Asean.
- The requirements of the airlines may change as they merge or form alliances, as happened at New York/Kennedy when many foreign carriers joined up operationally with major U.S. airlines and no longer needed a strictly international passenger building.

The airport configuration will need to evolve. It must also be able to evolve flexibly, to serve different types of customers and needs. Designers are not able to anticipate correctly exactly what will be required. As Chap. 4 emphasizes, forecasting is not a science and the forecasts are "always wrong." This means that the configuration of today must be able to adapt to circumstances other than those that now appear most likely (de Neufville and Scholtes, 2011).

In general, centralized complexes of passenger buildings can accommodate change more easily than decentralized facilities. This is because, in a centralized facility, as one airline or type of service grows relative to another, it can move over gradually into other parts of the passenger building (see discussion in Chap. 15 on shared use). Such incremental change is difficult in decentralized buildings, as the case of Baltimore/Washington demonstrates. As Sec. 14.1 indicates, this airport could not make use of the international gates vacated by US Airways when Southwest needed extra capacity, in large part because the facilities were decentralized.

A prime way to achieve flexibility is to ensure that space is available for eventual future requirements. Having space available is more than having land available somewhere on the airfield. The space should be next to the existing facilities, and the facilities should be designed so that they can grow into this space. The Aérogare 2 complex at Paris/de Gaulle provides a good example of this. In this case, the Aéroports de Paris developed their initial phase of buildings along a spine road with lots of space, so that they could extend the facilities almost indefinitely according to need (Andreu, 1997).

Hybrid configurations are also generally more flexible. Kuala Lumpur/International provides a good example of flexibility in general and in particular due to a mix of configurations. This complex of passenger buildings not only allows plenty of space for various forms of midfield concourses but also carefully provided ample space to lengthen the main passenger building, as well as clear zones available for the expansion of baggage facilities and the introduction of rail access. Furthermore, it combines a midfield concourse and finger piers that can flexibly handle varying proportions of international and domestic traffic. This combination allows it to expand easily the specific facilities that may be most appropriate in the future.

14.5 Assessment of Configurations

The evaluation of the configurations demonstrates that it is impossible to define a solution that is best in all cases. Table 14.9 provides a subjective summary of the results of the analyses. Others may look at these results and other factors and come to different conclusions. In any case, it should be clear that no single design is best overall. Some designs, such as finger piers and transporters, are less attractive, but no design is dominant for all stakeholders.

Configuration	Passenger		Airline	Owner	Retail
	Local	Transfer			
Finger pier	Fair	Poor	Fair	Fair	Good
Linear	Fair	Poor	Good	Fair	Poor
Transporter	Fair	Poor	Poor	Poor	Good
Midfield linear	Fair	Good	Good	Fair	Good
Midfield X	Fair	Fair	Fair	Good	Good

TABLE 14.9 Subjective Comparison of Configurations of Airport Passenger Buildings

The transfer passengers may be the single most important factor for designers to consider in the choice of configurations. This is because what is good for them may be poor for local traffic and vice versa. When designers know they are developing a destination airport that will not have much transfer traffic, they may feel freer to develop linear passenger buildings. On the other hand, if they know or anticipate that the airport will be a transfer hub, they should focus on midfield concourses. The shape and indeed the possibility of a midfield concourse depend on the availability of space. When space is particularly tight, designers may find that satellites are a better solution, as they did in developing Milan/Malpensa. In this case, satellites similar to those at Seattle/Tacoma may provide attractive possibilities as they can offer many of the advantages of midfield concourses (see Sec. 14.3).

No consensus exists about centralized or decentralized passenger buildings (Table 14.10). Airport operators and transfer passengers tend to prefer centralized buildings. The owners appreciate the advantages of concentrating services and retail areas (as at Amsterdam/Schiphol). The transfer passengers find transferring between buildings difficult, especially when the distances are huge as they commonly are between the domestic and international passenger buildings in Australia and at Chicago/O'Hare. On the other hand, major airlines, especially those operating a transfer hub, like to control their own space in a self-contained building as they do at New York/Kennedy (American, Delta, and United buildings), London/Heathrow (British), and Tokyo/Narita (Japan Airlines).

Configuration	Passenger		Airline	Owner	Retail
	Local	Transfer			
Centralized	Good	Good	Fair	Good	Good
Dispersed	Good	Poor	Fair	Poor	Poor

Table 14.10 Subjective Comparison of Centralized and Dispersed Airport Passenger Buildings

The final choice of the configuration of new airport passenger buildings should depend on its specific circumstances, the site, the type of traffic, and the needs of the several stakeholders. Good designers will pay attention to each of these prospective users of the airport, although they are not the immediate clients. They will thus provide the better overall design for their main clients over the long run.

14.6 Hybrid Configurations in Practice

Major established airports as a rule do not exhibit a single configuration of passenger buildings. Over the years, they build a variety of facilities designed to serve the particular needs of many users. For example,

- New York/LaGuardia first built a Marine Air Terminal that became a decentralized building exclusively serving Delta Shuttle services to Boston and Washington. Its next major building featured finger piers and provided centralized service to many airlines. Its latest buildings are decentralized unit terminals serving specific airlines. (It functions well with these distinct buildings because it has virtually no transfer traffic.)
- Paris/de Gaulle started with a centralized unit terminal that was going to be the first of five. However, its Aérogare 2 is completely different; it is a decentralized linear design (see the discussion under flexibility in Sec. 14.4). To meet the special needs of charter traffic for economical facilities, the airport then built Terminal 3 as a larger hangar by itself. Finally, it built a special facility for domestic flights (Terminal 2G).

Airports routinely evolve to include features that cater to the specific needs of their variety of clients and stakeholders. The configuration loses any original simple concept. In practice, airports eventually adopt a hybrid concept, one that brings together important features of distinct configurations. The hybrid concept emerges either because needs change or because the airport changes its preferences. Either way, airports implement hybrid configurations as they learn what they really need.

Designers should plan for hybrid configurations because they are the likely final configuration of the airport. At a minimum, they need to make sure that their initial plans can respond flexibly to the changing needs of the airport clients and stakeholders over the life of the project. Better, they should from the start build in the elements that most appropriately serve the various needs."

Exercises

14.1. Assess the configuration of your local airport or some airport with which you can become familiar. How would you describe the configuration? To what extent does it appear to meet the needs of the principal stakeholders? How flexible does it appear to be, to meet the requirements of plausible future traffic?

14.2. For an airport of your choice, use its information on departure and arrival of aircraft to identify how its management allocates gates to flights. To what extent does it locate flights intelligently to minimize walking distances for priority classes of traffic (such as heavily traveled domestic flights), for airlines, and for connecting passengers?

14.3. Describe the movement of aircraft near the passenger buildings at some major airport. To what extent are these patterns direct and free of congestion? Require many turns and involve possible delays as other aircraft block flow on the taxiway? How might an alternative configuration of the passenger buildings improve these flows, if at all?

14.4. Get monthly traffic data from one or more airports in which you are interested. To what extent do they exhibit seasonal patterns? How many gates might the airport need only for the peak periods? Think about how transporters could serve these peak requirements.

References

ACRP (2012) *Addressing Uncertainty about Future Airport Activity Levels in Airport Decision Making*, Airport Cooperative Research Program Report XX, Transportation Research Board, Washington, DC. www.trb.org.

Andreu, P. (1997) *Paul Andreu, the Discovery of Universal Space*, L'Arca Edizioni, Milan, Italy.

Ashford, N., Mumayiz, S., and Wright, P. (2011) *Airport Engineering: Planning, Design, and Development of 21st Century Airports*, 3d ed., Wiley, New York.

de Neufville, R., de Barros, A., and Belin, S. (2002) "Optimal Configuration of Airport Passenger Buildings for Travelers," *Journal of Transportation Engineering*, ASCE, *128*(3), pp. 211–217.

de Neufville, R. and Scholtes, S. (2011) *Flexibility in Engineering Design*, MIT Press, Cambridge, MA.

Horonjeff, R., McKelvey, F., Sproule, W., and Young, S. (2010) *Planning and Design of Airports*, 5th ed., McGraw-Hill, New York.

IATA, International Air Transport Association (2004) *Airport Development Reference Manual*, 9th ed., IATA, Montreal, Canada.

ICAO, International Civil Aviation Organization (1987) *Airport Planning Manual, Part 1, Master Planning*, 2d ed., Doc. 9184, ICAO, Montreal, Canada.

---

[1] Any translation between capital and annual costs is imprecise. With depreciation and maintenance, the annual cost of a building is about 10 percent of its capital cost. This is the basis for using a factor of 10 to translate annual savings into an approximation of justifiable additional capital expenditure.

[2] People movers are either self-propelled or pulled by cable as an elevator. Self-propelled vehicles are preferable over longer distances with complex, multiple routes. Tokyo/Narita uses cable-driven people movers to serve its satellite passenger building. See Chap. 17.

CHAPTER 15

Overall Design of Passenger Buildings

This chapter presents ideas essential for defining the amount of space desirable for the several functions of passenger buildings. It shows how planners can translate a concept and traffic estimates into an architectural program that can proceed into detailed design of the interior spaces of passenger buildings. It focuses on the analyses needed in the early phases of design, whereas the next chapter treats the detailed design and management of the airport. Readers should approach the two chapters in sequence.

The emphasis is on process rather than on design standards and specific numbers. No single set of standards can be valid for all airports. This is because the design context can differ substantially between countries and cities. The range of traffic patterns at various locations and times affect the appropriate design standards. Moreover, the clients—airport operators, airlines, government agencies, and concessionaires—may insist on specific levels of service. The text presents illustrative, commonly used standards. These cannot be definitive.

The process for determining the amount of space to be provided consists of two steps. The first is to specify the design loads, for example in terms of thousand passengers per hour. Section 15.1 treats this task, whose outcome is inherently uncertain as it depends on forecasts of future traffic (see Chaps. 4 and 19).

The second step translates the projected traffic into space requirements. Knowledgeable designers recognize that this is a subtle process, contrary to the common engineering notion that projected loads translate mechanically through formulas or tables into design requirements (such as square meters of space). The subtleties of this process arise from two considerations that mediate the translation of traffic loads to design sizes. These concern the following:

- *Sharing of facilities*: To what extent is it possible to share facilities between different categories of users, and thus in effect to double or triple count space? To the extent this is possible, designers can economize on space and provide more efficient designs. Section 15.2 presents the concept, and Sec. 15.3 the analysis required.
- *Performance objectives*: How does the management of the passenger building want to balance economy (which calls for smaller spaces) and quality of service (which implies more space)? Indeed, the translation of design loads into space requirements

is not a purely technical process; it reflects economic and social policy. Sections 15.4 and 15.5 show how to do this when the loads are either static (waiting areas) or dynamic (passageways).

Finally, Sec. 15.6 addresses the issue of sizing space for baggage systems. Baggage space is crucial for the effective functioning of the passenger building, and designers should include it in the original overall space estimates. The analysis for baggage rooms is very different, however, from the estimation of passenger space.

Designing space for sharing among different functions or users is a prime way to reduce overall requirements. Common use space can reduce facility requirements substantially compared to the traditional practice of providing each user and each function with its own separate space. In some cases, common use facilities can reduce the required space by as much as 50 percent. Additionally, common use space is much more flexible than space dedicated to a single user or function. It provides insurance against fluctuations and uncertainties in traffic.

To design space appropriately, planners and airport operators must be clear about their economic objectives. They need to appreciate the tradeoffs they are prepared to make between economy and quality of service. These management choices influence the technical analyses. For some facilities, or indeed entire airports, the airport operator will be looking to emphasize economy. Some airports create "budget terminals" for low-cost carriers: Singapore did this, as did the Aéroports de Paris with Terminal 3 at Paris/de Gaulle. Other airports are completely dedicated to low-cost service, as are London/Luton and Frankfurt/Hahn. Conversely, airports may emphasize quality of service and build generous spaces as Singapore has done in its Terminals 2 and 3.

The formulas for translating design levels of traffic into space requirements are arithmetically simple. However, they implicitly make assumptions about what operators of passenger buildings want to achieve and how they will manage daily operations. Failure to appreciate these assumptions has led to notable design errors.

15.1 Specification of Traffic Loads

The Issue

Planners seeking to define the overall dimensions of the various spaces in a passenger building must first face a difficult task. They must define the traffic for each of the functions these spaces serve. To do this, they must translate general forecasts of overall annual traffic into statements concerning the design loads at peak periods, for detailed operations.

For instance, they must specify the design loads for passport control facilities during its peak, which may be 30 minutes or 1 hour, depending on prevailing traffic patterns.

Simple formulas cannot adequately translate annual traffic into peak loads. Too many local variations exist. For example, international arrivals at some airports cluster around a few peak hours, as they do in Sydney, Australia; elsewhere however, such as Miami/International, they arrive throughout the day. The translation between forecast annual international passengers and the loads on the specific facilities needs to recognize such differences. To define the design loads correctly, analysts need an understanding of local conditions and the evolution of traffic patterns.

Planners need to appreciate that their calculation of design loads cannot be accurate. The process for determining these numbers is not scientific, and the results are not conclusive. Different analysts may look at the same overall forecasts and develop quite different conclusions about the design loads. The U.S. Federal Aviation Administration (FAA) has stressed this point in its graphs of the relation between the level of annual and peak-hour traffic. It pointedly labels each of these graphs with the caption, in capitals in the original for emphasis (FAA, 1988, pp. 10 and 11):

CAUTION!
NOT TO BE USED FOR DESIGN OR ANALYSIS.
FOR USE IN OBTAINING ORDER-OF-MAGNITUDE ESTIMATES PRIOR TO IN-DEPTH ANALYSIS

This warning about the inaccuracy of estimates of design loads needs emphasis. The unavoidable inaccuracy of the estimates of future traffic loads has an important consequence. Planners need to build considerable flexibility into their designs to ensure that their buildings will function properly when the actual loads differ from the design loads. This is a prime motivation for the use of shared or common use facilities.

Peak-Hour Basis for Design

Planners should design airport facilities for peak traffic but not for the absolute maximum traffic. Clearly, they should design for the maximum range of loads rather than average loads. If they did not, the facilities would be under capacity at all the times when the passengers and airlines needed them the most. On the other hand, if planners design facilities for the single busiest period in the design year, the facility would be overcapacity and oversized for every other hour of that year. In practice, the design load is a compromise between economy and the provision of enough capacity to meet peaks.

As a practical matter, it is impossible to define the design load that provides the best compromise between economy and quality of service. To balance these factors precisely, it would be necessary to know both the local cost of the specific facilities and the value of the quality of service to the particular users. As there is no satisfactory way to measure these quantities accurately and as these are changeable in any case, no calculation of an optimal balance can be absolute. All procedures for defining the design load are thus approximate.

Traditional procedures define the design load from statistics that analysts should be able to obtain from available airport data. In this connection, readers should note that commercial airports typically have good data on aircraft movements, even by hour of the day. The air traffic controllers will keep records of their activities both for administrative and billing purposes. Airports have much greater difficulty getting data on passenger flows, especially by hour of the day. They generally must rely on airlines to supply data on passengers, and this may be difficult in some locations.

Each definition of a design peak load represents a compromise between efficiency and quality of service. Different organizations use different measures. Different practices prevail in different regions. Two widely used definitions of the design load have been the *design peak hour* (DPH) in North America and the *standard busy rate* in Britain.

One procedure for defining the DPH applies a variable factor to the "average day of the peak month." (Chapter 21 gives details on calculations.) The factor is about 9 percent, decreasing steadily as the traffic at the airport builds up (see Table 21.3). The standard definition of the "average day of the peak month" is the traffic in the peak month divided by the number of days in the month (ACRP 2007). (Note that this level of traffic does not necessarily represent any actual day, let alone a median day.) As almost all airports have data on daily traffic, this estimate of the design load is widely applicable.

In Britain, an accepted means of estimating design loads for airports has been the *standard busy rate*, defined as the level of traffic during the 30th busiest hour of the year. The disadvantage of this method is that it presumes that the airport operator has reliable statistics on hourly traffic and can calculate the number. However, many airport operators have not collected such detailed statistics, and thus they cannot use this approach.

The International Civil Aviation Organization (ICAO, 2006) proposed a compromise approach to defining DPH loads. This combines elements of the two previous methods. It also presumes that the airport operator has hourly data and can calculate the historical ratio of traffic in the peak hour to that of the peak day. It also suggests, similarly but slightly differently from the FAA approach, calculating the traffic during the average day of the *2* peak months of the year. The analyst then obtains the design load by multiplying this average peak day by the ratio of peak-hour to peak-day traffic.

Most recently, the FAA-sponsored Airport Cooperative Research Program (2013) has published a guidebook for estimating peak-period loads. This latest procedure can be obtained free from their web site. This approach seems useful and may eventually become the new standard. However, as the discussion indicates, none of the available methods is definitive.

For rapidly checking plans, it is possible to estimate peak loads by using rules of thumb, such as these for airports with about 10 million annual passengers:

$$\text{Average peak-day traffic} \approx \frac{\text{annual traffic}}{300}$$

$$\text{Design peak-hour traffic} \approx \frac{\text{annual traffic}}{3000} \qquad (15.1)$$

Adjustment for Decreasing Peaks

An important issue for all approaches to defining peak-hour loads is that the ratio of peak-hour to peak-day traffic usually decreases over time. Uncongested airports typically have sharp peaks in traffic over the day. As traffic builds up, relatively more traffic occurs during the off-hours and the "valleys" in the profile of traffic over the day tend to fill up. Saturated airports tend to have steady traffic throughout their periods of operation. Figure 15.1, comparing the airline traffic over 10-year intervals at Toronto/Pearson, illustrates this effect. It shows how the variability of the traffic, measured in terms of the standard deviation as a fraction of the average level, decreased as traffic grew over that period. Table 21.2 provides similar data.

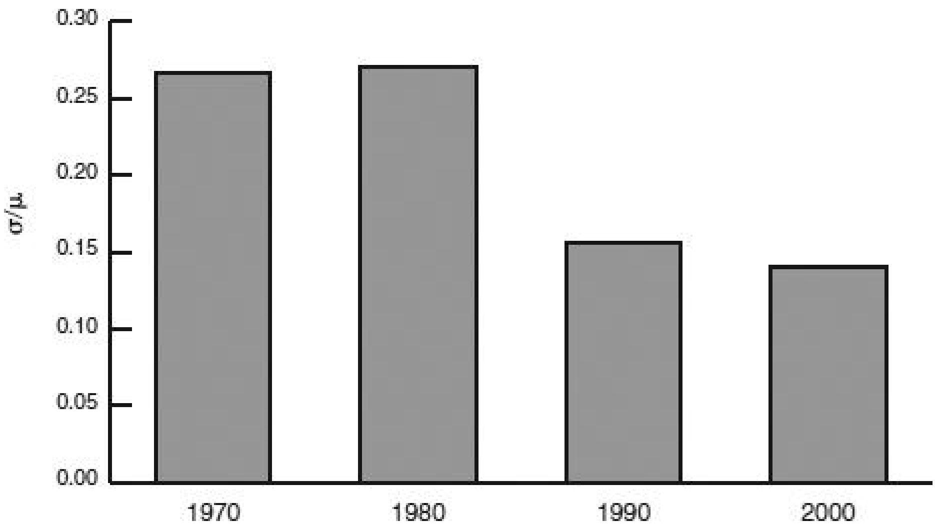


FIGURE 15.1 As traffic increases, its relative variability tends to decrease (data for daily airline traffic on typical Wednesday in October for Toronto/Pearson).

Because daily traffic normally evens out as yearly traffic grows, estimates of peak loads based on historical ratios of peak-hour to peak-day traffic can overestimate future peak loads. This is the case for the standard busy hour and the ICAO approach. The FAA approach, however, deals explicitly with this phenomenon: For larger airports, it lowers the factor for estimating the "average day of the peak month." It systematically accounts for the way traffic evens out as traffic increases.

Nature of Loads

Planners must estimate the loads for each distinct activity for which they intend to estimate the space. Major categories to keep in mind for this purpose are the following (see also Section 14.3):

- Arriving passengers, terminating their travel at the airport
- Originating passengers starting their trip at the airport and needing check-in facilities
- Transfer passengers going on to other flights
- Departing passengers who will need waiting areas
- International and domestic passengers
- Passengers on low-cost airlines who will use special facilities
- Shuttle or commuter passengers needing minimal check-in, lounge, and baggage facilities

Planners should specify the loads according to the crucial periods for the activity under consideration. This may be the peak hour or it may be some other period. For example, the critical period for facilities serving arriving passengers may be the peak 30 minutes after one or more large aircraft land. This may be the governing period for passport control, baggage rooms, and customs facilities. Likewise, planners should recognize that some forms of traffic have seasonal patterns different from other categories of traffic. Vacation travel may be concentrated in a few months, for instance. In short, planners seeking to estimate the overall space required for facilities in passenger buildings need to recognize the distinct local patterns for the traffic they intend to serve.

Note carefully that the total space for an airport passenger building is, as a rule, far less than the sum of the requirements for the distinct groups. Common areas and other shared-use facilities reduce overall requirements. This point needs emphasis. Failure to recognize this fundamental fact is a source of major errors in initial planning documents for airport buildings. Airport operators interested in reducing the cost of a construction program should investigate whether the conceptual planners have properly reduced the total size of the program by incorporating the effect of shared use of facilities.

15.2 Shared Use Reduces Design Loads

Planners need to recognize that the total space they need to provide can be much less than the sum of the space needed for individual activities. Two different activities often can share the same space or facility at different times. This means that this space meets two separate requirements, so the total space that planners need to provide is less than the sum of the requirements for each of the different types of activities. For example, when the peaks of international and domestic traffic do not coincide, boarding and waiting areas can serve international passengers at one time of the day and domestic passengers at another. Such operations have been routine in Atlanta's Concourse E since the early 1990s. By designing facilities that different users can share, planners can reduce the total amount of space. This is a crucial fact that designers need to recognize before they proceed to translate the design loads into space requirements.

Economic efficiency is a prime motive for the design of shared-use, common use facilities in airport buildings. Facilities that several uses can share increase economic performance. They lead to greater rates of utilization and correspondingly lower costs per unit served than facilities designed to serve only one client or function. Moreover, shared-use space provides flexibility to meet unexpected and varying loads. Multifunctional facilities, such as "swing gates" that can serve both international and domestic passengers, have become increasingly common internationally (Table 15.1). For example, Toronto/Pearson has operated 15 swing gates, enabling management to switch 12 gates between U.S. "transborder" and Canadian domestic flights, and 3 between other international and Canadian flights.

Region	City/Airport	Region	City/Airport
Canada	Calgary	Asia	Osaka/Kansai
	Edmonton	Europe	Athens
	Montreal/Trudeau		Birmingham (U.K.)
	Toronto/Pearson	United States	Atlanta
New Zealand	Wellington		Denver/International

TABLE 15.1 Some Airports with International/Domestic Swing Gates

Two features of airport traffic drive the desirability of shared use:

- *Peaking* of traffic at different times
- *Uncertainty* in the level or type of traffic

The time over which these drivers take place—their *cycle time*—defines the type of sharing that is appropriate (Table 15.2). Moreover, each combination of driver and cycle time requires a distinct form of analysis. This section presents four cases to provide a comprehensive set of tools to define the desirability of shared use.

Primary Driver	Cycle Time	Examples	Use
Peaking at different times	Hours	Swing space	Sharing gate between flights
	Days	Swing gates	Sharing between international and domestic flights
Uncertainty in type of traffic	Days	Additional gates	for peaks in weather, etc.
	Years	Reserve gates	for uncertain future growth

Table 15.2 Primary Drivers and Periods Motivating Shared Use of Facilities

Peaking of Traffic

Shared use of facilities is economical when distinct parts of the traffic peak at different times. This is because the facility required for the peak of traffic A could be used for traffic B as the traffic A drops and traffic B peaks. Shared use reduces the size or number of facilities that the airport needs to provide for a given total traffic, and thus increases productivity and the return on investment.

This driver of shared use has nothing to do with uncertainty. This observation is important: It means that the analyses and resulting consequences appropriate to peaking are quite different from those associated with uncertainty.

To make it possible to share facilities between operations that peak at different times, airport operators must be able to "swing" the facility from one use to another. Designers make this possible by building in the features necessary to implement such swings in use. For example, to enable shared lounge space, they have to create *swing space*, that is, joint lounges that can serve several gates instead of individual gate lounges separated from each other by walls or other barriers. To permit sharing of departure gates, designers have to build corridors that connect these gates with the appropriate users. Specifically, when regulations require the segregation of users (as is typical for international and domestic operations), designers must provide corridors through a system of doors that the operators can securely and reliably open and close as needed. For instance, Atlanta equipped international gates with electronic switches that immigration authorities use to control the opening of doors and the flow of arriving passengers.

The cycle time between the distinct peaks of traffic influences both the types of analyses and the design of the shared facilities. Two intervals are salient regarding peaking:

- *Hours*: The peaks occur in the range of about 1 hour, as for example in the case of passengers waiting to board three to four aircraft leaving 10 to 20 minutes apart.
- *Days*: The peaks occur at widely separate times of day, as for the peaks of international and domestic traffic at many airports.

When the interval between different peaks is in the range of 1 hour, two consequences arise. First, the traffic flows associated with the different peaks interact. Second, it is impractical to separate these flows easily. The interaction means that the advantage of sharing the facilities is only a fraction of the total. It also implies that arrangements requiring substantial operator intervention, such as the opening and closing of secure doors, are difficult to implement.

When the interval between peaks occurs over a much longer period, however, it may be possible to dedicate a facility to an alternative use for a portion of that time. For example, a swing gate can serve international traffic during its peak and domestic traffic at another time. Alternatively, airlines can also "swing" the designation of an aircraft from a domestic arrival to an international departure or vice versa. This could happen when a Star Alliance United Airlines flight arrives domestically at Chicago/O'Hare from St. Louis and proceeds on to Frankfurt, Germany, as a code-shared Lufthansa international flight, or the other way around. Airlines thereby avoid the cost and delay of towing aircraft. Over longer periods, airport operators can also accomplish the tasks necessary to create alternative secure paths to the gate, as when regulations require the separation of domestic and international traffic.

Uncertainty of Traffic

Uncertainty in the levels of traffic is the other principal driver motivating shared use. The issues in this case are that

- Additional facilities are required to "buffer" the system against peaks greater than the scheduled peaks. These extra peaks arise through either short-term delays or uncertainties in long-term requirements.
- The efficient size of the buffer depends on either the frequency of the peaks or the range of the uncertainty in the long-term needs.

- Thus, it is economical to provide these “buffers” jointly for several users, rather than individually for each user.

Shared use of the buffer facilities leads to savings because the peak needs for the entire system are normally considerably less than the sum of the possible peaks for each element of the system. This is because traffic drops for some users counterbalance the peaks of other users and smooth the variations in the overall traffic. This fact may not reduce the maximum possible peak, such as might occur when all users suffer delays during a major storm, but it does reduce the frequency of peak loads. The economically efficient size of the buffer represents a tradeoff between the cost of the buffer and the cost and inconvenience of not having extra space when needed. Therefore, reduced frequency of need reduces the economically efficient size of the buffer space.

In the long term, over many years, the total requirements for a specified level of total traffic are also less than the sum of the anticipated possible maxima for individual uses. In this case, this is because some users do not meet the forecasts. When this happens, the airport manager can reallocate the space reserved for future use to the users that then actually require it.

To enable the economies of shared buffer space, designers should physically place this capacity between the core facilities of major users, so they can easily use it when needed. This implies that the airport passenger buildings should be connected rather than independent, as they are in a “unit terminal” configuration. At Minneapolis/St. Paul, for example, all the buildings connect with each other, so it is easy in principle for one airline to use additional gates when its neighbor does not require them. On the other hand, the separate terminals at Boston/Logan impede the reallocations of space as the local needs of one airline grow while those of some other airline shrink.

The timing of the uncertainty influences both the types of analyses and the design of the shared facilities, as it does for the peaking factor. Two types of interval appear meaningful regarding uncertainty:

- *Days*: The uncertainty arises from operational factors such as mechanical and weather delays, and is resolved over a few days.
- *Years*: The uncertainty is about the level of future operations for different users, and may only be resolved over years.

When the uncertainty is resolved over days, the analysis is conceptually simple. Its essence lies in a tradeoff between the cost of the additional facilities and the costs associated with delays and schedule disruptions that arise when facilities are not available when needed. However, airport designers have no credible basis for estimating the costs of disrupted schedules for airlines years in the future, so it is impossible in practice to calculate these tradeoffs accurately. Fortunately, exact calculations are unnecessary. Because the base requirements at any airport change constantly with the level of traffic, so does the size of the remaining buffer capacity. In this circumstance, the analyses and rules of thumb cited in the next section are adequate.

The analysis is more complex when the uncertainty covers several years. This is because designers need to trade off the cost of constructing flexible buffer facilities now against benefits far in the future that need to be properly discounted. A flexibility analysis examines the possibilities. The discussion under "Uncertainty, Long-Term Variation," at the end of the next section explains this procedure.

15.3 Analysis for Shared Use

The appropriate analyses for shared use are different for each of the four major cases discussed previously and indicated in Table 15.2. These are associated with each of the two drivers of shared use and shorter and longer periods.

Peaking, Hourly Variation

A prime example of this situation is shared use of gate waiting areas. Other facilities, such as check-in counters, baggage handling, and security checks, operate under similar load patterns. In all these cases, the total space required for all users is less than that needed by the sum of several classes of users, because their peaks occur at different times. In practice, a shared-use departure lounge is simply a large room serving several gates (Fig. 15.2).

FIGURE 15.2 Shared-use gate lounges at Washington/Reagan. (*Source:* Zale Anis and the Volpe National Transportation Systems Center.)

The analysis determines the maximum requirement by calculating the sum of the fluctuating needs of many individual users. It tracks the dynamics of how the traffic builds up and abates for each user, and how these flows add up overall. The procedure divides the period of interest into many smaller intervals. Thus, if the focus were a bank of departures over 2 hours, the smaller intervals might be 5 or 10 minutes long. For each interval, the analysis calculates the total traffic in place, as derived from the cumulative flow patterns appropriate to this activity (see Chap. 20). To calculate the waiting room requirements, for example, the analysis calculates the number of passengers who have arrived for the flights for each small interval and takes away the number who have boarded. These numbers depend on the size of the aircraft, its load factor, the pattern of arrival of the passengers, the time between departures, and the boarding procedures.

Note carefully that the patterns of arrival of passengers vary by time of day and type of flight. For example, passengers arriving for early-morning flights tend to arrive much closer to departure time than passengers showing up for flights late in the evening. Likewise, passengers on international flights generally arrive much earlier than do passengers on domestic flights. Table 15.3, for Example 15.1, shows a tabular description of a cumulative arrival pattern.

Number of Passengers	Minutes before Departure of Aircraft						
	60	50	40	30	20	10	0
Arrived at gate	10	30	50	100	150	180	200
Boarded aircraft	0	0	0	0	10	140	200
In lounge	10	30	50	100	**140**	40	0

TABLE 15.3 Passengers into and out of a Waiting Lounge for a Single Aircraft

Spreadsheets provide a versatile and cost-effective way to analyze the effect of overlapping peaks. These computer programs easily handle the additions and subtractions needed to estimate the number of passengers who need space in the waiting room. Analysts can define entries in the spreadsheet cells parametrically and can vary these factors at will. They can also explore all kinds of combinations automatically by means of the "data table" function of spreadsheet programs (de Neufville and Belin, 2002). Analysts can develop an appropriate spreadsheet model for such situations in about a day.

Space sharing can cut requirements dramatically. The space needed depends mostly on the time between departures. The greater the time between departures, the more a preceding flight has time to empty space and make room for passengers for the next flight. If as little as 10 minutes separate the scheduled departure times, this is enough time to allow many passengers from earlier flights to leave the waiting rooms and provide space for passengers in the later flights. Example 15.1 illustrates the analysis, and Tables 15.3 and 15.4 present spreadsheet results showing the advantage of shared space.

Example 15.1 A simple case illustrates the procedure for calculating the savings due to shared use of lounge space and the results. Consider departure lounges serving 235 passenger aircraft operating at an 85 percent load factor and thus boarding 200 passengers. Assume that passengers arrive over an hour and begin to board the aircraft about 20 minutes before departure. Table 15.3 shows the number of passengers who have arrived at gate, boarded the aircraft, and—by subtraction—who remain in the departure lounge serving a single aircraft. For this case, designers would have to provide space for a maximum of 140 persons.

Suppose now that an airline operates a bank of departures at a shared lounge involving three aircraft, with the same pattern of arrivals and boarding procedures. Assuming that the aircraft leave at 10-minute intervals, Table 15.4 shows the number of passengers for each flight in the waiting lounge, and the cumulative number for all three flights. For this case of a shared lounge, designers would only have to provide space for a maximum of 290 passengers. This requirement for shared space is only 69 percent of the total for 420 passengers in three individual lounges!

Number of Passengers for	Minutes from Start of Bank of Departures								
	0	10	20	30	40	50	60	70	80
First flight	10	30	50	100	140	40	0	0	0
Second	0	10	30	50	100	140	40	0	0
Third	0	0	10	30	50	100	140	40	0
Total in lounge	10	40	90	180	**290**	280	180	40	0

TABLE 15.4 Passengers into and out of a Waiting Lounge for Three Aircraft Leaving at 10-minute Interval

It is important to understand how passengers share space over time. Figure 15.3 illustrates the dynamics of the process, in this case for passengers sharing a departure lounge serving four gates. The area starts filling up with passengers for the first flight. The number of travelers for this flight peaks and then diminishes as passengers start to board their flight. Meanwhile, passengers for the second, third, and eventually fourth flight arrive as the passengers for the first flight all leave, followed by those of the second, and so on. The shared-use space accommodates waves of traffic that peak and recede to leave space for the next wave.

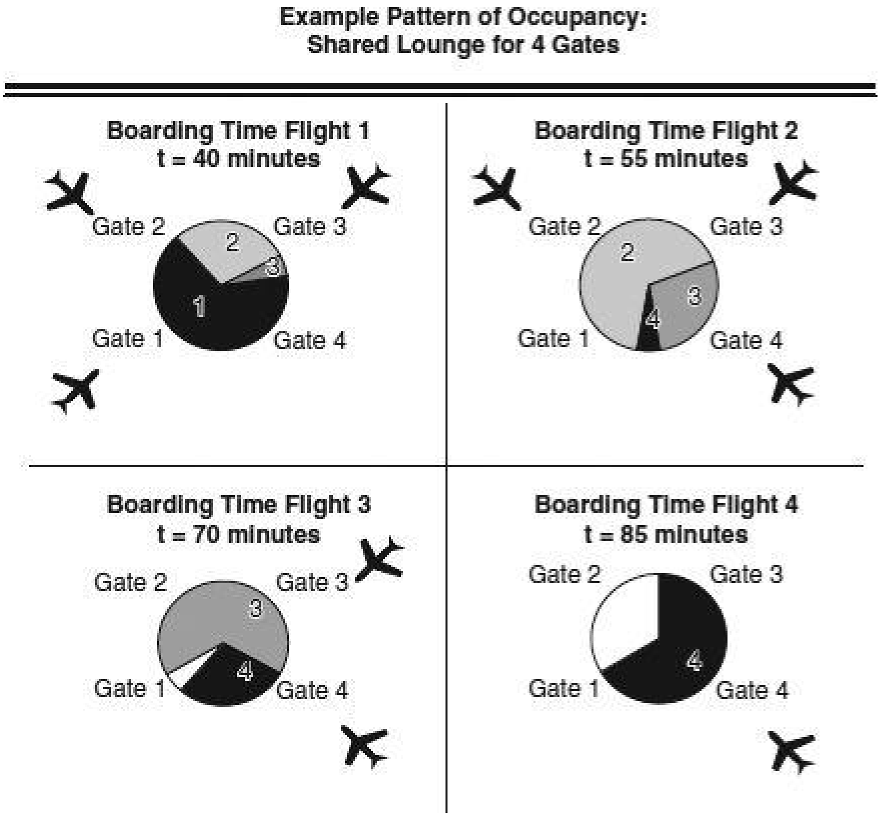
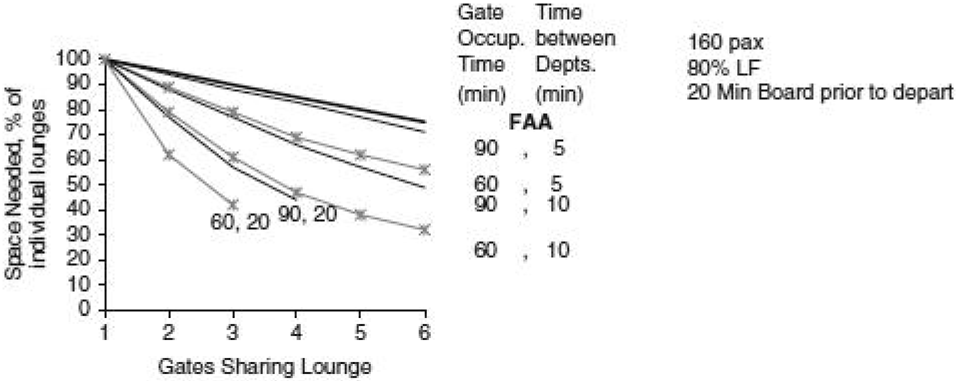


FIGURE 15.3 Schematic evolution of occupancy of departure lounge used by four aircraft. (*Source*: de Neufville and Belin, 2002.)

Sharing of lounge space can easily reduce the total size needed by 30 to 50 percent, as extensive analyses along the lines of Example 15.1 demonstrate. Table 15.5 and Fig. 15.4 show consolidated results for specific flight characteristics (aircraft size and load factor) with many combinations of operational characteristics (time available for boarding, time between aircraft departures, and time of gate occupancy).



A/C Size 160 Board time	20	FAA	60/5	60/10	60/20	90/5	90/10	90/20
	1	100	100	100	100	100	100	100
	2	95	89	79	62	94	88	77
	3	90	79	61	42	88	77	57
	4	85	69	47		83	66	44
	5	80	62	38		77	57	
	6	75	56	32		71	49	

FIGURE 15.4 Space required for departure lounges depends importantly on number of gates sharing the space, as well as on the time between departures and the size of aircraft. (*Source:* de Neufville and Belin, 2002.)

Flights (N)	Minutes between Flight Departures					
	0	3	6	9	12	15
2	100	94	87	81	75	70
3	100	87	75	64	56	48
4	100	81	63	51	42	37
5	100	76	54	41	34	n.a.
6	100	72	48	34	n.a.	n.a.
7	100	67	42	n.a.	n.a.	n.a.
In lounge	100	63	37	n.a.	n.a.	n.a.

*Analysis assumes 200-passenger aircraft, 60-minute occupancy at gate. "n.a." indicates impossible combinations of hourly frequency and time between departures.

Source: de Neufville and Belin, 2002.

TABLE 15.5 Shared-Use Lounge Space Needed for *N* Flights, as Percent of Space Needed for *N* Separate Lounges*

The more flights sharing the lounge space, the greater are the savings. The effect diminishes as the number of gates increases. As Fig. 15.5 indicates graphically, combining more than about six gates leads to relatively small additional improvements. Good designers recognize this pattern and usually have four to six gates sharing lounge space, unless constrained by space or local regulations.

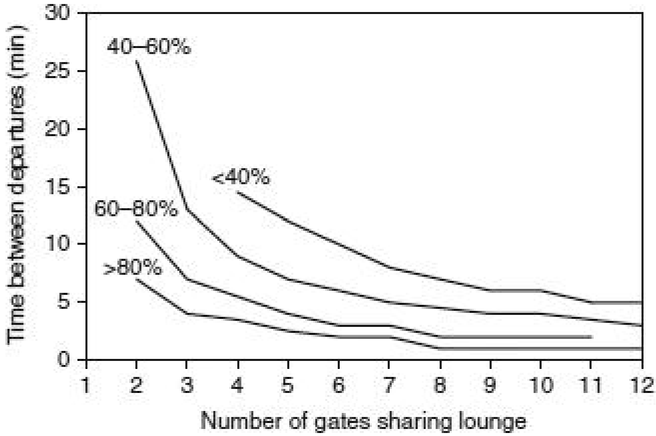


FIGURE 15.5 The percentage of space required for a shared-use departure lounge as a fraction of the amount needed for separate lounges depends on the time between departures and the number of gates sharing the space. (*Source:* de Neufville and Belin, 2002.)

The above analysis comes with an important warning. Being deterministic, it does not inherently take into account variations in the times between successive flights. Delays in departure shorten the time between some departures, and imply the need for more space in the shared lounge to accommodate passengers who cannot board on schedule. By giving results for different times between departures, Table 15.5 indicates the increase in space needed if these times change. The amount of extra space needed to provide a buffer to accommodate variations in times between departures depends on local conditions, such as weather and airline practices, that affect the frequency and distribution of the delays in boarding aircraft.

Peaking, Daily (or Longer) Variation

A prime instance of this situation is the use of gates. Different airlines or services will often exhibit distinct patterns of peaks over a day. Short-haul business traffic, for example, may have traffic peaks in the morning and evening. On the other hand, intercontinental services may have peaks determined by time zones, as when European flights arrive in New York in the early afternoon and leave later in the evening. Distinct peaks in traffic generally occur most significantly between international and domestic services.

The opportunity for sharing arises when peaks of different users do not overlap. Then one set of users can use facilities when other users do not need them. Airlines can share not only gates but also all the supporting facilities such as tugs and other vehicles for servicing aircraft, baggage services, and check-in kiosks.

Shared gates between international and domestic services require a carefully conceived system of doors and passageways that can channel passengers appropriately. Figure 15.6*a* and 15.6*b* illustrate how Edmonton arranges for shared space among three different types of traffic: Canadian domestic, transborder traffic that clears into the United States while still in Canada, and international traffic. Their solution is three-dimensional. In plan view, they separate holdrooms by movable walls to provide more or lesser space for the international and domestic traffic. In elevation, they use secure corridors to channel arriving and departing passengers.

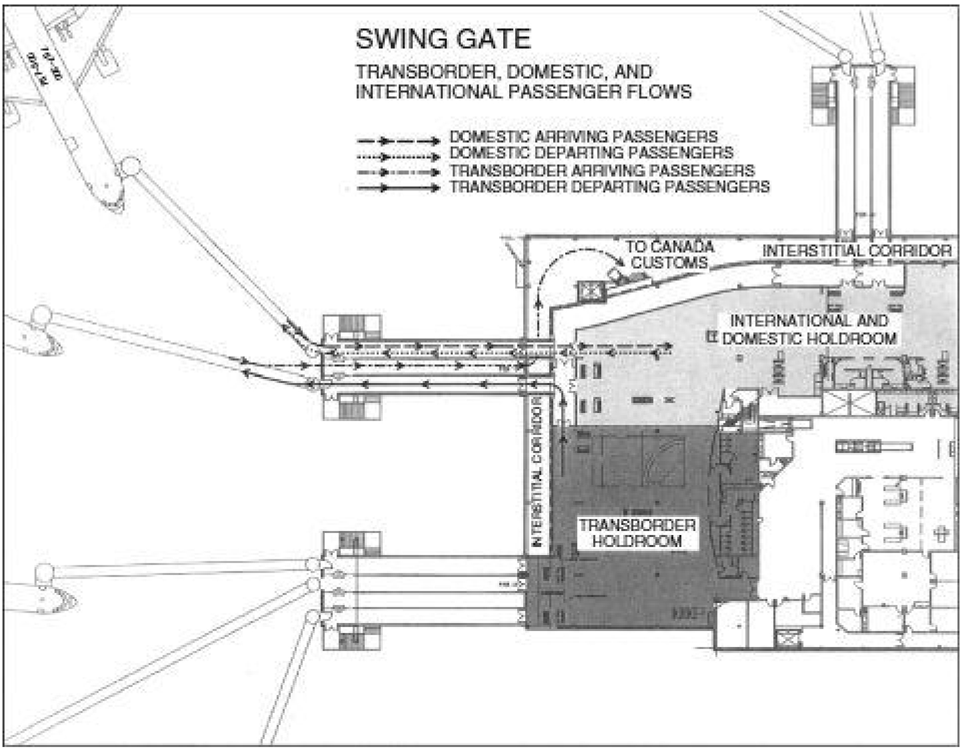


Figure 15.6*a* Shared-use space used by domestic, transborder, and international passengers at Edmonton. (*Source:* Lionel Oatway and Edmonton International Airport.)

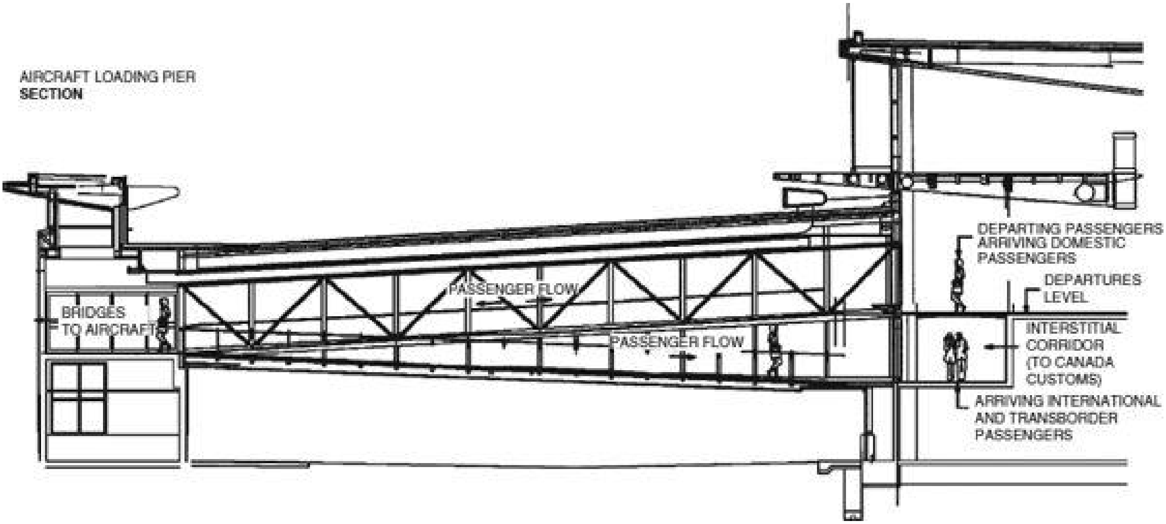


Figure 15.6*b* Shared-use space used by domestic, transborder, and international passengers at Edmonton. (*Source:* Lionel Oatway and Edmonton International Airport.)

Spreadsheets provide a versatile and easy way to analyze the possibilities for gate sharing. The analysis involves creating a table of the requirements for gates by time of day and type of use (airline, aircraft, type of service, etc.). Designers can obtain the total requirements of gates or other facilities separately, or for categories merged in various ways. They can thus easily test different forms of sharing. They can display the results in the Gantt charts that airlines use to plan and display their gate assignment schedules, as Fig. 15.7 shows. These can effectively communicate the advantages of sharing to airport managers.

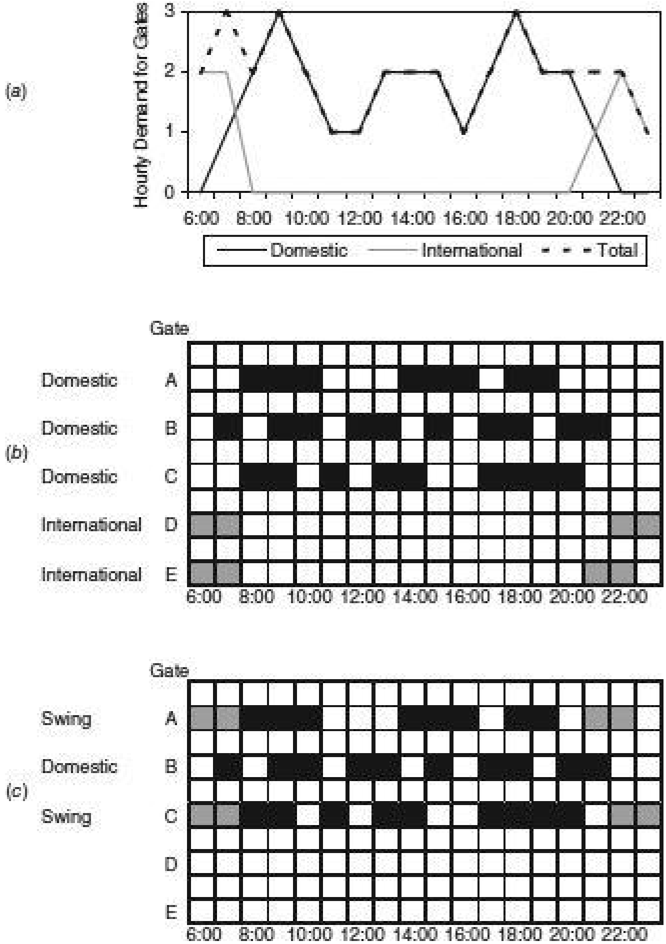


FIGURE 15.7 Analysis for the benefits of shared-use of gates: (*a*) hypothetical international and domestic operations; (*b*) Gantt chart showing gates needed for separate international and domestic gates; (*c*) Gantt chart for the number of shared gates. (*Source*: de Neufville and Belin, 2002.)

The savings from sharing gates can certainly be considerable. In Kenya, for instance, tourist flights from Europe tend to come and go at night, out of phase with the domestic flights that operate during the day. This offers the potential for great savings. Designers for Mombasa substituted a single building with shared gates for an original plan with separate international and domestic buildings. In this case, sharing gates between international and domestic services reduced the need for gates by about 35 percent, with corresponding savings in capital cost.

Uncertainty, Daily Variation

Weather, traffic, and mechanical delays lead to short-term delays. These drive the airlines to request additional gates, G^*, and backup facilities beyond their scheduled peak needs, G. They want the flexibility of this buffer space so that they can service their flights when late or delayed aircraft block gates scheduled for other flights. They know that they will need extra facilities to cope with random occurrences.

Probabilistic analyses can estimate the total amount of buffer space needed when the schedule requires more gates (Bandara and Wirasinghe, 1988, 1990; Hassounah and Steuart, 1993). A rule of thumb to estimate the approximate number of gates needed by any airline is

$$\text{Total gates required} = G + G^* = G + \sqrt{(G)} \tag{15.2}$$

This formula assumes that peak demand for gates during a busy period is a Poisson process. This is a reasonable approximation when considering 10 or more gates. The formula then exploits two facts: (1) The standard deviation of a Poisson random variable equals the square root of its mean; and (2) for a sufficiently large number of gates, an upside buffer equal to one standard deviation will cover about 85 percent of the random variations in the Poisson process. The formula leads designers to provide a total number of gates that should give immediate access to a gate for about 85 percent of the peak-hour flights.

Equation (15.2) reflects a most important phenomenon with important implications for design. The fraction of extra gates needed decreases as the number of scheduled gates increases. Thus

$$\frac{G^*}{G} = \frac{\sqrt{(G)}}{G} = \frac{1}{\sqrt{(G)}} \tag{15.3}$$

As G becomes larger, the size of the desirable buffer against uncertainty decreases. This is because random effects then tend to cancel out more. The practical consequence is that combining the requirements for individual airlines reduces the total number of gates needed. As airlines share their requirements, they increase G and reduce the relative size of the total buffer space, resulting in savings to all concerned. Example 15.2 illustrates this result.

This phenomenon provides a rationale for large-scale centralized management of gates at an airport. In the United States, the major airlines do this, because they are large enough by themselves to achieve the benefits of combining the need for gates. In the rest of the world, with some exceptions, airport operators assign gates to airlines and aircraft.

> **Example 15.2** If three airlines each provide for simultaneous peaks of 10 flights, they each independently need about $13 - (10+\sqrt{10})$, that is, 39 in all according to Eq. (15.2). If the airport defined their joint requirements, it would establish the overall number of gates at around $36 - (30+\sqrt{30})$ and obtain a savings of three gates around 8 percent.
>
> If the airlines were smaller, the savings would be greater. If six airlines each wanted five gates at the peak, the sum of their individual needs would be $45 - [6(5+\sqrt{5})]$ gates. The savings then achieved through joint use would be 9 = (45 − 36) or 20 percent. Gaudinat (1980) confirmed these results empirically at specific airports.

A convenient design solution places the extra facilities needed for uncertainty delays between major blocks of airlines or airline groups. This arrangement allows airlines to establish their brand at the airport and enables the airport operator to manage the overall facilities efficiently. For example, Toronto/Pearson allocates a core of gates and check-in facilities to airlines in proportion to their traffic. It places the balance of the facilities in a common pool from which it assigns positions at peak hours according to the varying needs.

The bottom line is that designers can reduce the number of gates by incorporating shared of gates, particularly among smaller airlines. They must balance this opportunity against both the management costs and passenger confusion associated with varying gate assignments.

Uncertainty, Long-Term Variation

The long-term issue is that the mix of traffic at an airport varies over the years. The future proportion of traffic represented by international passengers, by an airline, by a type of aircraft (e.g., narrow or wide body) will almost certainly neither be what it is, nor what it is forecasted to be at the time of design. As airport normally build facilities in anticipation of future demands, designers have to decide the proportions of facilities to create for each class of user, at a time when these are highly uncertain. Designers run the risk of getting the proportions wrong.

The uncertainty in the mix of traffic is the reason designers should create facilities that different kinds of traffic can share at different periods over the life of the project. This capacity needs to have the flexibility to serve the variety of future uses. The airport operator can then allocate this flexible space to the users who will need it in the future.

Flexible facilities that different users can share over the long term provide insurance against the risks associated with uncertain future needs. By enabling the airport to provide appropriate facilities easily available when needed, they remove the risk of avoidable extra new construction. Airports that fail to provide such insurance may end up wasting considerable space and resources. The experience of the Baltimore/Washington illustrates this phenomenon. The airport built a major international terminal principally for US Airways in the 1990s. However, that airline soon moved much of its international traffic to Philadelphia. Meanwhile, Southwest Airlines was growing rapidly and required more space. It would have been sensible to allocate the gates vacated by US Airways to Southwest. Unfortunately, the design of the international terminal was not flexible and did not permit the adaptation of that space to the needs of Southwest. The airport therefore could not fully use its existing facilities and wastefully had to build new gates (see ACRP, 2012).

The provision of facilities with the capability of serving several future purposes is known as *flexibility in design*. The flexible approach to design contrasts with traditional processes that focus on fixed designs to meet specific anticipated requirements. Whenever future needs are uncertain, as they certainly are for airports (see Chaps. 4 and 19), flexible design greatly increases the expected value of a system over its lifetime (see de Neufville and Scholtes, 2011).

Flexibility analysis permits designers to determine how much flexible space to include in the design. It provides the essential analysis comparing possible investments in shared-use facilities, with the prospective expected value of these investments over the scenario of future outcomes. It compares the immediate costs of implementing flexible design (that is of buying the insurance) with the future value of this insurance. In short, it is a risk analysis that explicitly considers the range of possible futures, their likelihood of occurring, and their consequences. This analysis of cost and value of possible levels of insurance defines the desirable level of insurance. In the case of airports, the analysis leads to the fundamental choice of the amount of capacity that will be flexible, that the airport operator will be able to swing from one use to another.

Decision analysis may be the best way for airports to determine the amount of long-term flexible space. This is because their choices are generally limited to a few sizes and few periods. The original decision node reflects the possible designs; the subsequent chance nodes reflect the mix of capacity actually needed at the end of the planning period; and the outcomes sum the extra cost of providing the swing capacity (in terms of sterile corridors and other necessary mechanisms) and the cost of meeting any shortfall in the capacity required for any particular use.

The analysis for Phase 1 of Bangkok/Suvarnabhumi illustrates the structure of the procedure (Fig. 15.8). The designers originally projected a capacity of 30 million annual passengers (MAP) but thought that international traffic (A) might contribute from 21 to 25 MAP, and the domestic traffic (B) from 5 to 9 MAP. As can be imagined, economical efficient design does not provide for the sum of the maximum possible for each use, 34 MAP (= 25 + 9). It should provide for a lesser amount consisting of some gates dedicated to each use, and swing gates that the airport can allocate to one use or the other. A possible solution in this case might have been to dedicate enough gates to serve 21 MAP of international traffic and 5 MAP of domestic traffic, with flexible gates in between that could serve up to 4 MAP or more of either traffic.

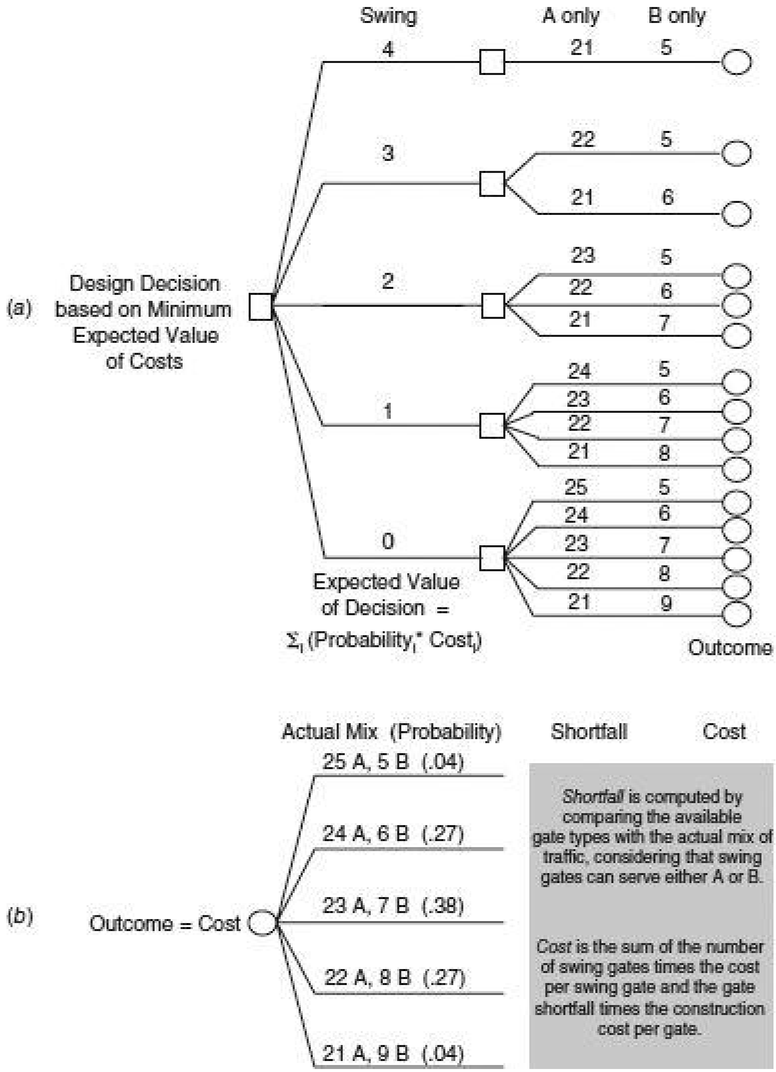


FIGURE 15.8 Decision trees for the analysis of the choice of the number of shared gates: (*a*) overall decision tree; (*b*) detail of one outcome node. (*Source:* de Neufville and Belin, 2002.)

A flexibility analysis should deal with normal and extraordinary variability. These are quite different:

- *Normal variability* refers to the routine variation around historical trends in the mix of traffic. For example, as the proportion of international traffic at Bangkok decreased from 80 to 71 percent in the 1990s, fluctuations occurred around this trend.
- *Extraordinary variability* is due to major shifts in the mix of traffic. These occur when operators radically reorient traffic, as United Airlines and AirAsia did when they quickly built up hub operations at Washington/Dulles and Kuala Lumpur respectively, or US Airways did when it pulled much of its international traffic out of Baltimore/Washington. Extraordinary variability implies much greater risk and provides the stronger motivation for shared facilities.

Analysts can estimate normal variability from historical records. The following illustrates how to do this. Consider an airport with two types of traffic, *A* and *B*, which could be international and domestic. Airport data over the years will allow the designer to calculate

- The past share of traffic in any year *t*, $A/(A + B)$
- The trend in that share over time using regression through the *n* years of data
- The standard deviation of this share around the trend

$$s = \left[\sum_{n} \frac{(\text{actual share in a year } t - \text{trend estimate of share for year } t)^2}{(n-1)} \right]^{0.5} \tag{15.4}$$

Assuming a Normal distribution of outcomes, these data permit an extrapolation to the end of the planning period of both the expected mix and likely range.

The probability of extraordinary variability is speculative. Analysts can only subjectively estimate the probability of a major change in traffic and its maximum shift. They might reasonably assume that whereas a congested airport such as Boston/Logan has little future as a transfer hub, an airport with large capacity and convenient runways might become one. For instance, there is a possibility that Orlando/International could become a hub at the expense of Miami. Conversely, Miami/International runs the risk of losing transfer traffic. The fact that these risks are subjective and impossible to estimate precisely does not mean that planners should neglect them. Good designers have a responsibility to provide some appropriate level of flexibility for dealing with these real risks.

To provide insurance against normal variability, the percent of shared use gates should be on the order of the standard deviation around historical trends calculated in Eq. (15.4). This is true so long as the additional relative cost of implementing swing gates is not too high. This makes sense intuitively. One standard deviation covers the bulk of the distribution of the risk, and it is natural that designers should provide less insurance as the risk diminishes. Figure 15.9 therefore is a reasonable general guide for conceptual planning.

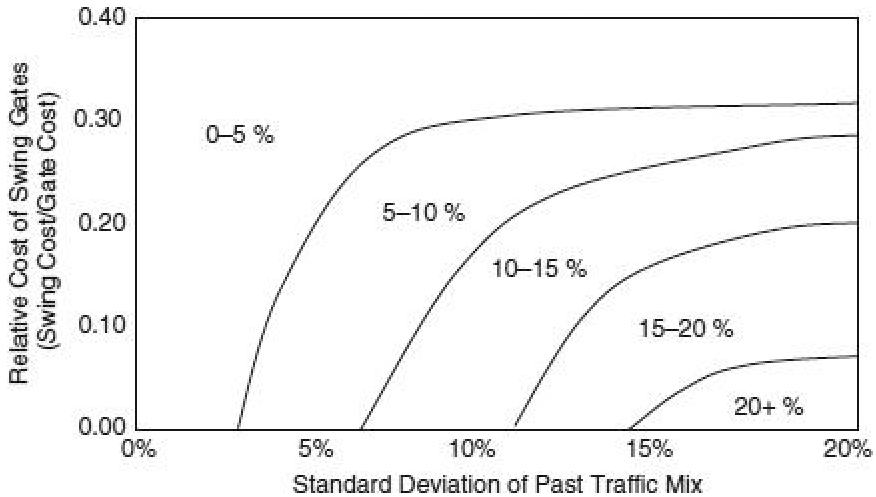


FIGURE 15.9 The proportion of swing gates to provide against normal variability depends on the extra cost of swing gates and the amount of variability. (*Source:* de Neufville and Belin, 2002.)

To provide insurance against extraordinary variability, the desirable percent of shared gates depends on the size of the possible shift, and the probability this may occur (Fig. 15.10). The upper limit is about the size of the possible shift. Thus, if the share of the international traffic might increase by 15 percent if an airport became an international transfer hub, designers might plan for as much as 15 percent of the gates to be international to domestic swing gates to provide the flexibility to deal with this extraordinary variability. The desirable percent of shared gates decreases as the probability of the increase is smaller, and the cost of the shared gates increases, as the comparison of Fig. 15.10*a* and 15.10*b* illustrates.

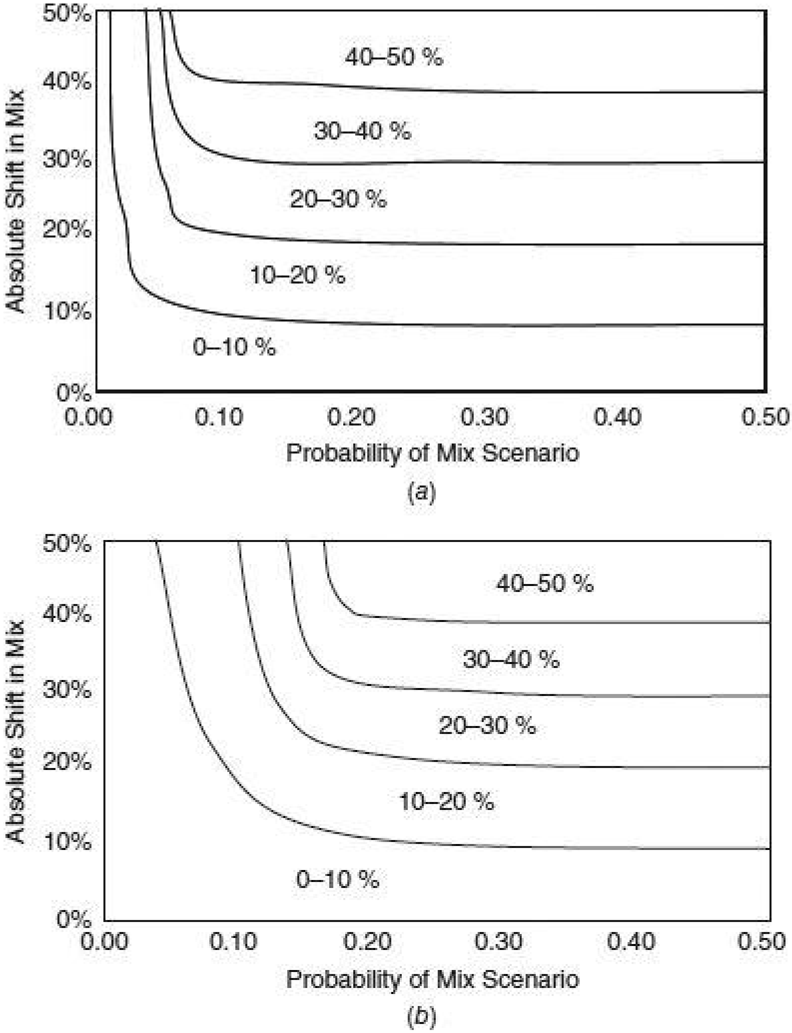


FIGURE 15.10 The proportion of swing gates to provide against extraordinary variability depends on the possible size of the shift and its probability as well as on the relative extra cost of swing gates: (*a*) lower cost, 5 percent extra; (*b*) higher cost, 20 percent extra. (*Source:* de Neufville and Belin, 2002.)

Overall Implications of Sharing

Common use, shared facilities are desirable because the airport passenger buildings are expensive. First-class terminal space easily costs over $\$3000/m^2$ or $\$350/ft^2$ to build,[1] and entails annual costs for cleaning, climate control, and maintenance. Responsible planners will incorporate facility sharing to the extent possible in their preliminary designs.

Planners should correspondingly incorporate the implications of sharing in their overall estimates of space requirements. These should recognize that some spaces will serve several purposes. They need to do this before they proceed to calculate the overall space requirements. Naturally, they should estimate the cumulative effect of the several different kinds of sharing. Example 15.3 illustrates how they can do this.

Example 15.3 Suppose that, for a specific project, planners have determined that they need to serve 2000 domestic and 1000 international departing passengers in their separate, distinct peak hours. This corresponds to 16 domestic and 8 international gates, assuming an average of 125 passengers per flight and a 1-hour turn-around time for aircraft.

Waiting room space: If the design provided separate waiting rooms for each flight, it would provide room for about 125 passengers per gate. However, it can reduce the need for waiting room space by clustering gate around a common waiting room. Supposing that they design modules of four aircraft gates and that they plan flight departures as close as 9-minute intervals at the peak period, they can estimate (referring to Table 15.5) that the space required for a shared lounge serving four gates is only about 50 percent of the amount required for individual lounges. This means that designers could size each module for half the maximum peak load for the four gates, that is, for 250 passengers.

Swing gates: Suppose that the peaks for domestic and international traffic do not fully overlap at this airport, so that one of the modules can serve international traffic during the peak for that traffic, and domestic traffic when that peaks. Designers could then provide five modules (20 gates in all), instead of the 6 that would seem to be needed by simple addition of the peak requirements without shared facilities (16 gates for domestic and 8 for international flights). Specifically, they could provide three modules exclusively for domestic service, one module exclusively for international service, and one module that could swing between domestic and international service as required.

Net result: The cumulative effect of sharing waiting rooms and gates is most significant. Instead of providing waiting room space for 3000 passengers (for the separate peak hours of the international and domestic services), the design consists of five waiting rooms for 250 passengers each, or space for 1250 passengers in all. Sharing in this case reduces the space requirement by almost 60 percent. This translates into enormous economies. Simply reducing the number of gates by four in this example might save about $40 million. The reduction in the size of each module leads to further savings.

15.4 Space Requirements for Waiting Areas

Standards exist to translate design loads for an activity to its space requirements. Those published by the International Air Transport Association (IATA) are used most widely for determining the areas needed for the several types of waiting areas in passenger buildings. These and any other standards represent subjective judgments about what is desirable. They are not scientific facts. The generally accepted standards change over time. Different organizations may use different standards.

The IATA definition of space requirements properly incorporates two features that reflect on how the airport operator plans to run the passenger building. Most obviously, it specifies standards according to “level of service.” It also subtly refers to operational practices, specifically to how fast the airport operator intends to move passengers through spaces. Specifically, it refers to the passengers' "time" in these spaces. Understanding these two issues is crucial for the correct use of IATA or any other space standards. The next two sections discuss these points in detail.

Importance of Level of Service

Level of service refers to the quality of the context in which a service takes place. Regarding passenger buildings, it refers specifically to the amount of space available for the activity. The idea is that the level of service is higher when passengers have more space. For example, passengers waiting for check-in might use larger or smaller spaces, for example 1 or 2 m^2 per person. In both cases, they could expect to receive their boarding passes. With the greater space, they will be able to move around easily and feel comfortable. With the smaller space, they will be squeezed and may feel uncomfortable.

Traditional practice uses six levels of service, from A (best) to F (worst).[2] Their explicit definitions are subjective (Table 15.6). They are also ambiguous, as some people may perceive an environment to be uncomfortably crowded when others may not. Notions of when people are too close for comfort are certainly personal and often cultural.

Level of Service	Description of Standard		
	Quality and Comfort	Flow Condition	Delays
A	Excellent	Free	None
B	High	Stable, steady	Very few
C	Good	Stable, steady	Acceptable
D	Adequate	Unstable, stop-and-go	Barely acceptable
E	Inadequate	Unstable, stop-and-go	Unacceptable
F	Unacceptable	Cross flows	Service breakdown

TABLE 15.6 Definition of Level of Service Standards

Note that the concept of level of service involves ideas about both the flows and the delays. As Chap. 20 indicates, these characteristics of any service system are tightly related. As the traffic increases toward the level of saturation of the server, both the delays and the variance of the delays increase, leading to increasingly unsteady flow. The descriptions of the several levels of service reflect this fact.

The level of service planners should provide for a facility depends on the performance objectives of the airport operator. More space costs more money for construction and subsequent operation and maintenance. The question is: How does the management of the passenger building want to balance economic efficiency (which leads to smaller spaces) and quality of service (which implies more space)? Airport operators differ significantly in this regard, for both the overall airport and specific facilities. For example, Singapore Airport has traditionally favored higher levels of service, consistent with their national objectives to make their city-state a premier location for business and tourists (recently they added a "budget terminal" with lower levels of service). Frankfurt/Hahn serves Ryanair, a low-cost airline, and offers lower levels of service. Aéroports de Paris provides a different level of service in its low-cost Terminal 3 than for Air France in Terminals 2.

Some airport operators have explicit ideas about the level of service they wish to provide. However, many airports have not defined their space standards and planners then have to use standard assumptions about the desirable levels of service. As these standards are not absolute, planners should feel free to adjust them up or down if the airport operator desires a more luxurious or a less expensive passenger building, respectively.

The standard assumption is that planners should design for level of service C for ordinary use. This implies that, on the worst days, the facility may fall to level of service D. Practitioners generally assume that level D is tolerable for short periods. Note that level of service in any area varies automatically as the number of people in it changes. Any facility therefore actually provides a distribution of service, as Fig. 15.11 suggests. These distributions can be estimated by simulation analyses, as Sec. 15.5 indicates.

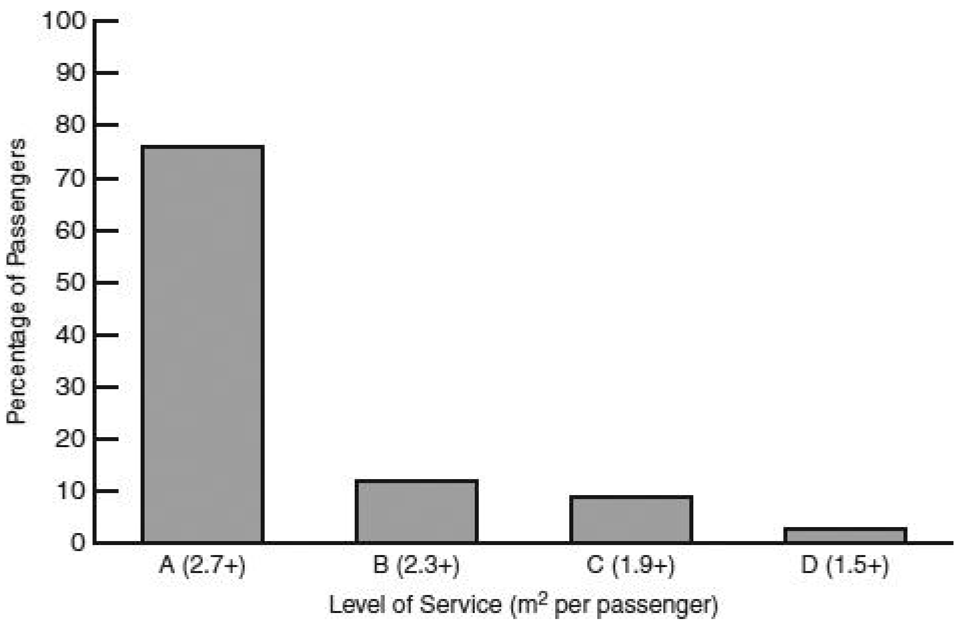


FIGURE 15.11 Distribution of level of service provided by a space over a period. (*Source:* Greater Toronto International Airport.)

Note further that if planners adopt level of service C for the design year, the level of service when the passenger building opens will probably be about A on average. This is because the design year is normally 10 years or more in the future, and anticipates growth of 30 percent or more in traffic. Thus, at opening the space per person will be over 30 percent greater than at design-year, and therefore at level A as Table 15.7 indicates.

		Level of Service Standard					
Activity	**Situation**	**A**	**B**	**C**	**D**	**E**	**F**
Waiting, circulating	Moving about freely	2.7	2.3	1.9	1.5	1.0	Less
Bag claim area	Moving with bags	2.0	1.8	1.6	1.4	1.2	Less
Check-in queues	Queued with bags	1.8	1.6	1.4	1.2	1.0	Less
Hold rooms	Queued without bags	1.4	1.2	1.0	0.8	0.6	Less

Source: IATA, 1995.

TABLE 15.7 Original Guidelines for Space to be Provided for Passengers in Different Functions (m^2/passenger)

It is useful to look carefully at the basic structure of the space standards shown in Table 15.7. Most obviously, better levels of service call for more space per person. Furthermore, the amount of space depends on whether passengers are moving or standing in place and on whether they have large bags with them or not. For any level of service, passengers standing in line without large bags take the least space. They require more space, 0.4 m^2 per person according to Table 15.7, when they are standing in line with large bags. They need even more space when they are moving around with their baggage, as in the bag claim area—according to Table 15.7, a further 0.2 m^2 per person at any level of service. The passengers need the most space when they are actively moving about. Using this understanding of the factors that define the space standards, planners can adapt the standards to situations that Table 15.7 does not specifically identify.

The 2004 IATA version of the space standards modifies the values in Table 15.7 in many subtle ways. However, it does not change in basic format: higher level of service entails more space, and passengers moving with bags require more space than those standing in line without their checked luggage. This latest version accounts for many current practices (such as the use of baggage trolleys). Designers needing details should consult this document or further revisions as they appear.

Importance of Dwell Time

Dwell time refers to the typical length of time passengers stay in an area waiting for service. It is a fundamental concept that needs to be carefully understood in order to use the space standards correctly. Failure to understand this concept has led to numerous design errors, as Example 15.4 indicates. The fundamental idea is that space is available over time, in units of time-area. Correspondingly, people use space over a specific time; they consume some amount of time-area according to the length of time they occupy an area. A given space offering a specific amount of time-area can then cater either to many people each staying a short time—provided their arrival is reasonably spread out—or to fewer people staying longer. Fruin and Benz (1984) and Benz (1986) developed this "time-space" concept.

Dwell time is important because it indicates how fast a space can be reused by another batch of passengers. The shorter the dwell time, the sooner a first group will leave the space, and another group can refill it and use the space again. Thus, if 1000 passengers per peak hour have to wait for 1 hour in a departure lounge, this space should be sized to fit these 1000 people. However, if the airport operator controls the traffic so that passengers wait on average only 30 minutes in the lounge, only 500 passengers use the space in the first half-hour and a second 500 use it in the next half-hour. When the dwell time drops in half—from 1 hour to 1/2 an hour—the amount of space needed also drops in half. The space needed is directly proportional to the dwell time.

Dwell time is implicitly incorporated in the space standards. Unfortunately, this fact is not immediately obvious. A superficial glance at the IATA standards seems to indicate that the space required for a design load of 1000 passengers per hour is simply 1000 times the figure in Table 15.7 appropriate for the activity and the design level of service. Proceeding on this basis is wrong, however, as Example 15.4 demonstrates. A dimensional analysis shows why this is so. The design load is in terms of (persons/hour) and the space standard in terms of (area/person). Multiplying these two factors gives a result in terms of (area/hour). To obtain the dimensionally correct answer in terms of (area), it is also necessary to factor in the dwell time, expressed in hours. The correct expression for estimating the area required for an activity is

$$\text{Area} = (\text{design persons/hour}) \times (\text{space standard m}^2/\text{person})(\text{dwell time in hours}) \qquad (15.5)$$

Example 15.4 This example, taken from an actual design review, shows how designers can misuse the IATA space standards by neglecting dwell time. It concerns the space to be provided for international arriving passengers to wait for passport control services. The design load was 2000 passengers in the peak hour. Naturally, these passengers do not arrive all at once; they filter into the space as each of their aircraft arrives and allows the passengers to disembark over about 15 minutes. The airport operator in this case had a design standard specifying that passengers should not wait more than 20 minutes for passport control. They could achieve such a standard by speeding up the process, most obviously by providing and staffing more passport control booths.

The question was: How large should the waiting space in front of the passport control be? The designer's original answer was to multiply the design load by the space standard for government inspection at level of service C, which was 1.0 m^2/person as in Table 15.7, to obtain 2000 m^2. This answer was wrong because it ignored the dwell time limit of 20 minutes, or 1/3 hour. The correct answer was

$$\text{Area for passport queue} = (2000 \text{ persons/h})(1.0 \text{ m}^2 \text{ standard}) \times \left(\frac{1}{3} \text{ hour dwell time}\right) = 667 \text{ m}^2$$

The mistake implied construction of 1333 m^2in unnecessary space. Assuming about \$3500 per m^2 for passenger buildings, this mistake would have cost close to \$5 million!

The space required for an activity in a passenger building depends on the management objectives of the airport operator in two ways. The airport operator can influence the space requirements not only by setting standards for level of service but also by defining the speed of service and the dwell times in specific areas. Example 15.5 illustrates how this can be done.

The ability of management to alter dwell time and thus to reduce space requirements is most important. Management can reduce space requirements by speeding up service in many ways. For instance,

- Airlines in the United States almost never weigh baggage on check-in, another U.S. practice that has been different from common European operations (see Chap. 3). This speeds up the check-in process by up to 30 seconds per passenger, or about 25 percent. The airlines can translate this faster service into either smaller areas in front of the check-in counters, along the lines of Example 15.4, or fewer check-in agents.
- Electronic kiosks similarly speed up the check-in process and reduce the dwell time and space requirements. Because they can also be situated almost anywhere, in parking garages or train stations, for instance, they further reduce the need for check-in space. Huge check-in halls may be obsolete; new ones may be much smaller.
- Low-cost airlines typically emphasize rapid turnarounds for their aircraft, on the order of half an hour compared to the hour or more usual for other airlines. Roughly speaking, such shorter dwell times double the capacity of their gates and halve their requirements for these facilities.

- Airport operators can reduce the size of waiting rooms for passengers by accelerating the boarding of aircraft. In Singapore, for example, the waiting rooms between the security check and the wide-body aircraft are often small and incapable of holding more than a fraction of the passengers destined for the aircraft. This design is successful, however, because the airport operator boards passengers practically as soon as they clear security and thus cuts the dwell time in the secure waiting room to only a few minutes.
- Governments can speed up and reduce the space required for passport control. For example, the U.S. Global Pass program pre-clears citizens for entry, which results in a faster process on arrival and shorter dwell times. Other countries are developing similar processes.

The conclusion to this section is that dwell time is an essential factor for determining space requirements. To estimate overall space requirements properly, planners need to consult not only with the airport operators but also the airlines and inspection services about how they intend to manage the dwell time of passengers in the different spaces.

Example 15.5 An airport operator planning the processing of arriving international passengers through passport control and the bag claim recognizes that passengers will somehow have to wait to get their bags. Bags will easily take 15 to 30 minutes to reach the claim area at a busy airport.

The airport operator can influence where this wait takes place. By slowing the passport control processes, say by having fewer control booths, the passengers will wait longer at the passport control—and then less in the bag claim area. What should the airport operator do?

Note that the space standard for government inspection is about 40 percent less than for bag claim areas. At level of service C, it is 1.0 m^2 instead of 1.6 m^2 according to Table 15.7. Therefore, the airport operator may achieve substantial savings in space if passengers wait at passport control. If the airport operator and the government agencies can manage the traffic cooperatively, passengers can arrive at the claim area at the same time as their bags.

Estimation of Areas

Once planners have developed an understanding of the management objectives of the airport operator concerning level of service and the operation of the facilities, the estimation of overall space requirements is easy. All analysts have to do is apply the relevant standards. It is the preliminary work to determine the appropriate level of service and the dwell times, which takes effort.

Note that the formulas for defining space needs, such as Eq. (15.5), can be reversed to provide estimates of the capacity of a space that already exists or to define its level of service under specific conditions. Example 15.6 shows how this can be done. Based on Eq. (15.5), the relevant formulas are

$$\text{(Capacity persons/h)} = \text{area}/[(\text{space standard, m}^2/\text{person}) \times (\text{dwell time in hours})] \quad (15.6)$$

$$(\text{Space, m}^2/\text{person}) = \text{area}/[(\text{design persons/h}) \times (\text{dwell time in hours})] \quad (15.7)$$

Example 15.6 A secure hold room for departing passengers dedicated to a single 300-person aircraft is 8 m by 25 m. What is its capacity at level of service C, if passengers flow through and have a dwell time of 30 minutes? What level of service does it afford if aircraft boarding is delayed and the dwell time increases to 1 hour?

Level of service C implies 1.0 m^2 per person. By Eq. (15.6)

$$\text{Capacity, persons/h} = \frac{(200\ \text{m}^2)}{[(1.0\ \text{m}^2)(1/2\ \text{h})]} = 400\ \text{persons/h}$$

The hold room thus has plenty of space under normal conditions. For the extraordinary conditions specified, by Eq. (15.7)

$$\text{Space per person} = \frac{(200\ \text{m}^2)}{[(300\ \text{persons})(1\ \text{hour dwell time})]} = 0.67\ \text{m}^2$$

The hold room would then offer level of service E, which is unacceptable except under very unusual circumstances.

15.5 Space Requirements for Passageways

The analysis for determining the size of corridors and stairways is similar to that for determining the size of gate lounges or other areas in which people spend time. It features level of service and considerations of dwell time. However, the formulas are quite different.

In designing passageways, width is the dimension that designers must specify. Width is a prime determinant of the capacity of passageways. All the formulas for corridors focus on this dimension.

The parameter that determines the capacity of a corridor is the rate of flow. The convention is to specify it in terms of the width of the passage. The units are “persons per unit width per unit time,” specifically, persons/meter/minute (PMM). As with the space standards for waiting rooms and similar activities, higher levels of service imply less congestion and less capacity for a given design. We can specify space standards in terms of PMM, as in Table 15.8, which represents the authors' recommendations. These recognize that airport pedestrians routinely haul baggage using trolleys or wheelies. To account for this fact, the numbers in Table 15.8 are half those used for pedestrians without baggage (Fruin, 1971). Experience indicates that this adjustment is a reasonable approximation.

Type of Passageway	Speed of Walking	Level of Service					
		A	B	C	D	E	F
Corridor	Regular	10	12.5	20	28	37	More
Stairs	Slower	8	10	12.5	20	28	More

Source: Modified from Fruin, 1971, to account for airport realities.

Table 15.8 Level of Service Standards for Passageways in terms of PMM (passengers/meter of effective width/minute)

The Formulas

The capacity of a corridor for any level of service is simply the standard times the "effective width," a concept explained in the next section. Note carefully that analysts should multiply the result by 60 to give the capacity in hours, to compare to standard design loads specified in terms of passengers/hour. Thus

$$\text{Corridor capacity per hour} = (\text{effective width}) \times (\text{level of service standard})(60) \tag{15.8}$$

Conversely, the effective width of corridor needed to carry a design flow is

$$\text{Effective width needed, meters} = \begin{array}{c}(\text{design flow/h})/[(\text{level of service standard})(60)] = \\ (\text{design flow/min})/(\text{level of service standard})\end{array} \tag{15.9}$$

The capacity of corridors is very large. For example, if a corridor has 3 m (10 ft) available for traffic flowing in one direction, its capacity at level of service C is

$$\text{Capacity (persons/h)} = (3\text{ m effective width}) \times (20\text{ persons/m/min})(60\text{ minutes}) = 3600$$

The great capacity of passageways needs emphasis. Designers often fail to grasp this point. They commonly plan corridors to be much wider than necessary. This can be very expensive. Corridors in passenger buildings extend over many hundreds of meters, so a couple of unnecessary extra meters of width can imply millions of dollars in wasted construction cost or space that could be better utilized. Hathaway (1999), for example, demonstrated "savings of approximately $10 million through reduction of corridor sizing" at Washington/Reagan. His use of simulation to do this is a good example of the value of these analyses (see Sec. 16.5).

Because designers frequently oversize corridors, airport operators often convert them into unplanned retail areas with pushcarts, storage space for miscellaneous facilities, and extensions of stores. All these activities may deserve space in the passenger building. Better design would provide for these activities directly rather than letting them randomly take over unused corridor space.

The capacity of corridors is much greater than that of comparably sized waiting rooms and other spaces because people pass through corridors quickly and their dwell times are short. Example 15.7 illustrates this phenomenon. This fact provides the analytic basis for the common-sense observation that the capacity of stairs, up or down, is much less than that of corridors. People move more slowly on stairs, the dwell time is longer, and the capacity is less.

An important design detail results from the difference in capacity between corridors and stairs. A stream of travelers going down a corridor will slow down and may form a queue when they reach stairs (Fig. 15.12). Designers need to anticipate this queue and provide space for it. Importantly, they need to consider carefully situations in which intersecting corridors meet in front of a staircase. The queue backing up in front of the stairs can easily block cross-traffic from the intersecting corridors and severely degrade the performance of that section of the building.

FIGURE 15.12 Jam of pedestrians around a stairway at Dallas/Ft. Worth. (*Source:* Harley Moore.)

Example 15.7 Consider a space that is 3 m wide and 50 m long. If this area is a lobby for people to circulate in for an hour, its capacity at level of service C that requires 1.9 m^2 per person is

$$\text{Capacity as lobby} = \frac{(150 \text{ m}^2 \text{ area})}{(1.9 \text{ m}^2)(1 \text{ hour dwell time})} = 79 \text{ persons}$$

Now consider this same space as a corridor through which people walk lengthwise. If they move at 3.2 km/h (2 mph), they are going at the rate of 53 m/min and will traverse the corridor in about 1 minute. The dwell time is then about 1/60th of an hour. Under these conditions, the capacity of this space as a corridor is

$$\text{Capacity as a corridor} = \frac{(150 \text{ m}^2 \text{ area})}{[(1.9 \text{ m}^2)(1/60 \text{ hour})]} = 5260 \text{ persons/h!}$$

Because the effective width (see below) of a 5-m corridor is 3.5 m once we deduct 1.5 m to allow for edge effects, the throughput of this corridor is about

$$\text{Corridor flow} = \frac{[(5260)/60]}{3.5} = 23 \text{ PMM}$$

This result is close to the value for level of service C in passageways shown in Table 15.8, and thus justifies these standards.

Effective Width

The *effective width* of a passageway is the width that is actually available to pedestrians. It is critical in determining the capacity of a passageway. The central idea is that, as a practical reality, pedestrians do not use some of the geometric width of the passageway. They avoid the edges and stay away from people coming in the opposite direction.

To obtain the effective width of a passageway, we must reduce the geometric width by the following three elements:

1. *Edge effects*: To reflect the fact that pedestrians shy away from walls, we need to deduct 0.5 m for each side of the passageway, that is, 1 m in all for this factor.
2. *Counterflow effect*: As pedestrians also avoid oncoming traffic, we need to subtract another 0.5 m from the geometric width of the passageway to account for this.
3. *Obstacles*: We should also subtract the width of any obstacle intruding into the passageway. These obstacles could include video monitors that attract a cluster of passengers looking, vending machines and shopping stands, etc.

The effective width of a passageway for pedestrians is thus a minimum of 1.5 m less than the geometrical width. To use the standards in Table 15.8 correctly, analysts should be sure to use the actual effective width of the corridor. Example 15.8 illustrates the calculation.

Designers should recognize that actual corridors in passenger buildings are usually wider than the minimum amount obtained from the calculations defined in this section. For esthetic reasons, architects will want wide passageways. For functional reasons, the airport management may also require wider corridors to allow for moving sidewalks or electric vehicles to convey disabled travelers. In short, passageways for passengers in airport buildings will almost invariably be much more spacious than what might be economically preferable (Seneviratne and Wirasinghe, 1989). Planners should accept this reality, while being aware of the danger of making corridors far wider than they need to be.

Intersecting flows of traffic provide an exception to this discussion. Managing crossing flows can be difficult. People slow down and stop, to avoid bumping into each other. This increases their dwell time and reduces the capacity of the space. Areas with streams of pedestrians crossing each other should recognize this phenomenon and allow extra space to account for it.

Example 15.8 What corridor width is needed to handle peak design loads of 600 passengers per quarter-hour, in each direction? Note that these peak flows of 1200 persons per quarter-hour correspond to about 4000 persons per hour, assuming that peaks over short intervals do not continue for the full hour. This flow implies about 10 to 12 million total passengers a year.

The design flow of 1200 per quarter-hour equals 80 persons per minute. From Table 15.8, the tolerable rate of flow is 20 persons/m/min for level of service C. The total required width is, using Eq. (15.9)

$$\text{Required width} = \text{effective width} + 1.5\ \text{m}$$
$$= \left[\frac{(\text{design flow/min})}{(20)}\right] + 1.5 = \left(\frac{80}{20}\right) + 1.5 = 5.5\ \text{m}$$

15.6 Areas for Baggage Handling and Mechanical Systems

Planners need to provide the right kind and size of space for baggage handling right from the start. Because bag systems are buried under the rest of the passenger building, it is generally difficult to expand them horizontally and probably impossible to do so vertically. Failure to provide adequate space for bag systems has led to extensive cost overruns, difficult workaround solutions, and excessive operating costs. The most glaring example of these difficulties occurred at Denver/International in the 1990s. In that case, a combination of planning failures resulted in a 17-month delay in the opening of the airport, the installation of a second baggage handling system, and numerous unanticipated operational difficulties. The delays alone cost $30 million a month in interest and other payments. All together, inadequate arrangements for the baggage handling system resulted in over $500 million dollars of extra cost and a 15 percent increase in airport costs (de Neufville, 1994; Dempsey et al., 1996).

Planners likewise should provide for the range of mechanical systems that are essential for the operation of a modern airport building. These systems include the following:

- Security devices for 100 percent screening of bags
- Heating, air conditioning, and ventilation ducts
- Water and sewer pipelines
- Electric substations and transformers
- Telecommunication lines
- Elevators and lifts
- And possibly right-of-ways for people movers, for use immediately or in the future

Planning adequately for these life-support systems of the building will save substantial costs for many years. The experience of the Port Authority responsible for New York/Newark Liberty airport provides a good example. The Authority designed its passenger buildings with an aerial right-of-way for a people mover. When the Authority later exercised its option to build this system, it was able to do so with a minimum of disruption to ongoing operations (see Sec. 17.5). As airport operators will have to reconfigure their systems many times over the life of the building, they should plan them with flexibility in mind.

Space planning for baggage and mechanical systems differs fundamentally from planning for passengers. In planning for humans, it is sufficient to think in terms of areas—architects will make sure that floor-to-ceiling heights will be adequate for people. In planning for machines, however, planners must think in three dimensions. Baggage systems need enough height to cope with layers of pathways crossing each other. The conveyor or tilt trays that carry bags through the building need unobstructed space to bend, rise, and drop gradually as necessary for the proper operation of the baggage system. Planners need to allow both enough area for baggage and other mechanical systems, and sufficient height and pathways for the connecting lines.

Baggage space is frequently most problematic for two reasons. First, it is normally in confined areas broken up by large columns supporting the building. Second, the elements of a large bag system are inflexible. Designers of baggage systems frequently find themselves trying to squeeze hard objects into an inadequate space. This is not easy to do well. Failure to anticipate the needs of the baggage systems can result in convoluted, inefficient arrangements that degrade the performance of the airport passenger building.

Space is also vitally necessary for the maintenance of all the mechanical elements. If there is no reasonable access to the equipment, maintenance crews will neglect it. This will lead to long-term difficulties, costs, and degradation of service. Unfortunately, access space tends to disappear as the detail designers cope with installing baggage handling tracks and sorters. In designing baggage facilities, planners need to consult carefully with experts in operating this process—the airlines, the baggage handlers, and equipment manufacturers.

To compound these difficulties, the design of baggage systems is not standard. Major new facilities tend to have unique systems. The lack of standardization has several causes:

- New security regulations introduce scanning devices whose design and performance is evolving.
- Changing technology for handling bags through advances in information technology, laser readers, and radio-frequency identification (RFID) systems.
- Major industrial countries prefer their own equipment.

This variety of systems for screening and moving bags means that there are few rules about planning baggage systems at the conceptual stage.

The space needed for baggage systems depends crucially on the number of bags checked per person. This number varies at different locations. It is about one bag per person. It might be half that for business travelers on for short trips—or almost double that for vacation or other flights for which passengers bring a lot of luggage. Moreover, the average is changing as airlines increasingly impose fees for checking bags, and passengers tend to carry on more bags.

The size and nature of the bag systems depends on whether the airport is a transfer hub and requires extensive sorting mechanisms and space to store bags between flights. As for passengers, the space needed depends directly on the dwell time of the bags in the system. This may be relatively short where transfers are rapid, as they are at Dallas/Ft. Worth, or long where transfers may easily be much longer, as at London/Heathrow.

For preliminary planning purposes, the area of the room for outbound bags should be a minimum of about 0.5 m^2 (5 ft^2) per peak-hour passenger for simple baggage handling systems. It should be up to three times as large for systems using some form of automated sorting. The space needed for outbound bags obviously depends on the number handled.

The height of the room handling outbound bags should be at least about 2.5 m (8 ft) more than the height of the first obstruction, with additional space for conduits and utilities—for the simplest systems. Rooms that will accommodate automated baggage systems need about an additional 2 m (6 ft) for each layer of belts or trays involved in the bag system. The total height of a modern automated baggage room might thus be about 9 m (30 ft). It may need to be as high as two levels of ordinary floors!

Because of changing technology and lack of general rules, planners should consult with security and baggage handling specialists early. Making sure that there is enough space for the crucial baggage handling function will enhance the operation and efficiency of the airport passenger building.

15.7 Take-aways

Correct translation of future design loads into specification of the size of facilities in an airport passenger building is a subtle process. It is much more than a technical application of standards and formulas. Properly done, it involves knowledge of management objectives, understanding of site-specific patterns of traffic, and appreciation of operating procedures.

Management objectives set the tone. Their choice of the level of service they wish to provide, overall or for specific customers, directly affects the size of facilities. At many airports, the default level of service C is increasingly inappropriate.

Patterns of traffic provide the basis for effective sharing of facilities by different users. Designs that enable sharing of gates, waiting areas, and other facilities can significantly reduce total requirements for space and provide cost-effective design solutions. Spaces that can serve multiple users also provide much needed flexibility as insurance against future uncertainties.

Operating procedures influence the time spent in facilities, and this "dwell time" directly defines the need for space and the capacity of a given space. Designers need to consider this factor carefully, as its neglect is the source of frequent mistakes in sizing facilities.

Exercises

15.1. For some airport for which you can obtain data on traffic, estimate peak-hour design loads for a specific activity, such as check-in.

15.2. Consider a passenger building that you can visit or for which you have plans:

a. If it does not have shared-gate lounges, estimate how one might be implemented. To the best of your ability, referring to data comparable to Table 15.5, estimate the savings in space that might result.

b. If it does have shared-gate lounges, estimate the average number of passengers per departure and the total size of alternative waiting lounges for individual lounges for each gate. How does this with the space actually available in the shared-waiting lounge? Using a local estimate of the typical time between aircraft departures, how does this ratio compare with what would be calculated from data comparable to Table 15.5? In your estimation, is the shared lounge oversized? Undersized? Or just about right?

15.3. For a passenger building you can visit or for which you have sufficiently detailed plans, estimate the capacity of spaces dedicated to specific activities, such as check-in facilities or gate lounges. Using a best estimate of peak loads, do these spaces seem adequate? How do the theoretical calculations compare to the reality at the site? What are the discrepancies? What might account for them?

15.4. Do Exercise 15.3, focusing on the adequacy of the corridors and passageways.

References

ACRP, Airport Cooperative Research Program (2007) "Airport Aviation Activity Forecasting: A Synthesis of Current Practice," ACRP Synthesis 2, Transportation Research Board, Washington, DC. www.trb.org.

ACRP, Airport Cooperative Research Program (2012) *Addressing Uncertainty about Future Airport Activity Levels in Airport Decision-Making*, ACRP Report 76, Transportation Research Board, Washington, DC. www.trb.org.

ACRP, Airport Cooperative Research Program (2013) *Guidebook for Preparing Peak-Period and Operational Profiles*, ACRP Report 82 Transportation Research Board, Washington, DC. www.trb.org.

Bandara, S. and Wirasinghe, S. (1988) "Airport Gate Position Estimation under Uncertainty," *Transportation Research Record, 1199*, pp. 41–48.

Bandara, S. and Wirasinghe, S. (1990) "Airport Gate Position Estimation for Minimum Total Cost—Approximate Closed Form Solution," *Transportation Research B, 24*, pp. 287–297.

Benz, G. (1986) *Pedestrian Time-Space Concept: A New Approach to the Planning and Design of Pedestrian Facilities*, Parsons Brinckerhoff Quade & Douglas, New York.

Dempsey, P., Goetz, A., and Szyliowicz, J. (1996) *Denver International Airport, Lessons Learned*, McGraw-Hill, New York.

de Neufville, R. (1994) "The Baggage System at Denver: Prospects and Lessons," *Journal of Air Transport Management, 1*(2), Dec., pp. 229–236.

de Neufville, R. and Belin, S. (2002) "Airport Passenger Buildings: Efficiency through Shared Use of Facilities," *ASCE Journal of Transportation Engineering, 127*(3), pp. 201–210.

de Neufville, R. and Scholtes, S. (2011) *Flexibility in Engineering Design*, MIT Press, Cambridge, MA.

Federal Aviation Administration (1988) *Planning and Design Guidelines for Airport Terminal Facilities*, Advisory Circular 150/5360-13, US Government Printing Office, Washington, DC.

Fruin, J. (1971) *Pedestrian Planning and Design*, Metropolitan Association of Urban Designers and Environmental Planners, New York.

Fruin, J. and Benz, G. (1984) "Pedestrian Time-Space Concept for Analyzing Corners and Crosswalks," *Transportation Research Record 959*.

Gaudinat, D. (1980) "A Study of Gate Sharing at Airports: Boston as an Example,"

S.M. thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA. Hassounah, M. and Steuart, G. (1993) "Demand for Aircraft Gates," *Transportation Research Record, 1423*, pp. 26–33.

Hathaway, D. (1999) "Landside Simulations of Washington National Airport," in S. Mumayiz and P. Schonfeld (eds.), *Airport Modeling and Simulation*, pp. 253–264, ASCE, Reston, VA.

IATA, International Air Transport Association (1995) *Airport Development Reference Manual*, 8th ed., Montreal, Canada. IATA, International Air Transport Association (2004) *Airport Development Reference Manual*, 9th ed., Montreal, Canada. ICAO, International Civil Aviation Organization (2006) *Aerodrome Design Manual, Part 1, Runways*, 3d ed., Doc 9157P1, Montreal, Canada. Seneviratne, P. and Wirasinghe, C. (1989) “On the Optimum Width of Pedestrian Corridors,” *Transportation Planning and Technology, 13*(3), pp. 195–203.

---

[1]Construction costs depend on the site, the scale of construction, project details, the quality of the finish, and other factors. Terminal 5 at London/Heathrow cost about £4.5 or $7 billion.

[2]U. S. highway designers developed the notion of level of service. These roots account for the grading from A to F, parallel to the way U.S. teachers mark students from best to failure.

CHAPTER 16

Detailed Design of Passenger Buildings

This chapter addresses the detailed design of facilities within an overall configuration that either exists or will be built. It concerns design once the outer shell of the building and many of its interior spaces cannot be altered significantly. This assumption is often realistic even for buildings that have not yet been constructed. Country leaders, star architects, and financial backers will often have settled on an overall design, and designers of the details in the building will often not be able to challenge such agreements effectively. Thus

- The X-shaped configuration of the satellites for Kuala Lumpur/International was fixed once the president of Malaysia approved the initial layout.
- Despite the rapid differential settlements in soft soil that greatly complicated the design and construction of a glass roof about 1 km long, it was impossible for the engineers for the Osaka/Kansai passenger terminal to alter the design of the signature architect.
- The linear configuration for the Athens airport was an integral part of the financial package arranged by the build/operate/transfer consortium.

This chapter presents a process for design rather than specific standards or formulas. This is because there is no agreement in the airport industry about standards. These depend on cultural expectations, local practices, and the current objectives of the airport operator. Moreover, airport operators and their clients often want to achieve multiple, conflicting objectives, and will therefore compromise their standards to achieve a suitable balance. It is therefore pointless to look for a unique set of design standards. Section 16.1 addresses this issue.

The chapter emphasizes the need to identify "hot spots" that may seriously degrade the performance of the passenger building. Hot spots are concentrations of traffic in time and space that create bottlenecks that may define the capacity of all or part of the building. Experience indicates that failure to detect and resolve these situations during detailed design has been a primary source of difficulties for the effective and efficient operation of passenger buildings. Section 16.2 illustrates this issue and suggests how careful inspection of the design can identify hot spots.

To understand how hot spots arise, it is necessary to examine in detail the varying rate at which people arrive at potential points of congestion such as check-in counters, security points, doorways, and stairs. Section 16.3 shows how to do this using a graphical analysis of the cumulative arrivals and departures from a service area. The method enables the analysts to see how they might eliminate hot spots by adding servers, changing operational procedures, and altering design. This approach is effective for rapid analyses of specific bottlenecks.

Detailed simulation of the flow of passengers and bags is a prime way to explore the overall performance of the building and its facilities. It can be especially useful in showing how many different processes interact and sometimes create unexpected hot spots. Section 16.4 illustrates the effective use of simulation in the design of airport passenger buildings.

Finally, Sec. 16.5 provides information for the configuration and sizing of the range of specific facilities in a passenger building. These include check-in counters, security controls, waiting lounges, baggage claim areas, and curbside areas for leaving off and picking up passengers.

16.1 Design Standards

Design standards represent judgments about desirable service. They express the values of some organization. They are not scientific facts. Correspondingly, they differ between countries, as Chap. 3 indicates, and even within a country. A privatized commercial organization may be concerned primarily with economy and efficiency. Other objectives, such as the desire to build impressive gateways to a city or region, may motivate a public agency. Moreover, airport operators may apply different standards to different groups. Airports routinely apply higher standards for their national airlines than for foreign carriers, and for full-service rather than low-cost airlines.

Detailed standards for any particular part of the passenger building are not universal. Table 16.1 illustrates this fact by showing design standards various agencies have used for sizing departure lounges. At a general level, the standards were similar. They each provided seats for about half the passengers. In detail, however, the standards do imply different sizes for departure lounges for identical aircraft.

Source of Standard	Standard for Passengers m²/Passenger Seated	Standing	Percent Seated
Aéroports de Paris	1.5	1.0	50–75
Amsterdam	1.0	1.0	50
BAA	1.0	1.0	60
IATA	1.0–1.5	1.0–1.2	50

Source: Adapted from Ashford (1988).

TABLE 16.1 Illustrative Differences in Design Standards for Departure Lounges

Detailed standards often refer to time as well as space. Airport operators may specify how long they expect check-in, security and immigration checks, baggage claim, and other procedures to take. In this case, too, national standards may differ tremendously.

Note that there is a strong relationship between time and space standards. The faster the process, the less time people will be in it, and the less space it will require. (See the discussion of dwell time in Sec. 15.3.) This interaction depends critically on the service rate, that is, the speed with which a server (such as a check-in agent) can process people per unit of time.[1] As Chap. 20 indicates, the average number of persons served equals the product of the service rate, μ (pronounced "mu"), the number of available servers, n, over a specified interval, t:

$$\text{Number of persons served} = \mu n t \tag{16.1}$$

if all servers are operating continuously. Conversely, the minimum number of service positions required to serve a given number of persons in a period t, is

$$\text{Required minimum number of servers, } n = (\text{number of persons to be served})/\mu t \tag{16.2}$$

This formula defines the minimum number of servers because, when all servers are operating at capacity, the delays tend to increase indefinitely (see Chap.20, especially Sec. 20.4 and 20.6).

To achieve time standards, designers will provide enough servers to achieve the desired level of service (see Example 16.1). Note, however, that the actual operators of the service may thwart the intent of the airport owner or designer. Service times may exceed design standards because the operators do not staff the positions. Immigration authorities may not have enough personnel or airport operators may decide to save money by cutting services during weekends or holidays. Moreover, the service provider may simply have a different set of values than the designers. Low-cost airlines routinely staff a minimum number of check-in counters to save money, and expect passengers to wait.

Experienced airport operators and designers recognize the need to consider other factors besides space and time when evaluating the performance of passenger buildings. As Chap. 14 indicates, walking distance is a major determinant of preferable configurations of passenger buildings. The reliability of the service is also important: How often do long delays occur that result in major consequences such as missed flights and connections? Airport operators may be interested in parameters such as the maximum time people have to wait for a service, or the minimum time it takes transfer passengers or bags to connect from one flight to another. One of the advantages of properly executed simulations of passenger buildings is that they can explore these issues and suggest solutions (see Sec. 16.4).

Example 16.1 Consider an airline that has to check in 300 passengers in 50 minutes. Suppose it takes 1.5 minutes on average to serve a passenger, so its service rate is 0.67 passengers per minute. How many check-in counters should be provided? The answer is

$$n = \frac{300}{[(0.67)(50)]} = 9$$

Observe that the airline would need fewer agents if it reduced the average time their check-in agents took to serve passengers. This fact motivates efforts to speed up processes, for example by using mobile devices to read in passenger data.

16.2 Identification of Hot Spots

Successful passenger buildings enable travelers to flow through many different areas and processes on time and with enough space to meet expected levels of service. All too often, unfortunately, minor aspects of the airport building cause bottlenecks that jam a significant portion of the operation. Some process may clog up and restrict the flow—and capacity—of a major portion of the system. At one major European airport, an obstruction at a crucial point reduced the productivity of a new passenger building by about a third, implying extra capital expenses of about $100 million in current terms (de Neufville and Grillot, 1982). Generically, these obstructions are “hot spots”—specific points that cause major problems for the efficient operation of the passenger building.

Critical bottlenecks routinely occur even when designers have provided enough space overall for some activity. These hot spots arise subtly. They rarely result from gross errors in sizing spaces. They usually stem from some seemingly minor architectural detail or lack of understanding of operational procedures in the airport building. What happens is that airport users, following their natural instincts, concentrate at specific points at specific times. Example 16.2 illustrates how hot spots occur.

To identify potential hot spots in the design stage, it is necessary to focus on the dynamics of processes. We must think through in detail how the airport users will surge through the building over time. This analysis complements the validation of the detailed design to see that it meets overall standards that Chap. 15 describes.

The analysis for potential hot spots needs to be conducted with persons knowledgeable about airport operations and flows in many situations. This point deserves emphasis because the design teams for airports typically do not include operational experts. Architectural firms rarely have employees with long experience in airport operations; few have a steady stream of airport work to justify such staff. International airport consultants use local firms that know local codes but probably have little experience in airport design. Local airport employees with years of experience will be knowledgeable about that airport, but unlikely to have had the opportunity to learn how new airports operate in different situations or other countries. In short, designers need to recruit appropriate reviewers to help identify and avoid potential hot spots.

The review for points of critical congestion needs to follow passengers through the several processes they encounter. The analysts should carefully consider how the flows of different streams of traffic might interact—for example, the flows of arriving, departing, and transfer passengers. The designers need to focus attention on peak instances when traffic surges over short periods. These may occur over much smaller intervals than the peak hours of traffic. For example, the blockage in front of the check-in counters cited in Example 16.1 was an issue during the half hour following the opening of the counters for check-in. Designers need to focus on the pattern of cumulative arrivals of passengers at specific areas or processes, as Sec. 16.3 describes.

Example 16.2 Three instances of hot spots encountered at airports:

Check-in counters: Preliminary design arrayed check-in counters along one side of a 12-m (40-ft) wide corridor that served passengers going to gates farther to the left in Fig. 16.1. Designers intended to enable 1 h turnaround times at each gate, which implied that check-in for each flight would start 50 minutes before departure. However, most passengers arrive more than 1 hour ahead of flight time. They would flock to the check-in the moment it opened. The resulting concentration of people would completely block the passageway and prevent other passengers from reaching their gates. Management dealt with this problem by increasing turnaround times to about 1.5 hours, so that fewer people would be in front of the check-in counters. The productivity of each gate thus dropped from about 10 to 6 or 7 turns a day—a major and expensive loss of capacity (de Neufville and Grillot, 1982).

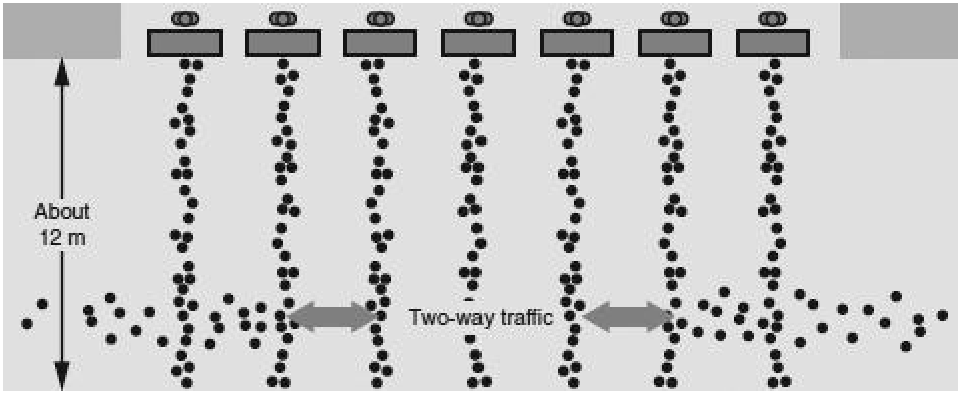


FIGURE 16.1 Check-in counters in a linear passenger building.

Secure passageways: A major passenger building features a secure corridor along the airside front. It channels international arriving passengers to immigration controls. Meanwhile, departing passengers cross the corridor to board their aircraft (Fig. 16.2). Problems occur whenever several major aircraft land in the same hour. The arriving passengers then occupy the corridor more or less continually. This congestion prevents the passengers in the waiting rooms from departing, because they should not mix with the arriving passengers who have not cleared customs. This situation substantially reduced the number of flights that could use the passenger building. The solution to this problem was to create a second immigration area for international arrivals, thereby freeing up space in the corridor.

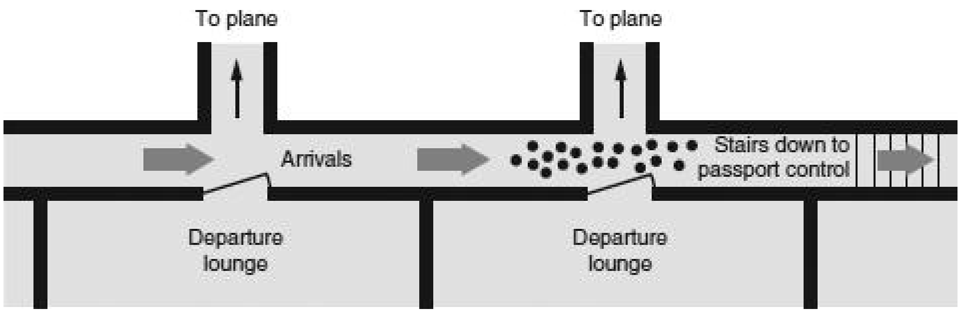


FIGURE 16.2 A secure corridor along the airside front of a passenger building.

Underground station: The platform in an underground train station serving a major airport provides enough space for a trainload of people. However, the detailed design placed a staircase in the middle of the platform (Fig. 16.3). Travelers descending to the platform with their bags naturally stopped at the bottom of the stairs and clustered there (area A). Few would either see or make the effort to get to the space behind the stairs (area B). As a result, too many people tried to board the train from area A, and many could not catch the train although the cars next to area B had lots of space. This situation reduced the practical capacity of the station.

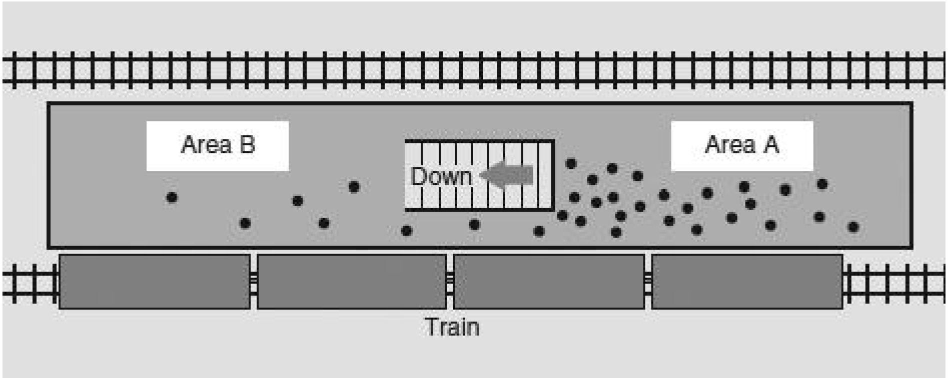


FIGURE 16.3 Underground train platform for a major airport.

The review should consider the psychological behavior of individuals and crowds. For example, people normally proceed to the first line of service they encounter, rather than turn aside to look for servers who have shorter lines. Coming down a corridor toward a row of security or border controls, for example, people will tend to cluster around those immediately in front of them, back up into the corridor, and block access to the other servers. Likewise, designers should be careful where they place passenger services such as flight information displays. The intuitive impulse is to place these services where they are visible to the most people. Unfortunately, the effect may be to create blockages just where traffic is most crowded.

16.3 Analysis of Possible Hot Spots

The cumulative arrival diagram is the basic element for the analysis of potential bottlenecks in passenger buildings. It represents the total number of arrivals over time on the vertical axis, keyed to a useful reference point in time on the horizontal axis (see Sec. 20.5). In the design of passenger facilities, it usually represents the number of persons who show up before some critical moment, such as the departure of a flight (Fig. 16.4).

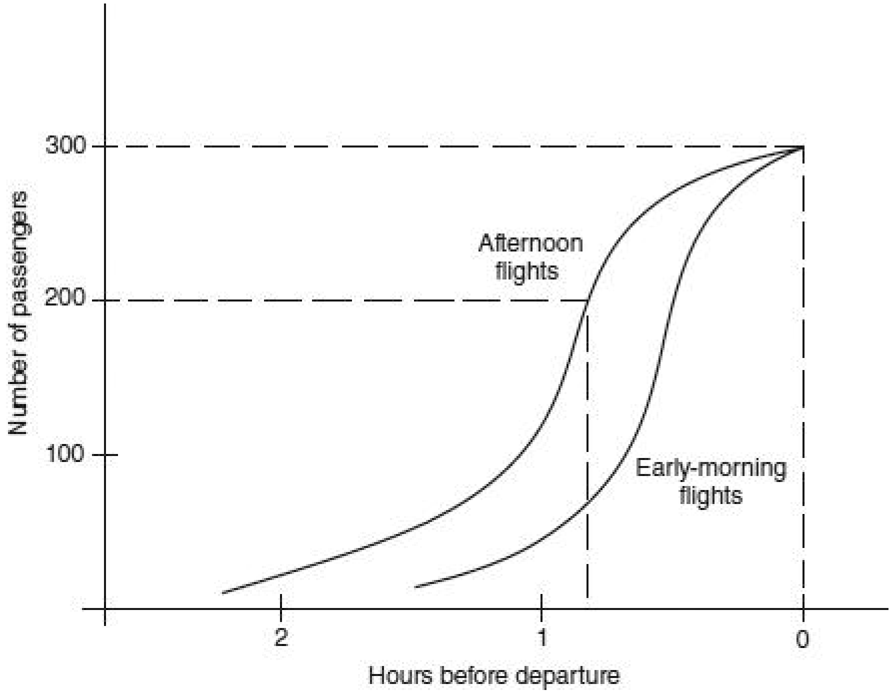


FIGURE 16.4 Arrival diagrams for a major European airport, for early-morning and afternoon departures.

Cumulative arrival diagrams represent an empirical phenomenon. Designers should base them on either observations or simulations of flows. They should expect different diagrams for distinct types of flights and locations. For example, the diagrams for departures early in the morning and in the middle of the afternoon for a major European airport have been quite different (see Fig. 16.4). People arrive at the airport for the afternoon flights as much as 3 hours early, for many reasons: their connecting flight is early, they want to avoid rush-hour traffic, or simply are anxious about the flight. For flights leaving at 7 a.m., however, people rarely arrive at 4 a.m.; they rise later and get to the airport closer to their flight time. In general, the cumulative arrival diagram for early-morning flights differs from those of later flights.

The many reasons for distinct cumulative arrival diagrams include the following:

- *International flights* tend to have passengers arriving much earlier than domestic flights, because of various additional controls.

- *Hourly flights* serving business travelers, as between New York and Washington, or Singapore and Kuala Lumpur, usually have most traffic arriving shortly before departure.

Cumulative diagrams also reflect the point at which they are taken. People do not flow directly through the airport from their point of arrival to their exit. They spend time in various airport facilities such as shops, restaurants, and other activities. As a rule, the earlier they arrive, the longer they spend in these ancillary activities. Therefore, the cumulative diagram for arrivals at the boarding gate is typically much more compressed than that for arrivals at the curb (Fig. 16.5).

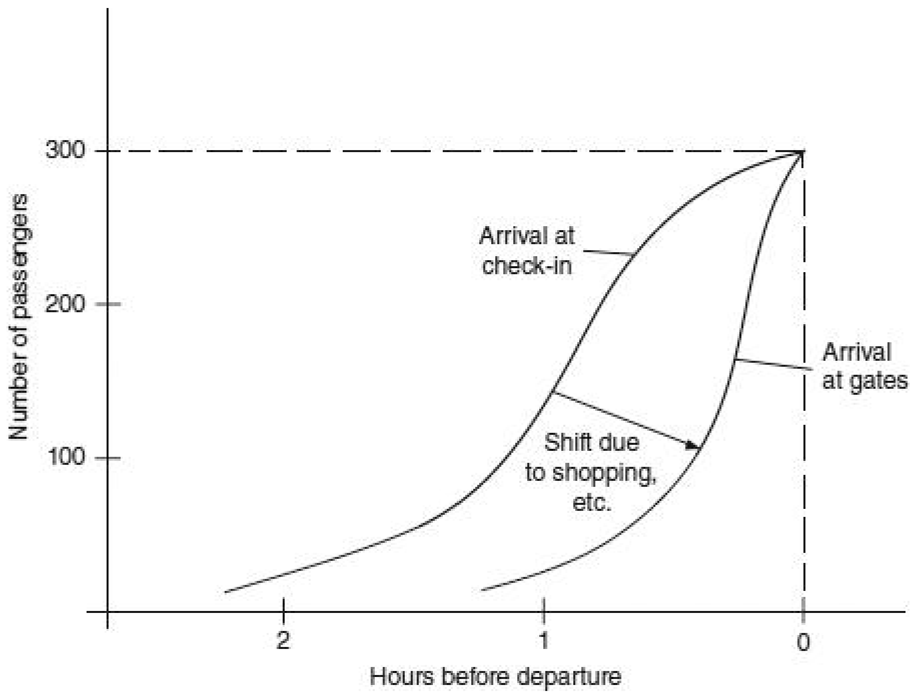


FIGURE 16.5 The cumulative diagram for arrivals at the boarding gate is typically much more compressed than that for arrivals at the curb.

The complementary cumulative departure diagrams represent the number of persons who complete a process over time. They look different than arrival diagrams because they feature straight lines (Fig. 16.6). These reflect the notion that whenever there is a queue of persons waiting for service, the average number of persons served in any period is a multiple of the rate of service per server and the number of servers, *n*, Eq. (16.1). The slope of any straight-line segment in the departure diagram then equals the number of servers then operating times their service rate. (If nobody is waiting for service, the service rate equals the arrival rate.) Furthermore, a departure diagram can indicate a definite starting time. Figure 16.6 indicates a service beginning 50 minutes before departure.

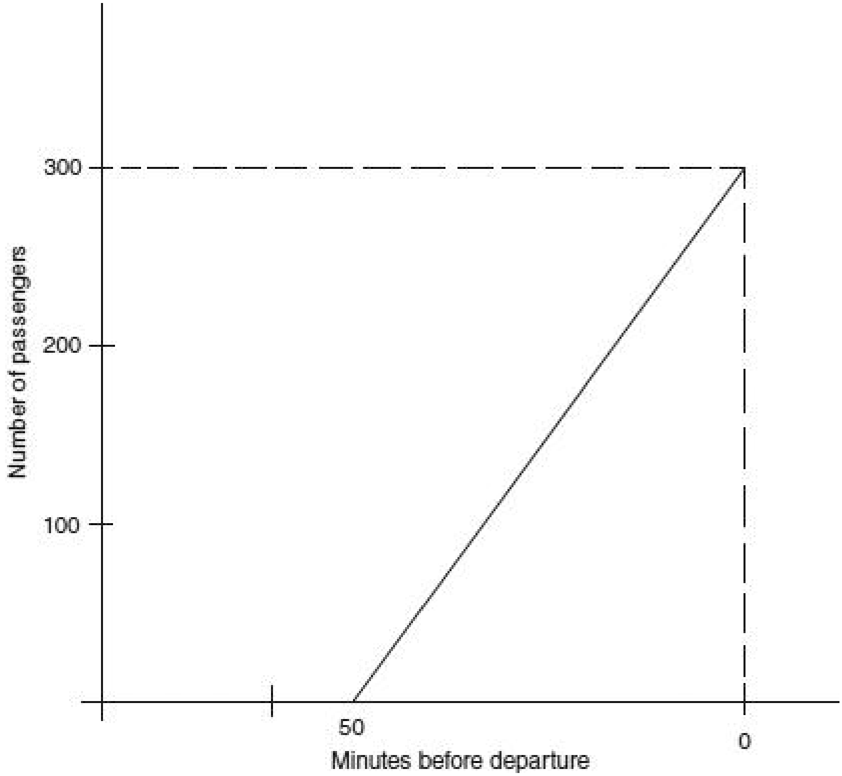


FIGURE 16.6 A cumulative departure diagram represents the number of persons who complete a process over time. It generally features straight lines.

Departure diagrams reflect management decisions about when to start service and how fast to run it. They can change daily according to short-term operational decisions. They differ conceptually from arrival diagrams, which describe long-term passenger patterns that airport managers cannot easily influence.

Designers can estimate the length of queues and the waiting time by combining the arrival and the departures diagram (Fig. 16.7). They do this by reading the differences between these two diagrams. At any point in time, the vertical distance between the two lines is the difference between the number of arrivals and departures, and therefore is the number waiting for service. Similarly, the horizontal distance is the average time between when a person arrives and is served, and indicates the average waiting time for an arrival at any particular time. (See Sec. 20.5.)

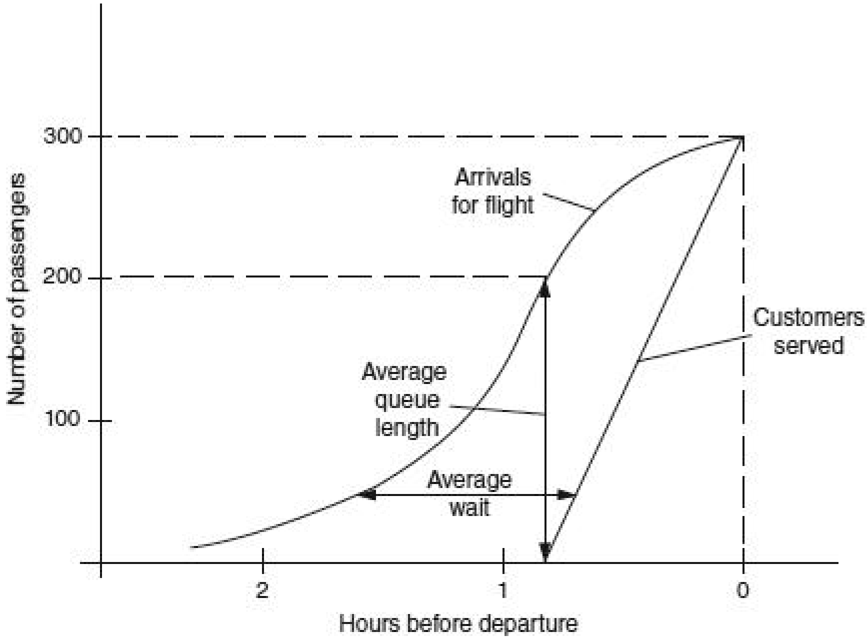


FIGURE 16.7 Combining cumulative arrival and departure diagrams to estimate length of queues and waiting time.

Designers can use this analysis to estimate if a potential hot spot is likely to cause a difficulty. They can then also use it to identify possible solutions to these difficulties by exploring alternative forms of operation that change the cumulative departure diagram—and thus the average length of queue and wait time. Example 16.3 shows how to do this.

Example 16.3 *The problem*: Consider the hot spot in front of the check-in counter described in Example 16.2. Allowing space for people to move along the corridor, less than 12 m (40 ft) was available for the queues. The cumulative arrival and departure diagrams in Fig. 16.7 describe the evolution of the traffic at the check-in process. They show that the airline operator could expect 200 passengers to be waiting for service when the counters open 50 minutes (T-50) before the flight closes. There would then be about 22 people in each queue, assuming that the process has the nine servers required to serve all customers, as calculated in Example 16.1. Because travelers in line with bags require about 0.6 m (2 ft) per person (see discussion for check-in counters in Sec. 16.5), the queues would be about 13.2 m (44 ft) long and exceed the space available. The result is a hot spot. *The solution*: Cumulative diagrams allow us to explore alternative solutions. An obvious possibility is to open the counters earlier, moving the cumulative departure line to the left. However, as the number of people served at any time cannot exceed the number who have arrived, this solution implies that the cumulative departures would rise until it met the arrivals curve, and then track it as long as the arrival rate was less than the service rate—then implying that some check-in agents were idle. A better solution is to open some check-in counters early, to reduce the maximum length of queue, and to open the remainder of he counters later. Figure 16.8 illustrates this possibility: Five counters opening 30 minutes earlier (T-80) would serve 90 persons by T-50, halving the queue at that time and ensuring that it would easily fit into the space available. Other solutions are possible and can be investigated with the cumulative diagrams.

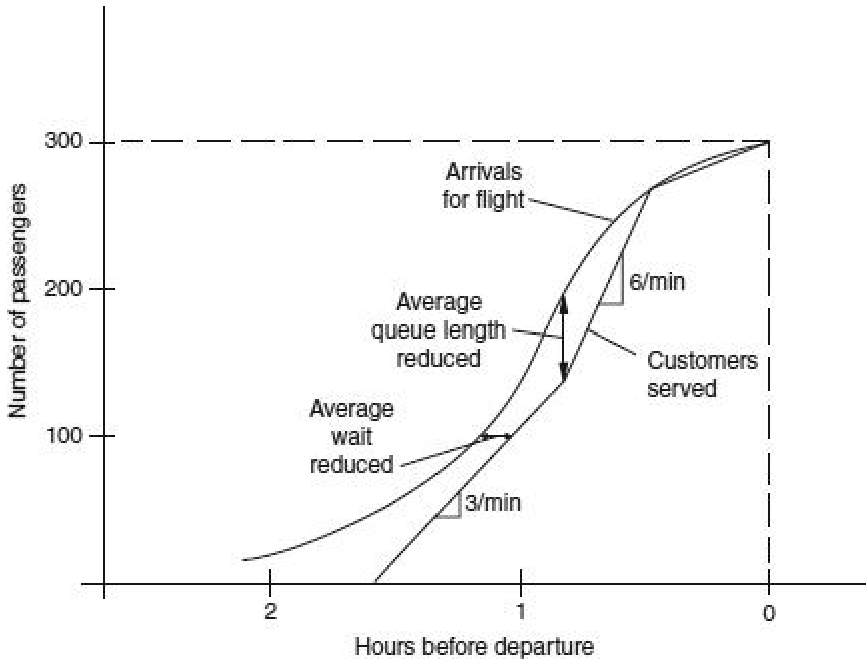


FIGURE 16.8 Opening some check-in counters earlier to reduce the wait time and maximum length of queue.

Analyses using cumulative arrival and departure diagrams provide a simple way to identify and resolve major difficulties successfully. They constitute a powerful tool that any professional designer should be able to employ. However, these analyses have limitations. They are the following:

- *Deterministic*: They represent typical situations and do not describe the variations that occur in practice.
- *Suitable for one process at a time*: They give no indication of the interaction between processes, as when delays in one process affect the patterns of arrivals at subsequent processes.

They are therefore not suitable for all situations. A simulation analysis is necessary to explore the flux of traffic over time throughout the passenger building.

16.4 Simulation of Passenger Buildings

Good simulation models can be powerful tools in the detailed design of passenger buildings. They can be particularly useful in "fine-tuning" a design by studying the operation of the building *in its entirety*. Some of the tasks that planners can accomplish efficiently using simulation include the following:

- Checking that the successive processes and services that passengers go through interact smoothly, so that there are no obvious bottlenecks in the building
- Estimating the total processing time for all the services that each class of arriving, departing, or connecting passengers utilizes
- Identifying the number of agents, machines, desks, etc., required at each part of the building
- Assessing the level of service at each part of the building both regarding space per occupant and waiting times

The design of the jetBlue terminal at New York/Kennedy provides a salient example of effective use of simulation for a passenger building (see Example 16.4). Mumayiz (1990) and Mumayiz and Schonfeld (1999) review landside simulation models and their characteristics.

Simulation can provide a most cost-effective way to improve design. It is inexpensive compared to the large capital and operation costs associated with the construction and operation of a major passenger building. However, good simulations are not cheap. Successful simulations require significant efforts to collect relevant data suitable to the specific airport under the range of significant conditions (such as morning and evening flows, seasonal patterns, and differences between full-service and low-cost airlines). They also depend on accurate descriptions of alternative plans that designers are considering. Effective, useful simulations entail significant staff costs.

Example 16.4 The interior layout of the jetBlue Terminal at New York/Kennedy is largely the product of one of the more intensive uses of simulation to explore the implications of different arrangements and sizing of the areas serving customers: check-in, security controls, waiting and restaurant areas, bag delivery, and so on.

Seeking the most efficient possible arrangements, the airline worked closely with its simulation consultants (Arup and Weston-Wong Associates) to appreciate when and where flows would concentrate, to see how delays in some areas would affect downstream processes and to resolve as best they could all potential issues they could anticipate and control within their building. Their simulations analyzed the potential of shared space, in particular for common waiting and eating areas (along the lines indicated in Sec. 15.3).

The result is a uniquely compact and efficient design that sets new standards for efficient layout of services and use of space.

Airport designers and other prospective users should choose their simulation software carefully. They need to make sure that the simulation correctly models the flows and services in detail, and that persons knowledgeable about airport operations have created the simulations and will be responsible for their use in design. This warning is important. Unfortunately, much of the simulation software available for sale is superficially attractive but operationally undesirable. This is because it is relatively easy to program generic simulation software, and some application developers focus on presenting attractive graphical displays to impress prospective buyers. What really counts, however, is not the wrapping but the content. Prospective users of airport simulations should assess their alternative choices carefully, making use of advice from airport experts.

Good simulations will include many sensitivity analyses. Specifically, they will consider a broad range of airline schedules. The performance of any passenger building depends crucially on the way airlines operate. Traffic peaks are higher when airlines cluster their departures around some specific time, such as the top of the hour, as they often do. Analysts need to recognize the great uncertainty in airline schedules, which airlines alter frequently. Good simulations will explore the effects of schedule variations and will look to identify designs that perform well, or have the flexibility to be easily adapted to perform well, over the range of possible future conditions.

16.5 Specific Facilities

This section offers guidelines sizing specific parts of the passenger building. These are suggestive starting points. Indeed, no statement about facility sizes could be definitive, given the variety of standards that exist at different airports, as Sec. 16.1 discusses. Persons needing more detailed guidance should consult the latest version of the International Air Transport Association (IATA) *Airport Development Reference Manual* (IATA, 2004) and the FAA Advisory Circular 150/5360-13 (FAA, 1988). Perhaps unfortunately, there is no rule-book on detailed design of passenger buildings.

Table 16.2 gives an example of design standards for a major international passenger building. Experienced designers will notice that the values and objectives implicit in these standards might not be acceptable elsewhere. For example,

Item	Area		Per	Remarks
	ft²	m²		
Check-in counters	4.64	0.42	Peak-hour passenger	Based on 3.5 min/ passenger
Check-in queues	9.68	0.87		25 ft (7.5 m) depth
Circulation	15.48	1.4		40 ft (12 m) after check-in
Arrivals hall	15	1.35		
Toilets	3.5	0.31		Based on comparable airports
Baggage claim	85	7.65		
Concessions	10,000	900	Million annual passengers	
Public circulation	10,000	900		
Gate lounges	14	1.25	Seats in aircraft using gate	90% load factor, 80% seating, and 20% standing

Source: Adapted from R. L. Brown Associates and HNTB Joint Venture.

TABLE 16.2 Example Planning Factors for a Major U.S. International Passenger Building

- The allowance for concessions is small compared to the area set aside for circulation and the arrivals hall, and does not reflect the kind of aggressive marketing now practiced by some airports.
- Circulation corridors in the ticketing area and at the gates are 40 to 50 ft (12–15 m) wide and might be considered excessively generous (see Sec. 15.4), especially when this space might be devoted to concessions.
- The amount of seating at the gate lounges is much greater than necessary and does not reflect any sharing of the facilities (see Sec. 15.3).

In practice, designers often have to establish standards suitable for their project. They may best adopt criteria associated with airports that are comparable to the ambitions of the airport operator or owner. Table 16.2 illustrates this approach; its standards reflect practice at then competitive airports in the United States.

Queues

All passenger-handling processes involve lines of people waiting. How much space should designers allow for this? As a rough rule of thumb, the length of a queue is about 0.6 m (2 ft) per person. This number assumes that families and friends stand side by side. The actual average separation depends on whether people are using trolleys and on whether many people are in line, which tends to compress the total length of the queue. As Chap. 15 indicates, the level of service for a queue depends on the available space. Figures 16.9 and 16.10 illustrate levels of service C and E.

FIGURE 16.9 Snake line at level of service C.(*Source*: Zale Anis and Volpe National Transportation Systems Center.)

FIGURE 16.10 Snake line at level of service E. (*Source*: Zale Anis and Volpe National Transportation Systems Center.)

Snake queues generally use space more efficiently. This arrangement involves a single queue for persons waiting for several check-in agents or other servers. It is called a *snake queue* because it follows a back-and-forth channel between stanchions. Because it uses all the available floor space, instead of having gaps between queues in front of individual servers, snake queues reduce the amount of space needed for queues in front of counters. (This is a reason the depth of queues in front of counters in Table 16.2 is relatively short compared to standards elsewhere.) Snake queues have the further advantage of being fairer, because it is impossible for any passengers to be held up for a great length of time by a single customer who has some difficulty. However, snake queues have not been universally accepted outside the United States.

An operational disadvantage of usual snake queues is that the person at the head of the queue may not immediately notice when an agent is free and will in any case take time to move to the agent. This phenomenon means that agents can be idle for some time, thus reducing productivity. This problem can be overcome by having the snake queue feed one or two person queues in front of each agent. This approach can substantially increase the productivity of the service and its agents. This solution is becoming standard for passport inspection facilities in the United States.

Check-in Areas

Design standards for check-in facilities are changing. Norms presented in available references are unlikely to apply in the future. Security concerns, the Internet, and smart mobile devices are contributing to important shifts in airline and airport procedures. The end result of this process is speculative, but it is transforming the notion of check-in and may make conventional check-in halls obsolete. Electronic ticketing reduces the need for large check-in halls for two reasons. It

- *Reduces processing time at the airport*, as passengers arrive with "boarding passes" either downloaded to their mobile device or preprinted at home or office. This reduces the number of check-in positions per thousand passengers.
- *Distributes many aspects of the process away from the airport*, which means there is less need for large check-in areas.

Many passengers now bypass the check-in halls. Arriving with boarding passes and traveling with carry-on bags, they proceed directly to their boarding area, where they confirm their presence with the gate agent. Others arrive with computer-readable bar-coded or electronic receipts that can halve the time to check-in, compared to traditional twentieth-century processes. The common result is that airports need less space for check-in. Vast check-in spaces built around 2000, as at Boston/Logan Terminal E or San Francisco/International, now appear relatively empty and oversized.

Security Checkpoints

The crucial issue for security screening is the average service rate of the facility (or its inverse, the average time per person). This determines the number of devices needed. This time varies considerably due to different local practices, the nature of the passengers, and the weather (which influences what passengers wear). Domestic and business travelers on short trips take less time; families with children and several carry-on bags take more. During periods of special concern, the average time per person will increase, which will increase waits and queues and may necessitate additional screening devices and staff.

The formulas for calculating loads on security screening, the number of required devices and queue lengths are all permutations of Eqs. (16.1) and (16.2). The expected load on the security checkpoints can be estimated as

$$L = P(1 - T)(1 + K)R \tag{16.3}$$

where P = peak-hour enplanements; T = percent of transfers who bypass the security checkpoint; K = a factor that accounts for other airport traffic (employees, etc.) as a proportion of P; and R = a factor between 1 and 1.5 to provide additional capacity to cope with fluctuations in loads over the peak hour, a higher number for greater variability (FAA, 2001). From this we can calculate

$$\text{Number of checkpoints} = \frac{L}{(\text{service rate})} = \frac{L}{(S \times F)} \tag{16.4}$$

$$\text{Number of screening machines for carry-ons} = L\left(\frac{B}{X} \times F\right) \tag{16.5}$$

where F = a utilization factor reflecting breaks by personnel; S = the checkpoint nominal service rate; X = nominal service rate of the screening machine; and B = number of carry-ons trays per passenger. Note that whereas service rates may be relatively stable, the number of carry-on trays depends considerably on weather and destinations—are travelers wearing coats or not?

Armed with estimates of the loads on the system and the service times per person, analysts should be able to calculate the number of security checkpoints to provide good level of service (see Chap. 20). However, the observed result is that delays at security checkpoints can easily be 15 minutes and delays of up to 30 minutes occur regularly. In the context of normal airport standard expectations for level of service (see Sec.15.4), these results are unacceptable. To some extent these conditions occur because of inadequate design. More generally, the delays occur for a range of operational reasons such as:

- Variability in inspection procedures between airports, which confuse and delay passengers
- Inexperienced staff due to high turnover in the boring job of inspecting
- Inadequate staffing at critical hours, such as early in the morning, when people arrive before the inspection stations open

The bottom line is that security checkpoints are and can be expected to continue to be hot spots of inadequate service.

Good design of security checkpoints features secondary checks behind the primary detectors. These make it possible to deal separately with persons who set off the alarm when they first walk through the detector. If such devices are not present, each person who has to go back and walk through the metal detector again will delay the entire process while fumbling with forgotten change, taking out a cell phone, or removing a belt.

Good design of the screening process will include about 4 m (13 ft) or more of table space either side of the screening devices for carry-ons. These allow passengers time on entry to fill their trays and become ready for inspections, and on exit to pick up and put on their belongings and clear space for following trays. The original design of security checkpoints at many airports failed to provide sufficient tables—and in some cases none! The result was that the screening process was idle while passengers prepared their carry-ons for screening, and this reduced the throughput by as much as half. See Fig. 16.11.

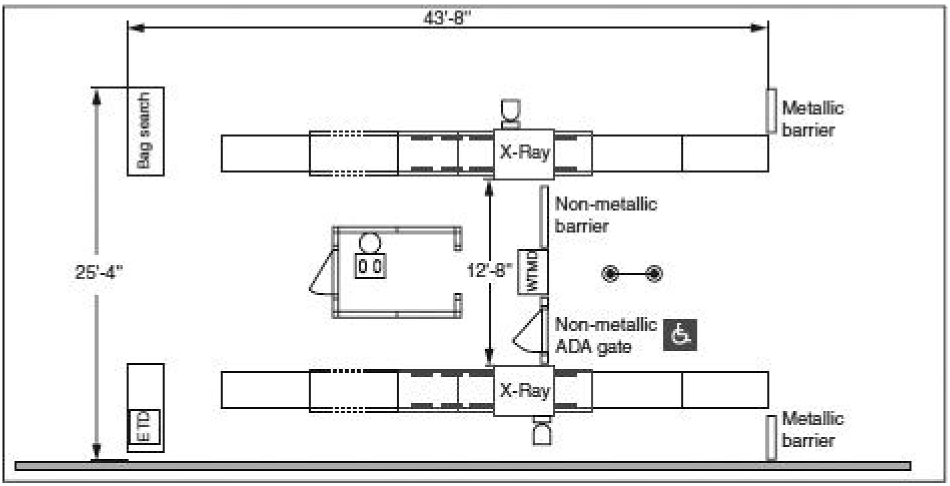


FIGURE 16.11 Two to one configuration (two streams of bag checks with one stream of body checks) of security checkpoint with holding/wanding station. (*Source*: Transportation Security Administration, 2006.)

Passport Control Processes

As with check-in, these processes are changing in the electronic era. Bar codes and chips in passports now reduce the time to enter data on passengers. For visitors to the United States, airlines forward passenger lists electronically to immigration authorities so that they can process clearances before passengers arrive. Although border authorities worldwide move cautiously, their processes are becoming more efficient—and require relatively fewer positions. Political integrations, such as the Schengen agreement among many European countries, can also have a tremendous impact in this regard insofar as they minimize passport controls for passengers moving between signatory nations.

Moving Walkways (or Sidewalks)

Many airports use moving walkways to reduce the distance people have to walk and to facilitate their movements. These devices extend up to about 50 m (165 ft) each. ACRP (2012) provides an extensive empirical database on the performance of these conveyances (and also on elevators). Airports typically deploy them in chains to cover long distances, as at Denver/International and Hong Kong/Chek Lap Kok. Transfer airports usually install them in pairs going in opposite directions.

Moving walkways do not, on average, speed up pedestrian traffic. Some people will move faster using these devices, to the extent that they can walk freely on them. However, many people will stop and rest on the moving sidewalks, and block others from moving on them. The speed of moving sidewalks is about 40 m/min or 2.4 km/h (1.5 mph). This is half the average pedestrian speed in airports, based on empirical observations of pedestrian movements in various airport terminal corridors. On balance, moving sidewalks do not seem to speed up pedestrian flows (Young, 1995 and 1999). They act as a convenience when distances are long.

Waiting Lounges

The industry is not agreed on a standard for space devoted to waiting lounges at the gate. What makes sense depends on how the airport and its airlines will operate the facility. In any case, designers should plan on accommodating a number of passengers less than aircraft capacity. They do not have to provide all the waiting space for passengers at individual gates. Passengers can and will wait in space shared with other gates, in concession areas, and airline clubrooms. In any case, they normally filter into the boarding area as others begin to board. In Singapore, for example, the gate rooms are quite small; the airport tends to call flights only when they are ready to board, and thus accommodates very few persons in these lounges. Similarly, the design of London/Heathrow Terminal 5 counts on most passengers spending their time before flights in shops and restaurants, and provides minimal gate lounges. Whereas in some operational circumstances it might be necessary to provide individual gate lounges for a large fraction of aircraft capacity, current practice generally makes this unnecessary.

Finally, the amount to be provided depends greatly on the extent to which gates share common waiting lounges. As Sec. 15.3 indicates, it is possible to achieve up to 50 percent reductions in the total space allocated to waiting areas if these are arranged so that they can be shared (see also de Neufville et al., 2002).

Economically efficient design will plan on shared waiting lounges and expect that many passengers will spend time outside these areas. Unfortunately, the gate lounges in some airports are sterile—closed areas without easy access to amenities, as is the case in Bangkok/Suvarnabhumi and San Francisco/International in the international building. Such design is neither attractive for passengers nor profitable for the airport operator, who would prefer to have passengers spend time in concessions. Good design of gate lounges features shared facilities with good access to shops and services.

Concession Space

The financial success of concessions in a passenger building depends on the amount of time people spend in the airport. The longer they have to wait around, the more likely that they will chose to shop. Put another way, commercial success depends to some extent on the inefficiency of the airport from the perspective of passengers who want to get to their plane or destination as quickly as possible.

Moreover, the longer people stay in the airport, the more space the designers will have to provide. The airport operator therefore faces a dilemma: Commercial success comes at a significant cost in terms of space and possible operational inefficiency. Decisions about how much commercial space to provide should reflect management objectives. They cannot be derived from strictly technical considerations.

Shopping at the airport is a definite convenience for many passengers. They may wish to use duty-free stores or take advantage of finding many brands and souvenir shops in one place. Local residents may also use the airport for routine shopping, as happens in countries where transportation centers are exempt from restrictive laws limiting when stores can operate. This is why Frankfurt/International and Amsterdam/Schiphol feature grocery and clothing stores. Those situations are unusual.

Overall, commercial activities at airports have become much more significant worldwide as knowledgeable retail operators have taken over the management of this space. In a nutshell, retail experts have been changing the previously prevailing paradigm. Particularly in the United States, public airport operators used to award concession spaces to the highest bidders. Smart operators of shopping malls recognize, however, that they can promote shopping and overall revenues by creating areas of interest that will attract customers—and they do so by subsidizing or giving away space to selected tenants and activities.

Planning commercial space is risky. Many—but not all—shopping developments at airports have been successful. Some companies have obtained their highest sales volumes per unit area at airports, and a number of airports have achieved remarkable success. However, some commercial developments in new airport passenger buildings have been significant failures. Creating the best concession spaces takes skill and experience, not just space. Airport operators need to plan their commercial space carefully with the help of retail experts.

As the saying goes, the three most important factors in successful sales are “location, location, and location.” This is a key factor influencing the success of any commercial space. As a rule, the best locations for commercial space are in the direct line of passenger traffic, after the various barriers between them and the gates. Commercial space needs to be seen. Stores placed in out-of-the-way places, on mezzanines people reach by changing floors, or where there is little traffic, are at a disadvantage. For example

- Many stores in the linear buildings at Kansas City and Munich withered for lack of traffic.
- Passengers like to move directly from check-in to their departure gates without making a detour to an upper floor, and thus avoid mezzanine commercial spaces, as at Tokyo/Narita Terminal 2.
- The expensive restaurant and various stores hidden from view behind the check-in counters in the International Terminal in San Francisco/International closed after only a few months of operation.

In short, the development of profitable commercial space requires expert attention.

Commercial space benefits from being in places where passengers feel comfortable browsing. This means that it works best “after security,” after passengers have passed through all the barriers that impede their way to the gate. For example, in the new passenger building at Washington/Reagan, people walk by the stores “before security,” wanting to be sure to make it to their gate and catch their flight. Once passengers have passed through security and are at their gates, however, they relax, shop, eat, and drink.

In countries with especially high value-added or sales taxes, it can be advantageous to place duty-free stores in the arrivals areas of airports. This enables travelers to avoid the hassle of carrying goods, especially liquids such as alcohol and perfume, as carry-on baggage.

Operators of airport stores estimate the area needed per store in terms of the number of persons passing by, the percentage of passengers they can attract (the “penetration rate”), and the time passengers remain (“dwell”) in the shop. They also add extra space to account for inventory space. A simple version of this calculus is

$$\begin{aligned}\text{Passengers in store} = {} & [(\text{passenger flow rate (passengers/hour)}] \\ & \times [\text{target penetration}] \\ & \times [\text{average dwell time in shop in hours}] \\ & \times [\text{peaking factor}]\end{aligned} \tag{16.6}$$

The target penetration might be 0.65 for a duty-free store, and much lower at gate lounge shops, for example, 0.25. The peaking factor might be 1.2. The calculation of passengers in the store leads to an estimate of the appropriate its size. For example,

$$\text{Space in m}^2 = 4\ [\text{number of passengers in store}] \tag{16.7}$$

The factor of 4 allows 1 m^2 for personal space, 2 m^2 for circulation, and 1 m^2 for cashiers, checkout space and stock displays (Freathy and O'Connell, 1998).

Baggage Claim Areas

Baggage claim areas must first provide enough *claim presentation length*, that is, length along the conveyor belt or racetracks, for people to identify and pick up their bags. The IATA (2004) standards recommend about 70 m (230 ft) for wide-body and about 40 m (130 ft) for narrow-body aircraft being served at the same time. This standard implies about 0.3 m (1 ft) of claim presentation per passenger. In practice, the amount available appears to deviate ±50 percent from this recommendation. The FAA (1988) alternatively defines the length required in terms of the number of aircraft arriving in a peak 20 minutes and assumes that passengers check 1.3 bags per person. Either standard leads to approximately the same results. However, these standards should be modified according to local realities such as the average number of bags checked. Responding to airline fees, passengers in the United States have been checking fewer bags.

The standards most particularly need to account for the intensity of the transfer traffic at the airport. The greater the percent of travelers who connect between flights, the less space is needed in the baggage claim area. The FAA standards include this factor. A major exception to this rule concerns international arrivals in the United States. According to U.S. practice, all passengers entering the United States must claim all their baggage upon arrival and present it for inspection even if they are transferring to another city (they then have to recheck it at a special facility beyond customs). In this case, baggage claim devices need to serve all passengers.

Congestion hot spots can be a major difficulty in baggage areas. This is because passengers naturally cluster at the places where bags appear on the claim devices. These locations can become crowded even if the entire baggage area meets other space standards (see Sec. 15.3). As a precaution against such problems, IATA (2004) recommends that the distance between adjacent baggage claim devices be at least 9 m (30 ft).

Designers should also cater to the need to handle "outsize" accompanying items. These typically consist of sporting equipment such as skis, bicycles, or surfboards. Airports serving resort destinations obviously see the most of this traffic. However, these outsize items come and go through other airports, and need appropriate check-in and reclaim areas.

Baggage Makeup Space

Airports routinely discover that they do not have enough space, of the right kind and in the right place, to handle bags efficiently and effectively. Their problems have been especially acute since the turn of the century, as security rules have evolved and required new machines and procedures. Additionally, changes in the structure of the airline industry impose new functional requirements that existing systems often cannot accommodate easily. Larger aircraft and faster turnaround times demand more makeup and storage space to serve each gate. Airline development of new transfer operations at airports impose the need for 'tail-to-tail' distribution of bags between flights that designers had not envisaged for existing systems. In short, airports worldwide regularly scramble to improve their baggage systems.

The crux of the problem is that space for baggage systems is simultaneously inflexible, and must accommodate inherently inflexible systems. The space is inflexible because it is almost always tightly confined. Other buildings or immovable roadways typically surround the space on all sides, while it is sandwiched between the ground and passenger levels. In addition, beams and columns supporting the floors above cut up the space. In extreme cases, the bag rooms are actually underground in basements. Baggage systems are also inflexible. They consist of extensive tracks that restrict possible configurations: Inclines must be gentle, to prevent bags from tumbling out of the slots that identify their destinations; curves must be gradual to accommodate the trays or belts that compose the system. Moreover, tracks will cross each other and require sufficient vertical clearance for the bags, for maintenance staff, and for vehicles delivering containers to and from the bag rooms and the aircraft. (Section 17.9 discusses these issues in detail.) Designers often must shoehorn bag systems into confined spaces.

It is not possible to provide reliable guidelines as to the amount of space designers need to provide for bag systems, as their technology is evolving rapidly, and depends on the intensity and type of operations. The immediate advice is that the architects and designers of the overall configuration of any passenger terminal should consult with baggage experts from the start. Unfortunately, this rarely happens to the extent it should.

The most important advice as regards the overall design of space for baggage systems is that it should be flexible. It should have the capacity to expand easily to accommodate new equipment, more storage space, and additional tracks as necessary. Without such flexibility, the airport runs the risk that its bag system will become a hot spot of congestion and delays that will determine the capacity and usefulness of the passenger buildings.

Curbside and Equivalent Areas

Passenger buildings need extensive space for people to get out of and into cars, taxis, and buses. This space typically consists of one or more long stretches of sidewalk along a roadway in front or alongside of the passenger building. Passengers and visitors will either alight from their vehicles on this sidewalk and proceed into the building, or emerge from the building onto this sidewalk and get into their vehicle. Airport operators and designers refer to this space as the *curbside area*.

Many airports provide the equivalent of curbside space through short-term parking areas located in front of the passenger buildings. These areas permit drivers to stop their cars to drop off or pick up passengers, and are functionally equivalent to the curbside area. These short-term parking areas are beneficial both to the airport operator and the traveling public (de Neufville, 1982). They allow airports to expand their curbside area relatively easily, in situations where it might be impossible to extend the actual curb in front of the passenger building. Moreover, short-term parking generates revenue for the airport, which curbsides do not, generally at much higher rates than longer-term parking spaces. Short-term parking areas also allow drivers to stay in these drop-off and pickup areas longer than they could at the curb in front of the building. Boston/Logan and Montreal/Trudeau airports have such arrangements (see Sec. 17.4).

"Cell phone lots" have become a popular, effective device for relieving curbside congestion. These are open areas in which drivers can wait for a short time until their friend or client calls to let them know that they have gotten off the plane, retrieved their bags, and are ready to be picked up. These facilities provide workable alternatives for many friends and relatives who might otherwise idle at the curb or cruise around the airport until their pickup was ready (from 15 to 50%, as reported in ACRP, 2010). Because drivers now tend to get to the airport close to the arrival time of their air passengers (since they can usually track the actual arrival times of aircraft), their dwell times in the cell phone lots are typically only 20 minutes or so. As this dwell time is low, the hourly capacity of cell phone lots is high, perhaps three times the actual number of spaces. Existing cell phone lots at major airports have been small, with space for perhaps 50 cars.

Arriving passengers need considerably more curbside space than departing passengers. This fact needs emphasis. Inexperienced designers might assume wrongly that because the numbers of arriving and departing passengers are about equal, so should be their need for curbside space. This is incorrect, because people have different needs on arrival than on departure. In the departure process, drivers merely need to drop off the passengers—this can happen quickly. On arrival, however, drivers first need to locate the arriving passengers—this takes time. The dwell time for cars picking up passengers is therefore longer, and this translates into a need for more space. The exact amount of extra space needed for arriving passengers depends on local circumstances. It can be over twice as much as for departing passengers as Table 16.3 indicates (de Neufville, 1982; Mandle et al., 1982).

Note that while these data are old, they accord well with 2009 United States reports (ACRP, 2010).

City	Dwell Time in Minutes	
	Departing	Arriving
Denver	1.2–2.8	4.8–6.9
Miami	1.6–4.5	2.3–4.5
New York	1.0–1.6	2.1–4.8

Source: Mandel et al., 1980.

TABLE 16.3 Example Differences in Curbside Dwell Times for Arriving and Deplaning Passengers

A formula for estimating the desirable length of "curbside" area appears below (adapted from Mandle et al., 1982). It is the sum of the length provided for private cars and taxis, and for larger vehicles such as buses and courtesy vans that take longer to load and unload. Thus

$$\text{Length of curb space} = C = C \text{ for cars} + C \text{ for busses}$$

$$C_I = P[(M_I)(F_I)(D_I/60)(L_I)]/V_I$$

where $P =$ peak-hour number of originating and terminating passengers (no transfers)

$C_I =$ length of curb for passengers using mode I

$M_I =$ fraction of passengers using mode I

$F_I =$ fraction of passengers in mode I using the curb (as opposed to long-term parking)

$(D_I/60) =$ the average dwell time in hours for vehicles in mode I

$V_I =$ the average number of passengers in vehicles in mode I

$L_I =$ average length of space needed to park vehicles, about 7.5 m (25 ft) for cars and taxis, about 13.5 m (45 ft) for buses.

To put the curbside requirements into perspective, it is useful to express them in terms of length per passenger, *C*/*P*. As a very rough approximation, airports need about 0.15 m (0.5 ft) per 1000 annual passengers excluding transfers. This works out to about 1.5 km or almost 1 mile for an airport serving 10 million originating and terminating passengers. Such calculations demonstrate the potential importance of curbside requirements. However, airport operators can reduce the amount of curbside space considerably by careful design and management.

Airport operators can increase the productivity of their curbside primarily by reducing the dwell time. The length needed is directly proportional to the dwell time. Halving the dwell time halves the length needed or, alternatively, doubles the capacity of existing space. Airport managers can reduce dwell time by enforcing regulations on the amount of time cars can take to load or unload passengers. “Hiring policemen” is, in many cases, the simple answer to lack of curbside capacity. Los Angeles/International has been notoriously strict and successful in this regard.

Airport operators can also influence the number of people using the actual curbside along the passenger building by providing short-term parking convenient for the pickup and drop-off of passengers. This solution is inappropriate for taxis, but can reduce the amount of curbside needed by about a third, depending on the circumstances.

Major airports serving 10 million and more passengers annually must provide extensive curbside areas, even with clever design and extensive use of public transport that reduce requirements. They do this by

- *Double-decking the access roads*: placing the departing passengers on the upper level and arriving passengers on the lower level. This design doubles the amount of curbside available for a given frontage of the building.
- *Providing parallel curbsides*, consisting of long pedestrian islands that duplicate the actual curbside along the passenger building. Two separate parallel islands may exist at the busiest or most congested airports, such as Boston/Logan, London/Heathrow, and Atlanta (Fig. 16.12).

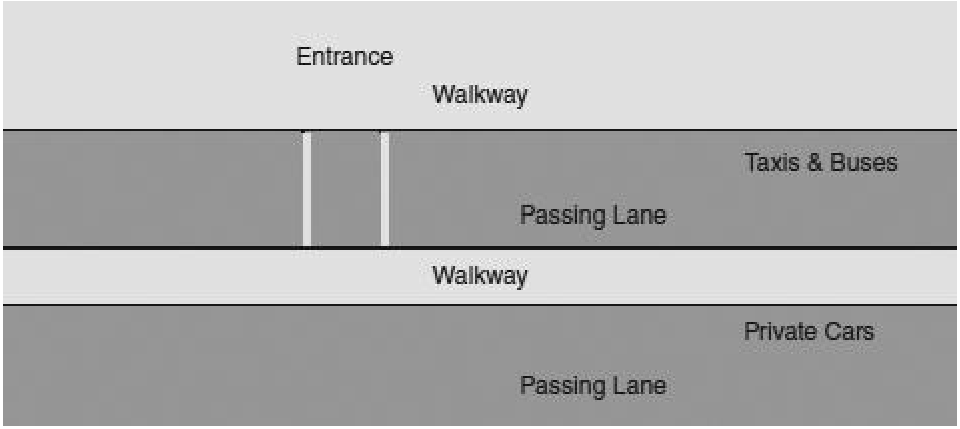


FIGURE 16.12 Typical layouts of curbs with multiple lanes.

- *Establishing taxi pools*, which hold waiting taxis until they are needed and called, which reduces their dwell time at the curb and the amount of curbside needed.

In designing parallel curbsides, good practice separates the buses and vans from taxis and cars. This is because larger vehicles need more time to load and unload. If they are mixed in with the smaller vehicles, they could block their flow. Similarly, when the curbside is particularly long, as in front of a linear passenger building, it is important to provide places where traffic can bypass vehicles double-parked somewhere along the curb. Airports can do this conveniently by using parallel curbsides with crossover points. Simulation analyses can usefully explore the exact design of these situations.

Taxi Operations

There is no industry consensus about how to design taxi pickup operations. Different airports variously line up taxis in single or double file, parking them parallel to the curb or diagonally. Some airports tightly control the flow of passengers; others expedite the process. There is little data or analysis on these operations. What is clear, however, is that taxi operations at many airports are often slow, involve long lines of customers exposed to the weather, and generally provide poor service.

At the curbside, taxi operations are open to improvement through the application of standard strategies for handling queues. These include

- Pre-positioning the next customers at taxi pickup points instead of at the head of a snake queue
- Establishing separate queues for different services (as by destinations)

- Organizing the taxi queues for greatest efficiency

As Costa and de Neufville (2012) have shown, careful simulation, paying close attention to operational details observed on site, can be used effectively to analyze alternative solutions.

A generic problem for major airports concerns the delays associated with taxi pools. During busy hours, hundreds of taxis may be waiting for fares at an airport. As there is no space for this fleet right at the curbside, airports often provide large remote parking areas for the taxis. This is the taxi pool. In operation, dispatchers receive requests for taxis, and direct taxis from the pool to the locations where they are needed. The delays occur for two reasons: The taxi pool is far away and the dispatch process is inefficient. Also, dispatchers may neither be able to react immediately nor send enough taxis to where they are needed. These issues severely degrade the level of service at a number of airports. The net effect is that taxis take excessive time to move from the pool to curbside pickup points, especially at congested times. Airport planners should anticipate this problem and deal with it.

Consolidated Rental Car Facilities

Major airports now collocate rental car agencies in a multistory parking facility, known generically as a consolidated rental car facility (Conrac). They then may also provide some kind of transportation between the curbside and this facility, which may be by bus or people mover, as at Atlanta and San Francisco/International.

The use of consolidated car rental facilities has two advantages. It

- Reduces curbside congestion and pollution. It eliminates the competing fleets of vans circulating past the curb trying to pick up customers for their companies.
- Saves valuable airport space for alternative uses, insofar as it replaces several ground level parking areas with a more compact, multistory garage.

This approach is often an effective way to improve the level of service to passengers.

Conracs can be large and expensive. According to reports, the Atlanta Conrac has over 8000 spaces and cost $480 million; the San Francisco/San Jose Conrac has about 3500 spaces and cost $260 million. The corresponding average costs per space were about $60,000 at Atlanta—and $74,000 at San Jose, which was deliberately built as an iconic work of architecture. These costs are greater than the average total cost of ordinary structured airport parking, which can exceed $40,000 per space (see Sec.17.4). This is because they require considerable space for the rental car operations (offices, customer lobbies, car washes and minor repair shops), and to accommodate the buses or people movers that ferry customers between the Conrac and the passenger buildings.

16.6 Take-aways

This chapter embodies three overarching ideas:

1. Available design standards for the detailed design of passenger buildings do not apply universally. They are indicative, and airport managers should adapt them to locally defined priorities and conditions.
2. Good design will identify and avoid potential "hot spots" that reduce capacity and level of service. It will do this by carefully considering the detailed flow of passengers and baggage through the building, paying close attention passenger behavior at critical times.
3. Simulation is a fundamental tool for analyzing and validating the detailed design and operation of passenger buildings. But not any simulation will do. Airports need to choose validated models that embody detailed appreciation of airport processes.

Exercises

16.1. Rework Example 16.1 assuming the airline could reduce the average check-in time by 20s. How would that affect the number of counters and personnel needed?

16.2. Rework Example 16.3 assuming that there are only six counters and agents take an average of 1 minute per passenger to check in travelers.

16.3. Visit an airport and observe the check-in process for one or more flights. (If you choose several, look at different kinds of flights, such as domestic, international, shuttle, or "cheap fare.") How long is the average processing time? How long are the queues? How much space does each person in queue take? What suggestions would you make about how this process could be improved?

16.4. If possible, observe the security procedures at some airport. On average, how far apart are the persons in the queue who have a minimal amount of baggage? What is the average time per person for this process? Are there secondary facilities for screening persons who cause an alarm on the primary metal detector? In your estimation, how well did this process work during your observations?

16.5. Observe how people use waiting lounges at gates in one or more airports. Considering the size of the aircraft for the flight, what fraction of the passengers is in the lounge at any time? Where else are they? What is the ratio of seats in the lounge to the size of the aircraft? Are there enough seats? What kind of changes would you suggest?

16.6. Choosing specific flights, note the time they actually arrive at the gate. Then observe the time their first and last bags appear in the baggage claim area. Also observe, if possible, when the first passengers from the flight arrive. Use these data to evaluate the quality of the baggage delivery service for that flight.

16.7. Visit an airport and observe its curbside operations in a peak period. How are they configured? How much curbside is there? How do these measurements correspond to the guidelines? Do these operations appear adequate?

References

Airport Cooperative Research Program (2010) *Airport Passenger Terminal Planning and Design*, ACRP Report 25, Transportation Research Board, Washington, DC. www.trb.org.

Airport Cooperative Research Program (2010) *Airport Curbside and Terminal Area Roadway Operations*, ACRP Report 40, Transportation Research Board, Washington, DC. www.trb.org.

Airport Cooperative Research Program (2012) *Airport Passenger Conveyance Systems Planning Guidebook*, ACRP Report 67, Transportation Research Board, Washington, DC. www.trb.org.

Ashford, N. (1988) "Level of Service Design Concept for Airport Passenger Terminals—A European View," *Transportation Planning and Technology, 12*(1), pp. 5–21.

Costa, D. and de Neufville, R. (2012) "Designing Efficient Taxi Pick-up Operations at Airports," *Transportation Research Record* 2300, pp. 91–99 Transportation Research Board, National Research Council, Washington, DC.

de Neufville, R. (1982) "Airport Passenger Parking Design," *ASCE Journal of Urban Transportation, 108*, May, pp. 302–306.

de Neufville, R., de Barros, A., and Belin, S. (2002) "Optimal Configuration of Airport Passenger Buildings for Travelers," *ASCE Transportation Engineering Journal, 128* (3), pp. 211–217.

de Neufville, R. and Grillot, M. (1982) "Design of Pedestrian Space in Airport Terminals," *ASCE Transportation Engineering Journal, 108* (TE1), pp. 87–102.

FAA, Federal Aviation Administration (1988) *Planning and Design Guidelines for Airport Terminal Facilities*, Advisory Circular 150/5360-13, U.S. Government Printing Office, Washington, DC.

Freathy, P. and O'Connell, F. (1998) *European Airport Retailing: Growth Strategies for the New Millennium*, Macmillan Business, London, UK.

IATA, International Air Transport Association (2004) *Airport Development Reference Manual*, 9th ed., IATA, Montreal, Canada.

Mandle, P., Whitlock, E., and LaMagna, F. (1982) "Airport Curbside Planning and Design," *Transportation Research Record 840*, pp. 1–6, Transportation Research Board, National Research Council, Washington, DC.

Mumayiz, S. (1990) "Overview of Airport Terminal Simulation Models," *Transportation Research Record 1273*, pp. 11–20, Transportation Research Board, National Research Council, Washington, DC.

Mumayiz, S. and Schonfeld, P., (eds.) (1999) *Airport Modeling and Simulation*, ASCE, Reston, VA.

Young, S. (1995) "Analysis of Moving Walkway Use in Airport Terminal Corridors," *Transportation Research Record 1506, Airport and Air Transportation Issues*, pp. 44–51, National Academy Press, Washington, DC.

Young, S. (1999) "Evaluation of Pedestrian Walking Speeds in Airport Terminals," *Transportation Research Record 1674, Pedestrian and Bicycle Research*, pp. 20–26, National Academy Press, Washington, DC.

---

[1] The service rate is the inverse of the average time to complete service. If a server takes 2 minutes on average to complete a service, its service rate is 1/2.0 = 0.5/min.

CHAPTER 17

Ground Access and Distribution

Good ground access is vital for an airport. It is an indispensable ingredient of good service and maintains the attractiveness of the airport for the users and the value for the airport operators. The issue is that providing adequate and appropriate means of getting to and from the airport is a major challenge for airport planners and operators.

Because most users and employees of the airport are widely dispersed over the metropolitan area, it is inevitable that cars and buses play a significant role in providing airport access. For particularly busy or remote airports associated with cities that have a substantial rail transit system, tracked systems increasingly constitute an important means of travel to and from the airport. Hong Kong/Chek Lap Kok, Seoul/Incheon, and Tokyo/Narita exemplify this situation. In general, however, airports need to provide substantial highway access and parking facilities.

Automated people-mover systems (APMs) can distribute users efficiently around airports. Planners are thus installing them increasingly at major airports. The use of these systems has been one of the major innovations in airport design of the last generation. On the landside, APMs can distribute passengers and employees efficiently to various airport facilities dispersed widely over the airport. On the airside, APMs allow designers to disperse gate positions widely over the airfield, thus facilitating efficient aircraft ground maneuvers, without making passengers walk excessively.

The distribution of baggage around the airport is a major issue especially as baggage systems are large, cumbersome, and located in basements and other areas that are difficult to expand. Effective sorting systems are crucial to the performance of major airports. These require operational and design procedures that minimize sorting, cost-effective and reliable mechanical devices, extensive information systems to track the bags, and fallback systems that deal with the eventual sorting failures.

In this regard, designers have to size mechanical distribution systems that serve many points, baggage systems in particular but also people movers, with about twice the nominal capacity as they expect to use. They have to do this to ensure that their systems deliver reasonable levels of service in terms of reliability and delays. This extra margin is needed to guarantee that there are enough empty spaces in the system to serve all the prospective users in good time.

17.1 Introduction

This chapter covers the range of networks for connecting passengers and their baggage to the airport and for distributing them around the airport passenger buildings. These range from metropolitan systems dealing with airport access to mechanical devices operating within the limits of the airport. The discussion proceeds from the large-scale, regional issues of access to and from the airport, to the smaller-scale issues of people movers and baggage handling systems at the airport.

The next two sections first define the overall patterns of flows between the city and the airport and then indicate which access systems best serve these needs. As much of the traffic to and from most airports almost inevitably moves by automobile, Sec. 17.4 discusses strategies for providing adequate parking at appropriate prices. Sections 17.5 to 17.8 then describe the several ways of moving passengers around airports, focusing on the APMs that have been the central innovation in the design and operation of airports in the past generation. Section 17.9 closes with a presentation of the crucial issues of baggage handling systems, vital elements of airports.

Airport connection and distribution systems are highly complex. The flows on these networks are typically "many-to-many." Passengers and bags originate from many different points and go to many distinct destinations. For example, people come from homes, hotels, or offices spread around the city, and go to one of several passenger buildings at the airport. Passengers and bags similarly flow from dozens of check-in counters to as many or more aircraft. In simpler cases, the flows may be considered to be "many-to-one," as when passengers coming to the airport go to only a single passenger building. For practical purposes, however, the traffic flows around the airport have multiple origins, multiple destinations, or both. Only rarely are they "one-to-one," serving a single origin and destination. These few cases are predominately APMs that shuttle passengers between a main passenger building and a single satellite or midfield passenger building, as at London/Gatwick and Tokyo/Narita. The interactions among the many intersecting flows increase the complexity of the connections between origins and destinations. Many traffic streams either merge with each other or, worse, merge and then must be sorted according to their several destinations.

Because airport distribution systems are complex, their effective capacity is much less than the maximum flow the system could carry. A complex distribution system involves many sequential queues, as passengers and bags wait for vehicles to take them somewhere, and as these vehicles wait their turn to merge into other flows of traffic. Such queuing systems can only consistently provide stable service with reasonable delays when they operate at a fraction of their maximum capacity (see Sec. 17.8 and Chap. 20). Moreover, a second factor worsens this generic problem. To provide adequate service to traffic from each of the origins on the network, the system must deliberately feature numerous windows of empty capacity. These empty spaces ensure adequate service to traffic originating down the line. The net effect of the queues and need for empty capacity is that the practical capacity of complex distribution systems is perhaps only half their mechanical ability to process flows. Failure to appreciate this point has led to numerous expensive and embarrassing mistakes.

17.2 Regional Airport Access

Getting to the airport can be a challenge. The airport is generally relatively far away, roads may be congested, and traffic patterns at the airport are often confusing. For most travelers, especially the many going to the airport at rush hours, the trip to the airport can be a most annoying part of their journey. Equally, the daily trip to and from the airport can be difficult and expensive for the airport and airline employees. Together, the difficulties of getting to the airport conveniently and reliably constitute the *airport access problem*. What should airport operators and regional authorities do about it?

A popular idea is to build high-speed rail connections between the airport and the center of the city. The concept of modern vehicles speeding travelers between the airport and their local destinations resonates with civic leaders and their constituents, as well as with the local construction industry. This vision has motivated long-term efforts to build such systems worldwide. In this vein, the BAA built the Heathrow Express connecting London/Heathrow with Padding-ton Station in London. Malaysia similarly built a rail link between the capital and Kuala Lumpur/International. Shanghai inaugurated a magnetically levitated, high-speed train from Shanghai/Pudong toward the city center. In San Francisco, authorities extended the Bay Area Rapid Transit (BART) to San Francisco/International and started a $500 million project to connect it with San Francisco/Oakland. The Port Authority of New York and New Jersey developed a rail connection to New York/Kennedy. Authorities have built or extended rail systems to 75 major airports (Tables 17.1 to 17.3). To what extent is this popular notion a useful and effective approach to the airport access problem?

City	Airport	Intercity	Metropolitan
Atlanta			yes
Baltimore			yes
Boston	Logan		yes
Chicago	O'Hare		yes
	Midway		yes
Cleveland			yes
New York	Kennedy		yes
	Newark Liberty	yes	
Minneapolis/St. Paul			yes
Philadelphia			yes
Phoenix			yes
Portland (Oregon)			yes
San Francisco	International		yes
	Oakland		yes
St. Louis			yes
Toronto CAN	Pearson		planned
Vancouver CAN			yes
Washington	Baltimore		yes
	Dulles		u.c.
	Reagan		yes

TABLE 17.1 Airports Served by Rail Systems in the United States and Canada

Country	City	Airport	Intercity	Metropolitan
Australia	Sydney			yes
China	Beijing			yes
	Shanghai	Pudong		yes
	Hong Kong	Chek Lap Kok		yes
Dubai		Al Maktoum		u.c.
		International		yes
India	Delhi			yes
Israel	Tel Aviv			yes
Japan	Nagoya	Chubu		yes
	Osaka	Shin Kansai		yes
	Sapporo	Shin Chitose	yes	
	Tokyo	Haneda		yes
		Narita	yes	yes
Korea	Seoul	Gimpo		yes
		Incheon		yes
Malaysia	Kuala Lumpur	International	yes	
Philippines	Manila	Ninoy Aquino		yes
Singapore	Singapore	Changi		yes
Thailand	Bangkok	Don Muang	yes	
		Suvarnabhumi		yes

TABLE 17.2 Airports Served by Rail Systems in Asia and Australia

Country	City	Airport	High Speed	Intercity	Metropolitan
Austria	Vienna				yes
Belgium	Brussels	International			yes
Denmark	Copenhagen			yes	
France	Lyon		yes	yes	
	Paris	de Gaulle	yes	yes	yes
		Orly			yes
Germany	Berlin	Schönefield	planned		yes
	Dresden				yes
	Düsseldorf		yes	yes	yes
	Frankfurt	International	yes	yes	yes
	Hamburg				yes
	Hannover				yes
	Köln-Bonn		yes		yes
	Leipzig-Halle		yes	yes	yes
	Munich				yes
	Stuttgart				yes
Greece	Athens				yes
Italy	Milan	Malpensa		yes	
	Rome	Fuimicino			yes
Netherlands	Amsterdam			yes	yes
Norway	Oslo	Gardermoen		yes	
Poland	Warsaw	Chopin			yes
Portugal	Lisbon				yes
	Porto				yes
Russia	Moscow	Domodedovo			yes
		Sheremetyevo			yes
Spain	Barcelona	International			yes
	Madrid				yes
Sweden	Stockholm	Arlanda			yes
Switzerland	Geneva			yes	yes
	Zürich			yes	yes

United Kingdom	Birmingham			yes	
	Glasgow				scrapped
	London	Gatwick		yes	
		Heathrow		yes	yes
		Stansted			yes
	Manchester				yes
	Newcastle				yes

TABLE 17.3 Airports Served by Rail Systems in Europe

To answer this question and to deal appropriately with the airport access problem, it is necessary to understand the nature of the traffic to the airport, the costs and performance of the alternative systems, and the preferences of the users. This section describes the nature of the traffic flows, their distribution, and the preferences of the users. Section 17.3 details the performance of the technical possibilities, how they match the criteria of actual and potential investors, and recommends preferred choices for different circumstances.

Nature of Airport Access Traffic

Airport traffic has three major components. Planners must appreciate their special characteristics.
Each group has the following distinct patterns and needs:

- *Originating and terminating travelers*, who have only one access trip per flight, since either their departure or arrival is by air (however, their access may involve a round trip if someone delivers them to the airport, or if the taxi is not allowed to pick up at the airport).
- *Employees*, who commute daily to and from the airport.
- *Supply, delivery, and other commercial vehicles* that service the airport.

Each of these categories of airport traffic normally accounts for at least about 20 percent of the total trips to the airport. This fraction depends strongly on local conditions, however. If the airport is a transfer hub, it will have relatively few originating and terminating passengers and thus these will account for a smaller proportion of the whole. Similarly, if the airport is a maintenance or training base for an airline, it will have more employees and other commercial traffic. Regardless of these significant variations, each of the categories of traffic is of the same order of magnitude. This observation is important because it corrects the mistaken popular definition of airport access that focuses on the passengers and neglects the other traffic. Airport passengers are only part of the airport access problem.

A misplaced focus on the number of access trips due to passengers is easy to understand. The number of airport passengers may easily be several times the population of the metropolitan area it serves. For example, the annual number of passengers through the Boston and Paris airports is six to eight times the local population. The relative number of airport passengers is large even in areas where air travel is less frequent. Thus, the number of airport passengers through Jakarta and Mexico City about equals the population of these cities. In contrast to the apparent importance of passengers, the number of employees at an airport is relatively low. Although the busiest airports have tens of thousands of workers, they are generally less than one per thousand passengers (Table 17.4). Such figures easily concentrate attention on passengers.

Airport	Employees	
	Average Daily	Per 1000 Annual Passengers
Chicago/O'Hare	40,000	0.64
Los Angeles/International	40,000	0.72
Denver/International	19,000	0.40
Boston/Logan	14,600	0.58
Houston/Intercontinental	14,400	0.37
Salt Lake City	13,000	0.66
Seattle Tacoma	11,400	0.37
San Francisco/Oakland	10,500	0.28
Las Vegas	8,000	0.21
Phoenix	8,000	0.22
Portland (Washington)	5,000	0.39
Miami/Ft. Lauderdale	4,700	0.15
San Diego	3,000	0.18
Omaha	2,500	0.61
Sacramento	1,500	0.17
Los Angeles/Wayne	1,000	0.11

Source: Based on ACRP, 2010, and FAA Airport Passenger Data.

TABLE 17.4 Sample Data on Employees at U.S. Airports in 2009

Data on the number of people are misleading, however, because they neglect the frequency of travel. Originating or terminating passengers each make one trip to or from the airport. They each account for about one vehicle trip (more than one if a driver has to return empty, less than one when they share rides—see Table 17.5). Employees and other commercial traffic, however, make round trips every day of the year—and they may make additional trips to go on various errands. Each employee thus makes about 500 or more access trips a year. This frequency of compensates for their low number and makes employee traffic the same order of magnitude as passenger traffic (see Example 17.1).

Access Mode	Trips
Pickup or drop-off	1.29
Taxi	1.09
Parking	0.74
Rental car	0.69
Courtesy bus	0.33
Scheduled bus	0.10

Source: Shapiro et al., 1996, from a Massport study.

Table 17.5 Vehicle Trips per Passenger by Access Mode

The salience of passengers in the airport access problem is due to their particular characteristics, not their number. First, they concentrate at the "main door" to the airport, whereas employee and commercial traffic disperse to locations all around the airport, such as cargo areas and other facilities away from the main passenger buildings. Second, passengers tend to be anxious. They need to make a flight and are often unfamiliar with the airport. Finally, the passengers are important customers. Thus, although they are only part of the airport access problem, they do command attention.

Example 17.1 Consider an airport serving 10 million total passengers a year. The number of passenger trips to and from the airport might be between 7 and 9 million a year, because transfer passengers do not leave the airport. This implies an average in the range of 20,000 to 25,000 passenger trips a day to and from the airport.

Suppose that, on average, 5000 people work at this airport every day, which is 0.5 employees per thousand passengers (see Table 17.4). They include the airport operators, the airline employees, and the staff working for concessionaires, hotels, freight companies, car rental agencies, and the like. If all employees commute both ways, they account for about 10,000 trips a day on average.

Trips by commercial vehicles depend on the situation. They include suppliers of all sorts, deliveries and pickup of cargo, persons coming to sell products, and other visitors. These easily account for as many trips as employees do.

These conditions imply about 40,000 trips to and from the airport every day. Passengers account for about half of these trips. Employees and commercial traffic each generate about a quarter of the airport access traffic in this case.

Distribution of Airport Access Traffic

The general rule is that only a small fraction of the traffic to and from an airport goes to or comes from any specific destination. Most of the ground traffic for the airport spreads out over a wide area. This fact is crucial for proper understanding of the airport access problem.

The center of the city generates only a small fraction of the trips to and from the airport. Passengers living in the metropolitan area typically start their trips to the airport from their homes, even if they are taking a business trip. Furthermore, their homes are generally located in the suburbs or at least in apartment buildings some distance from the financial or commercial center of the city. Travelers from elsewhere may be going to offices or hotels in the city center. However, if they are visiting friends or relatives, they are likely to go to their homes in the suburbs. Even business travelers may not be destined for the city center, because they are either calling on companies in industrial areas distributed around the metropolitan area or seeking less expensive hotels at some remove from the city center. According a survey, only about 8 percent of the passengers originating or terminating at Boston/Logan came from or went to the whole city of Boston, spread out over some 20 square miles (Leigh Fisher et al., 2000). Only a fraction of these trips was associated with the city center. In general, the percentage of passengers coming from or going to any city center is small.

Employee and commercial traffic goes primarily to the edges of the city. This traffic links the airport to the parts of the metropolitan area that are less expensive for housing and industry. Only exceptionally is it connected directly to the center of the city.

Example 17.2 The 10-million-passenger airport of Example 17.1 generated about 40,000 access trips a day. Suppose that 20 percent of the passengers connect with the city center. Suppose also that essentially none of the employees or commercial visitors goes there. This means that about 4000 to 5000 passengers a day go to or come from the city center. This is roughly 10 percent of the total.

Overall, only about a tenth of the airport access traffic wants to come from or go to the city center. Example 17.2 shows how this works out. The exact amount depends, of course, both on how planners define “city center” and the local situation. In any case, the fraction of airport access traffic connected with the city center is small. Most of the airport access traffic is spread widely over the metropolitan area.

The fraction of traffic destined for the city center tends to decrease for the largest airports. This is because the airports with the most traffic, such as Atlanta, Chicago/O'Hare, London/Heathrow, and Tokyo/Narita, generally are transfer hubs. They thus have relatively fewer originating or terminating passengers and therefore relatively fewer trips to the center of the city.

Preferences of the Users

Individually, a prime concern for passengers is getting to the airport on time. They tend to be most concerned about the reliability of their travel time to the airport. Missing a flight can have severe consequences in terms of the following:

- Delays until the next flight—hours or a day later
- Costs due to the need to replace or upgrade tickets
- Missed connections or appointments

Reliable connections are much more important than speed. To deal with unreliable access and to get to their flights on schedule, passengers routinely allow substantial extra time for their trip to the airport. This provides a buffer to ensure that they will not miss their departures. Accepting added time allowed to compensate for unreliable connections is equivalent to accepting a much slower average speed for the trip. Such actions demonstrate the passengers' willingness to sacrifice speed in favor of reliability.

Collectively, passengers also want access systems that can distribute them to their destinations spread widely over the metropolitan area. Employees likewise share this requirement for a system that can take them between the airport and their homes scattered over the suburbs and other bedroom communities. Systems that serve only a few points and do not connect conveniently with a wide network of transportation do not adequately address the requirements of most of the airport access traffic.

Price can be a significant consideration. It is generally a secondary concern for passengers, compared to reliability and accessibility. Business passengers paying hundreds or thousands of dollars for an airfare may be prepared to pay reasonable amounts to get to and from the airport. Families and others may be more careful. Moreover, the price of the airport access trip is salient for employees, who have to pay for daily round trips. In short, whereas some passengers may be willing to pay for premium service, many passengers and most employees cannot afford expensive access transport.

Price considerations tilt passengers and employees toward the use of automotive forms of access. When they have access to privately owned vehicles, these will generally provide cheap (and convenient) access. Passengers may also favor taxis because these vehicles normally cost the same for as many people as can fit inside—contrary to other forms of access, for which each traveler has to pay an additional fare. This makes the average cost per trip per person much lower for the many passengers who travel in groups, as with family, friends, or associates. When people consider the total cost of getting to the airport, automobile access can appear much more economical.

Overall, the users of airport access systems need reliable systems that can distribute them broadly throughout the metropolitan area. Most of the market is also unlikely to pay premium fares for individual trips so long as they have some form of inexpensive automotive transport available (see de Neufville, 2006).

Needs of Airport Operators

Airport operators often feel obliged to promote rail and other forms of public transport. Some major airports face an inability to build highway capacity for the airport and must either develop rail access or face gridlock. Toronto/Pearson, for example, foresees that as the city grows, it will not be able to count on the local highways to deliver traffic reliably. It has therefore acquired and is reserving right-of-way for a connection to the local mass transit system. Likewise,

- New York/Kennedy recently used APMs to tie into the metropolitan rail and mass transit network.
- Japan connected Tokyo/Narita to several rail lines.
- BAA built the Heathrow Express connection between London/Heathrow and the Central London rail and mass transit network.

Furthermore, rail links are sometimes not a choice but a necessity to ensure maximum use of the airport. Major island airports such as Hong Kong/Chek Lap Kok, Osaka/Kansai, Seoul/Incheon, and Nagoya/Chubu have rail links to the mainland.

Many airport operators face powerful pressures to restrict automotive access to the airport. Because they often are the largest single destination and economic activity in a metropolitan area, environmental interests see them both as good targets for environmental improvements and as “deep pockets” that can afford to subsidize public transport. Thus Massport, the operator of Boston/Logan, is paying for regional bus service and water taxis to induce a few people not to come by car. Meanwhile, however, its economic interest is to attract cars, because it derives about a quarter of its total revenues from parking and fees from rental cars (see Sec. 17.4).

Over the last generation, operators of many major airports have installed tracked forms of airport access. In the United States, they have focused on metropolitan rail systems (see Table 17.1). In Europe and Japan, where people use long-distance railroads regularly, many airport railroads connect to the national intercity and even bullet trains (see Table 17.2).

Correspondingly, the Asian and European systems tap into existing markets and generally serve a higher fraction of the airport traffic than the U.S. airport rail systems (Table 17.6).

United States		Europe and Asia	
Airport	Market Share	Airport	Market Share
Washington/Reagan	14	Oslo/Gardermoen	43
Atlanta	8	Tokyo/Narita	36
Chicago/Midway	8	Geneva	35
Boston/Logan	6	Zürich	34
San Francisco/Oakland	4	Munich	31
Chicago/O'Hare	4	Frankfurt/International	27
St. Louis	3	London/Stansted	27
Cleveland	3	Amsterdam	25
Philadelphia	2	London/Heathrow	25
Miami/International	1	Hong Kong/Chek Lap Kok	24
Washington/Baltimore	1	London/Gatwick	20
Los Angeles/International	1	Paris/de Gaulle	20

Source: Transit Cooperative Research Program, 2000.

Table 17.6 Market Share of Passengers Served by Rail Systems in the United States Compared to Those in Europe and Asia

17.3 Cost-Effective Solutions

The Issue

Highways provide the dominant mode of airport access. Automobiles, taxis, vanpools, and buses dominate the traffic. These are the people's choice. For most people going to the airport, automotive vehicles provide the best value for money. Passengers and employees appreciate that this form of airport access fulfills the essential function of distributing traffic conveniently throughout the metropolitan area. It also caters to their varied needs by providing a range of more or less convenient and luxurious service, more or less expensively.

Airport owners appreciate that, for all but the largest airports in the most congested areas served by an extensive rail network, highways are relatively easy and inexpensive to build—compared to rail connections. Moreover, highways create profitable demand for parking, which for many airports is a major source of revenue (see Sec. 17.4). Compared to tracked forms of airport access, such as railroads, highways generally require less investment. Highways are everywhere, and the airport operator often needs only to build a short connection to tie into the existing system.

The issue for planners is this: What kinds of complementary airport access systems should airport operators develop, under what circumstances? New rail projects—either intercity railroads or metropolitan transit systems—can be very expensive. Reportedly, the development of the Arlanda to Stockholm/Arlanda airport cost $600 million for a distance of 25 miles—$24 million per mile, and that of the Heathrow Express to London/Heathrow from London cost £450 million ($675 million). (See Fig. 17.1.) These were twentieth-century costs. More recently, the AirTrain connection between New York/Kennedy and the Long Island Railroad and the New York Metropolitan Transit Authority had a budget of over $1.6 billion. (Rail projects, being so expensive, can also be highly controversial and difficult to implement. The project for New York/Kennedy was under discussion for a quarter-century before construction began. Likewise, the railroad system to Tokyo/Narita opened over 20 years after the airport operator built the airport and the railroad station.) Rail systems to the airport can be highly visible and impressive, but when and where do they make sense?

FIGURE 17.1 Heathrow Express connecting London/Heathrow and Paddington Station, London. (*Source*: BAA plc.)

Door-to-Door Analysis

When considering a new airport access system, planners first need to estimate its prospective number of users. If only a few people use the system, it will not be serving its intended function. Moreover, if traffic is low, the service is likely to be infrequent, and this discourages prospective riders. Under such circumstances, fares may not cover the operating costs, let alone repay the initial investment, and the system could represent an enormous financial loss. Adequate traffic is essential for the operational and economic success of any airport access system.

It is generally difficult for alternative airport access systems to compete with highways. They typically enter the market at a disadvantage, after the highway systems are well established. By the time a new service opens, people have their cars, and fleets of rental cars, taxis, and buses exist. When it begins service, it may have few passengers and thus may find it difficult to afford to offer the frequent services that will make it competitive with automobiles that operate on demand. To make a rail or other mass transit system successful, planners need to define carefully how this complementary system will be competitive.

A door-to-door analysis of trips by competitive transportation modes establishes their relative advantages. It focuses on the total time, cost, and convenience of each mode for the entire trip, door to door, between the home or office and the airport. The need to consider the door-to-door experience is a most important point, often neglected. Travelers look at the entire experience of a trip, not just a single segment. Popular presentations of new rail systems, however, frequently focus on their speed from station to station and lose sight of the fact that people generally have to wait for departures and spend time and money getting to and from the station. For example, initial publicity about the rail connection to New York/Kennedy boasted that it "will cut the travel time to Manhattan to 45 minutes." This was true if the traveler entered the station just before the train departed, made the connection at Jamaica station with no delay, and wanted to go to Pennsylvania Station. Normally, however, travelers walk to the station, wait for the train, and then have to find their way from the train stop to wherever they want to go—which may take them another 20 minutes and cost $10 or more. All these connections add considerable time and cost to the trip, as well as the inconvenience of having to get in and out of several vehicles. Considering the total door-to-door experience, the trip by the fast train may be, for many people, slower and more expensive than a taxi directly between home or office and the airport.

To understand the competitive position of a proposed new airport access system, analysts need to estimate carefully the total door-to-door time and cost of the competitive modes.

In addition to the costs associated specifically with the mode (the taxi fare, the rail ticket, or the parking charges), they need to add the time and effort it takes to get to these modes. They may have to include the cost of taking a taxi from a hotel to the railroad and then the wait to buy tickets and for the train to begin its trip. Example 17.3 illustrates how this can be done.

Analysts should also consider the trip costs for the owner of a rail or alternative system. Indeed, the fare the passenger pays for a service often does not represent its true cost. A public agency may consider this part of normal subsidy of metropolitan transit. However, a private airport operator has to be particularly concerned with covering costs.

Rail Solutions

Comparative analyses of door-to-door travel times and costs generally demonstrate that rail systems of airport access can be competitive with highway modes in favorable circumstances. Among the factors that favor competitive rail service are the following:

- *Airport size*, to generate enough passengers to cover costs and to sustain frequent service that lessens the time people have to wait for trains
- *Existing local rail service*, which lowers the cost of the airport connection (this is the basis for the rail service at Amsterdam, Frankfurt/International, Geneva, London/Gatwick, Lyon, and Zürich)
- *Easy connections to a wide metropolitan transit system*, as exists at London/Heathrow, Paris/de Gaulle, and Washington/Reagan
- *Difficulty of automobile access to the airport*, a factor for airports on man-made islands, such as Hong Kong/Chek Lap Kok and Osaka/Kansai, and for distant airports such as Oslo/Gardermoen or Kuala Lumpur/International.

Example 17.3 Consider the Heathrow Express service between Paddington Station in Central London and London/Heathrow. It offers direct ride to the airport on trains leaving every 15 minutes during most of the day. It delivers one of the best services of its kind.

Table 17.7 estimates its door-to-door cost for the traveler from a central London hotel to an airport check-in counter, for peak and off-peak traffic. Table 17.8 shows a comparable estimate for a taxi journey. These approximate numbers are based in fares and estimated times published in 2012.

Journey Element	Cost (£) (£1 ~ US$1.5)	Time (min)	
		Peak	Off-Peak
Taxi to station	15	20	10
Get to train		5	5
Buy ticket	19	3	3
Wait for train		7	7
Train ride		15–27	15–27
Walk to check-in		5	5
Total	**34**	**55–67**	**45–57**

TABLE 17.7 Example Door-to-Door Times and Cost for a Passenger to London/Heathrow from a central London Hotel, on Heathrow Express

Journey Element	Cost (£) (£1 ~ US$1.5)		Time (min)	
	Total	Per Person	Peak	Off-Peak
Taxi to airport	65	22	60	30
Walk to check-in			2	2
Total	**65**	**22**	**62**	**30**

TABLE 17.8 Example Door-to-Door Time and Cost for a Passenger to London/Heathrow from a Central London Hotel by Taxi, in Total and per Person for a Group of Three

As the comparison of the two tables indicates, the Heathrow Express does not offer competitive service to many travelers. Although the train moves quickly and is reliable once a person gets to the station, making that trip through central London is not easy. The service is cheaper and faster than a taxi for people traveling alone at the peak hour. However, it is generally less convenient than a taxi because it requires repeated loading and unloading of bags. It is not economical for three or more persons.

Conversely, the same analyses demonstrate that rail access systems are difficult to justify for most airports, even in big cities. They cost too much. There are no easy connections to an efficient transit system. They serve too few people. They do not provide enough value for the money spent. They are not cost-effective. As Table 17.6 indicates, about half the rail connections to airports in the United States contribute only marginally to the solution of the airport access problem. ACRP (2013) provides guidance on planning for rail access to airports in the United States.

Highway Solutions

For most airports, rubber-tired modes of transport provide the most cost-effective forms of airport access. These systems can flexibly serve the entire metropolitan area without enormous investment. Operators can adjust the vehicles and routes of these systems to meet the needs of the range of travelers and employees. The vehicles can range from private automobiles and taxis to various forms of collective transport such as vans, luxury motor coaches, and ordinary busses.

Collective transport using rubber-tired vehicles is not glamorous and does not get the attention it deserves. Precisely because access systems using buses and vans distribute people effectively over the metropolitan area, they are small and spread out compared to rail systems that are big and concentrated. Public transportation based on buses and vans has proven, however, to be successful at many locations. For example,

- Courtesy buses and shuttle vans connect major U.S. airports to local destinations.
- London/Heathrow connects to its metropolitan areas with large fleets of buses.

Coping with rubber-tired traffic requires substantial effort. A 10-million-passenger airport may receive 40,000 vehicles a day, or about 4000 vehicles in the peak hour (see Example 17.1). Three lanes of traffic each way are necessary to serve this level of traffic. These figures give an impression of the importance of the effort that has to be made to provide suitable airport access. Readers should note, however, that the numbers cited are only illustrative. The actual amount of highway required depends on local conditions. The number of lanes needed is relatively less for transfer hubs that cater to many passengers who do not need airport access. It is also less in regions in which people do not use their own cars to get to the airport, because people either own relatively few cars or are accustomed to using public transport. Correspondingly, the access highways to airports in North America tend to be larger than they are elsewhere. These differences are only a matter of degree, however. Major airports inevitably require extensive connections to the regional highway network.

Major airports need to develop strategies to reduce the number of vehicles coming to their airports. Although cars provide parking revenues, they cause large costs in terms of the miles of curb frontage (see Sec. 16.5), the cost of elevated roads and garages, and their use of valuable space around crowded passenger buildings. Moreover, private cars cause air pollution (see Chap. 6). Airport operators thus need to encourage the use of various forms of high-occupancy vehicles. Rail access can be one good approach in some circumstances. Often, however, a range of access modes provides the best overall service. Analysts and planners need to consider each of the possibilities and try to optimize the overall mix of services they can provide to the variety of prospective customers.

17.4 Parking

Modern airports need a lot of parking. In 2012, Dallas/Ft. Worth advertised over 28,000 spaces! How the airport should provide for parking is a major issue. Airport operators need to think about both the quantity of spaces and also the range of ways they can serve this market.

Traditionally, planners estimated the amount they should provide as a proportion of the total number of passengers, as presented by the FAA (1988). This approach is no longer especially useful for many reasons. Most obviously, because many passengers simply transfer at an airport, they are not potential customers of parking. Second, context matters. In some environments, the availability of rail or taxi alternatives lowers the number of travelers who will use their own car to get to the airport (this is the case for Singapore or Tokyo). In others, access by private vehicle may be prohibitive (the case of island airports such Hong Kong/Chek Lap Kok and Osaka/Kansai). Finally, on-airport parking may face extensive off-airport competition from long-term lots (Chicago/O'Hare is surrounded by about eight facilities of this kind). Having said this, we can nonetheless provide an order of magnitude figure for major airports in developed countries. Using Dallas/Ft. Worth as a reference, we arrive at a crude figure of about 500 spaces per million passengers. The actual number needed or desirable depends on local circumstances, such as the level of automobile ownership in the population, the availability of public transport, and current environmental policy. In any case, operators of major airports need to consider thousands of parking spaces.

Airport parking is a big business. It generates substantial revenues. In the United States, according to ACRP (2009), in 2007 parking accounted for an average of 18 percent of total airport revenues for the largest airports, up to an average of 28 percent for small hubs. For the 35 largest airports in 2010, parking generated over $1.75 billion in total revenues and accounted for 7 percent of total airport revenues (see Table 8.5). As reported by their annual reports, the total revenues from parking in 2010 were over $97 million for Dallas/Ft. Worth and $93 million for Chicago/O'Hare.

Airport operators need to recognize that the market for parking has at least six distinct segments. The best designs will provide the facilities appropriate to each. The categories are the following:

1. Short-term parking for up to a few hours, used primarily to pick up arriving passengers
2. Premium parking, offering a variety of special services
3. Structured parking close to passenger buildings, mostly serving persons on short trips or business travelers who can afford this expensive facility

4. Long-term, remote, and less expensive parking (often provided privately near the airport)
5. Rental car parking
6. Employee parking

Additionally, airports increasingly find it expedient to establish small cell-phone lots for persons waiting in their cars to pick up arriving passengers (see Sec. 16.5)).

Short-Term Parking

Hourly parking serves drivers meeting arriving passengers and, sometimes, delivering passengers. These people want spaces close to the passenger buildings, so that they will not have to carry bags a long way. If this is not available, they will tend to wait at the curb in front of the building, circle around the airport, or use cell-phone lots where available. Short-term parking thus both meets a demand and helps relieve the congestion on the curbs in front of the passenger building (see Sec. 16.5).

To provide space for short-term parking, the airport operator needs to set aside some convenient area, often in a parking garage, segregated from daily parking. Operators can do this by installing separate entrances and exits. Toronto/Pearson, for example, has short-term (Express) parking on the floor of its parking garage that is closest to the entrance of its Terminal 1. The airport operator ensures that spaces are available for hourly parkers by charging high hourly rates that are affordable for persons staying only an hour or so but prohibitive to anyone leaving a car for one or more days.

The total number of spaces dedicated to hourly parking can be relatively small. Because people park in these spaces for only a short time, a single space is used many times a day. As emphasized in Sec.15.4, when the dwell time is short, the capacity is large. On average in the United States, a short-term parking space serves over 1000 cars a year, whereas a long-term space serves only about a quarter that number.

Short-term parking provides airport operators with two significant benefits. It relieves crowding at the curbs in front of the passenger buildings. It is also profitable, because the revenues per space are high. By differentiating the market for parking into those who do and do not want to pay for convenience, the airport operator can charge premium prices for the most desirable parking spaces.

Premium Parking

The introduction of premium parking is a recent innovation. Airport operators traditionally charged uniform prices for all spaces in a given location. Many airport operators now provide a range of high-margin, complementary services to clients willing to pay for them. These include valet parking, reserved close-in or shaded spaces, special access to parking areas, and even car detailing and similar attentions. (See ACRP, 2009.) For example, Boston/Logan reserves a floor of a close-in garage for subscribers; London/Heathrow provides personalized, automated transport to "business" parking (see Sec. 17.5).

Structured Parking

Multilevel garages are generally a practical necessity at busy airports. The demand for convenient space next to the passenger buildings is high. The only way to make efficient use of this valuable property is to build it up. The busiest airports tend to feature four-or five-level garages with spaces for thousands of cars.

Multilevel garages are expensive. The capital costs for the parking structure itself average about $20,000 per space as of 2009. The average total costs may easily be double, once one includes the capital costs for modern equipment (automated pay stations, entrance and exit controls, etc.) and the soft costs for design, planning, and project management. Specific projects may cost significantly more due to the difficulties of building around areas in active use and connecting to them. Thus Boston paid over $200 million to add 2500 spaces to its existing facilities—over $80,000 per space. These figures translate into high costs to the users (ACRP, 2009). The actual daily charges for parking depend on the financial objectives of the airport operator, but they are easily $20 a day and more.

Structured parking thus normally serves persons on short trips or business travelers who can afford this expensive facility. Its cost may be excessive for parkers who might want to stay a week or more. This reality motivates the development of cheaper parking for long-term users.

Long-Term Parking

At-grade lots provide less expensive parking suitable for long-term stays. At busy airports these are necessarily some distance away, typically in old industrial sites (as around Los Angeles/International) or other open spaces (as around Boston). These facilities often develop independently, in competition with the airport operator. When this is the case, they constrain the airport operator's profits from parking.

Rental Car Parking

Rental car agencies often require substantial parking for their fleets. They typically operate out of individual lots in various remote areas tucked into corners of the airport, or even off airport. The situation can be confusing to renters, who have to find their way to and from these lots and may not know the area. Such lots also require a constant flow of courtesy vans or buses to ferry clients to and from the passenger buildings. These require special parking spaces and generally congest the curb areas.

Airport operators in the United States are increasingly building consolidated rental car facilities, Conracs, as discussed in Sec.16.5. These are multilevel garages that house several rental companies. This innovation simplifies airport access at congested airports. By having a single major facility, the airport operator can direct all rental cars uniformly, thus reducing confusion, and can implement a single transfer system between the passenger buildings and the rental car building. This can either be a bus service, as at Washington/Reagan, or an APM, as at Atlanta and San Francisco/International. Conracs are likely to become prevalent at major airports.

Employee Parking
Employees may require hundreds, even thousands of spaces at the busiest airports. Many of these spaces will be dispersed around the edge of the airport, close to the cargo centers, maintenance bases, and other facilities distant from the congested passenger buildings. However, the airport operator does have to establish parking for persons who work in and around the passenger buildings, and links between the passenger complex and the parking or public transport locations the employees use. Bus services usually provide these connections.

17.5 Automated People Movers

Major airports typically feature a network of bus routes. Courtesy buses serve remote parking, local hotels, and move airline and airport employees around the airport as needed. Numerous shuttle services may link the airport with homes and businesses throughout the metropolitan area. Major bus companies may also serve a variety of downtown and suburban destinations. This traffic can be confusing to the passengers and congest the curb space in front of the passenger buildings. Airport operators need to organize these networks coherently.

To deal with the confusion of access and distribution systems around the airport, many airport operators have invested in automated transit systems to connect principal points of access to and on the airport. These are known generically as *automated people movers* (APMs). The use of APMs for moving people around the airport is perhaps the single most important innovation in the design of airport layouts since the early 1970s. It has enabled airport designers to disperse aircraft across the airport, thus facilitating their movement while still providing passengers with short walking distances and travel times within and between the passenger terminal buildings. Specifically, this development has led to implementation of mid-field concourses as a standard feature of most new major airport complexes.

In the 40 years since the first airport APM opened at Tampa in 1971, 46 (mostly large, international) airports throughout the world have acquired APMs. Of these, less than half are landside systems that connect the terminals with other facilities on-and off-airport and most are airside systems that connect multiple terminals or terminals and concourses. For many years U.S. airports built most airport APMs, but airports in China and the Middle East are increasingly adopting this technology. Europe and Japan also have a number of airport APMs. Airport APMs vary from short (1000–1500 ft) single-and dual-lane shuttle systems to loop and pinched-loop systems with over 5 miles of operating guideways.[1] About 80 percent have self-propelled vehicles and almost all are rubber tired. More recently, cable-propelled systems are being built and a personal rapid transit (PRT) system is now in operation at London/Heathrow Terminal 5.

This section introduces airport APMs. Sections 17.6 and 17.7 discuss specific landside and airside applications, highlighting the different design requirements and their implications for this technology. Section 17.8 discusses issues in planning airport APMs.

APM Technology Characteristics

Airport APMs share some basic characteristics. They are as follows:

- *Automated*: Operating without drivers and controlled by central computers that ensure safe and efficient operation with short (2- to 3-minute) headways between trains to minimize wait times.
- *Safe and reliable*: Featuring automatic control subsystems designed on rigorous fail-safe principles, most airport APMs meet an availability level of over 99.9 percent. APM guideway systems typically have emergency walkways should there be a failure or emergency on board.
- *Designed for people*: Airport APMs normally meet requirements of the Americans with Disabilities Act (ADA), including level boarding, small (elevator-like) gaps between station platforms and vehicle floors, and wheelchair accessibility and parking locations. APMs also carry passengers' baggage: landside systems all baggage, sometimes with baggage carts; airside systems only have to accept carry-ons.
- *Total systems, not a conglomeration of subsystems provided separately*: One company typically manufactures most subsystems, and supplies and integrates them all in the design and installation process. These subsystems include the vehicles, automatic train control, power distribution, audio and video communications, and station, guideway, and maintenance equipment.

All APMs operate on special-purpose guideways reserved exclusively for their use. This is due to safety issues related to automatic control and high-voltage power rails mounted on the guideway. APMs run as trains consisting of one to six vehicles, depending on the technology and the capacity requirements. They also operate horizontally or on minimal inclines. They usually run on elevated structures or in tunnels, although some have at-grade guideways or a mix of these three.

Like elevators, all airport APMs have both station platform doors and vehicle doors that operate together. In part this is a safety issue, particularly because, unlike urban systems, many passengers are unfamiliar with such transit. Additionally, because almost all stations are climate controlled, station platform doors help maintain the proper temperature.

Outside contractors design and build the physical infrastructure for APMs. The stations are often integrated with the design of the airport passenger and other buildings and so conform to a larger architectural plan. Civil engineering consultants initially design guideway structures and maintenance facilities generically but then tailor them to the specific APM technology.

APM systems differ technologically from each other. This has some important implications for their procurement:

- As each supplier "owns" its APM technology, this means that, with one possible exception, the supplier of the initial system will be responsible for any expansion to be done.
- APM systems should be procured using performance specifications defining needs in terms of capacity, level of service, code, and other project requirements.
- The initial supplier will operate and maintain the system, at least initially, and is likely to continue maintenance indefinitely, although the airport may take over operational control.

Beyond these common features, APMs come in a wide array of technologies and models. They have different suspension systems (rubber tires, steel wheels/rails, and air levitation) and operate on various guideway structures (usually concrete or steel like roadways, but also space frame steel tubes). Vehicles usually run on top of the guideway, but Düsseldorf airport has a suspended APM. Most (80 percent) are self-propelled with on-board motors; the others are pulled by cables (like elevators, except horizontally) or in-guideway fixed linear induction motors. APMs come in a wide range of sizes. Most airport APM vehicles are 10 m (40 ft) long, but some are somewhat smaller. At the extreme, there are monorails with one to six small cabins and, most recently, four to six passenger PRT vehicles.

APM Suppliers

The number of active suppliers of service-proven, reliable APM systems is limited. As of 2012, the suppliers of the typical large rubber-tired, self-propelled APMs include the following:

- Bombardier Transportation Holdings has provided nearly half of the airport APMs operating in the world today. Some systems use the older CX-100 vehicle (recent examples at Orlando, Denver, Sacramento, London/Gatwick, and Frankfurt/International); newer ones have the Innovia vehicle (Dallas/Ft. Worth, Phoenix, and Jeddah). Bombardier also provides a steel wheel-rail vehicle, the ALRT, which operates at New York/Kennedy, and has provided a small monorail APM at New York/Newark and Tampa (in a parking garage). See Figs. 17.2 to 17.4.

FIGURE 17.2 Bombardier CX-100 at Sacramento. (*Source*: Harley Moore, Lea+Elliott, Inc.)

FIGURE 17.3 Bombardier Innovia at Dallas/Ft. Worth. (*Source:* Harley Moore, Lea+Elliott, Inc.)

FIGURE 17.4 Bombardier ALRT at New York/Kennedy. (*Source*: Harley Moore, Lea+Elliott, Inc.)

- Mitsubishi Heavy Industries, that has provided its Crystal Mover vehicle at Washington/Dulles, Atlanta (Conrac system), Miami, Singapore/Changi, and Hong Kong/Chek Lap Kok. See Fig. 17.5.

FIGURE 17.5 MHI Crystal Mover at Singapore. (*Source*: Harley Moore, Lea+Elliott, Inc.)

- IHI/Niigata has systems at Osaka/Kansai and Taipei/Taoyuan. In 2011 it initiated a significant expansion to the APM at Hong Kong/Chek Lap Kok (where unusually both MHI and IHI vehicles operate on the same guideway). See Fig. 17.6.

FIGURE 17.6 IHI New Transit at Osaka/Kansai. (*Source*: Harley Moore, Lea+Elliott, Inc.)

- Siemens (previously Matra) has its VAL APMs at Chicago/O'Hare, Paris/Orly, and Paris/de Gaulle. See Fig. 17.7.

FIGURE 17.7 Siemens VAL at Chicago/O'Hare. (*Source*: Harley Moore, Lea+Elliott, Inc.)

Cable-driven systems can work well for relatively short distances and in a shuttle configuration between two points. This technology has several potential advantages. Their motors are in drive rooms at the end of the track, rather than on-board, so the vehicles and their supporting structure can be lighter and less expensive than self-propelled APMs, and high-voltage propulsion power rails along the guideway are not needed. In addition, the drive mechanisms and passive vehicles can be easier to maintain. Before 2011, this technology was not competitive for systems longer than about 600 m (2000 ft), but with multiple cables and drives and releasable grips they can compete successfully with self-propelled technologies in certain circumstances. For example, the 3.5-mile (5-km) BART-Oakland Airport Connector is to use cable technology with a cable transfer point (with the trains stopped) at an intermediate station.

As of 2012, there were two suppliers of cable-propelled APMs, both with a primary main business providing ski lifts and gondolas:

- Doppelmayr Cable Car Company (DCC), an Austrian company, has airport APMs at Birmingham (U.K.), Mexico City, and Toronto/Pearson, and in 2012 was installing others for the New Doha International airport (inside the terminal building) and at San Francisco/Oakland. See Fig. 17.8.

FIGURE 17.8 DCC at Toronto/Pearson Airport. (*Source*: Harley Moore, Lea+Elliott, Inc.)

- Leitner-Poma (previously Poma-Otis), an Italian-French company, has—among others—air-levitated APMs at Tokyo/Narita, Detroit (inside the terminal, see Fig. 17.9), and Minneapolis/St. Paul.

FIGURE 17.9 Poma-Otis at Detroit/Metro Airport. (*Source*: Harley Moore, Lea+Elliott, Inc.)

PRT systems feature small vehicles (four to six passengers) and direct station-to-station service in a network, rather than a route-based service. The first airport application opened in 2011 at London/Heathrow. Built by a British company, ULTra, it serves a station in Terminal 5 and two stations in a "business" remote parking lot (Fig. 17.10). Two other companies are developing PRT projects off-airports: the Dutch 2getthere and the Korean Vectus. PRT systems have limitations for airport use: Vehicles require passengers to sit and limit baggage and groups, and the system limits overall capacity and routing. However, PRT systems offer some advantages: The small vehicles fit in tight spaces and provide direct point-to-point service.

FIGURE 17.10 Ultra PRT at, London/Heathrow International Airport. (*Source*: Harley Moore, Lea+Elliott, Inc.)

17.6 Landside APMs

Airport designers can locate APMs either landside or airside. They serve different functions and have different design requirements in either application.

Landside APMs tie together airport facilities and access systems. They are "landside" because they are used by the general public, not just the passengers, crews, and employees who have been cleared through airport security. Landside APMs connect different unit terminals and the terminals with remote passenger and employee parking, consolidated rental car facilities, and airport access facilities such as rail stations and bus terminals. Table 17.9 lists airports with landside and airside APMs. About a quarter the airports have both airside and landside systems.

Landside APMs substitute for bus transport systems. Although bus systems have the great advantage of being inexpensive to build, easy to implement, and flexible, they have many disadvantages. Their level of service is relatively low, particularly when there is traffic congestion on the airport. They require many specialized drivers and can thus be expensive to operate and maintain. If they operate on conventional fuels, they emit air pollution, particularly as they idle at stops. Moreover, busses are neither modern nor glamorous, and most require climbing up stairs rather than level, fully ADA-accessible boarding. The development of landside APMs thus often depends on the objectives of the airport operator. Is it willing to spend large sums for higher levels of service, greater capacity, and cleaner air? In some cases, it may not have a choice.

Landside APMs have some different characteristics and requirements from airside systems. These, and design considerations can include the following:

- *Speed*: Given the distance these systems often must travel, cruise speeds of 50 to 80 km/h (30–50 mi/h) can be important for both capacity and travel time.
- *Seating*: Because of their longer routes, landside vehicles often have more seats than the same vehicle used airside. This reduces vehicle capacity but improves service levels.
- *Baggage*: Landside APMs that serve rental car and airport access facilities carry passengers and their baggage. The number and sizes of bags can be significant and reduce vehicle capacity. Some airports allow baggage carts on the APMs, which further reduce their capacity and require carts with brakes for safety. Baggage and carts also require wider doors and larger cabins for circulation.
- *Guideway design and location*: As landside systems typically cross over roadways and other facilities, they are usually on elevated structures, as opposed to airside APMs that are often in tunnels under the terminals and aprons. The guideway layouts are typically loop or pinched loop.
- *Routing*: Direct routes are preferable, but given the dispersal of the facilities served, the routes of landside APMs can be more complex than those for airside systems, including multiple routes on the same guideway. It is often cost-effective to have stations in the different areas. For example, the New York/Kennedy APM has one route that serves remote parking areas, the consolidated rental car facility, and a subway station; another route serves a combined off-airport railroad/rail transit/bus station at Jamaica.
- *Technology*: Given the need for speed, size, and route distance, airports normally select large, self-propelled APM technologies. The San Francisco/Oakland system will be an exception.
- *Passenger information*: Is generally much more complex for landside systems, particularly for those with multiple routes and stations. It can require a series of signs and station and train-based public address announcements, often in multiple languages at international airports.

Region	Airport	APM Type	
		Airside	Landside
Americas	Atlanta	X	X
	Chicago/O'Hare		X
	Cincinnati	X	
	Dallas/Ft. Worth	X	
	Denver	X	
	Detroit/Metro	X	
	Houston/Intercontinental	X	X
	Las Vegas/McCarran	X	
	Mexico City		X
	Miami/International	X	X
	Minneapolis/St. Paul	X	X
	New York/Kennedy		X
	New York/Newark		X
	Orlando/International	X	
	Phoenix		X
	Pittsburgh	X	
	Sacramento	X	
	San Francisco/International		X
	San Francisco/Oakland		X
	Seattle-Tacoma	X	
	Tampa	X	X
	Toronto/Pearson		X
	Washington/Dulles	X	
Europe	Birmingham		X
	Düsseldorf		X
	Frankfurt/International		X
	London/Gatwick		X
	London/Heathrow	X	X
	London/Stansted	X	
	Madrid	X	
	Paris/de Gaulle	X	X
	Paris/Orly		X
	Rome/Fiumicino	X	
	Zürich	X	

Middle East	Cairo		X
	Doha	X	
	Dubai	X	X
	Jeddah/King Abdul Aziz	X	
Asia	Beijing/Capital	X	X
	Hong Kong/Chek Lap Kok	X	X
	Kuala Lumpur/International	X	
	Osaka/Kansai	X	
	Seoul/Incheon	X	
	Singapore	X	X
	Taipei/Taoyuan	X	X
	Tokyo/Narita	X	

TABLE 17.9 Airside and Landside Airport APMs in 2012

U.S. examples of access-related APMs are in New York and San Francisco. New York/Kennedy has one of the largest land-side APMs with three routes (two as noted previously and another that circulates—in the opposite direction—among the unit terminals in the central terminal area). New York/Newark has a land-side monorail APM that passes through its three main passenger buildings, two rental car facilities, three parking facilities, and connects with the regional rail system. San Francisco/International has a landside APM that connects the consolidated rental car facility, the BART rail transit system, and four terminals. San Francisco/Oakland is also to have a connection to BART. In France, Paris/Orly has a connection to the metropolitan RER rapid rail network.

The desirability and cost-effectiveness of airport APMs for airport access must be considered carefully. The systems are expensive, often costing hundreds of millions of dollars for the operating system and its infrastructure. Their large extra cost over the alternative of bus routes between the passenger buildings and other facilities may not be justifiable. Boston/Logan, for example, decided it could not afford to pay over $300 million to eliminate its shuttle bus service that connected the unit terminals and the regional rail system. In many circumstances, however, people movers may be essential to the functioning of an airport. At some point, the buses necessary to move people around the airport cannot handle the loads, will overcrowd the roadways, and contribute excessively to air pollution.

17.7 Airside APMs

Airside APMs serve passengers and others who move through the airport on the airside of the security checks. They have been a prime characteristic of the designs of many of the airport since the late 1970s when Atlanta rebuilt its airport around the concept of midfield concourses (see Sec. 14.4). They typically connect a central passenger building with one or more midfield buildings and are often in tunnels under the buildings and aprons. Examples of these include Atlanta, Denver, Washington/Dulles, Seattle-Tacoma, Kuala Lumpur/International, London/Heathrow, Madrid, and Zürich. Dallas/Ft. Worth replaced its landside APM with an airside APM and, because the passenger buildings, roadways, and many other facilities were already built, this Skylink system is on elevated guideway with stations at the roof level of the terminal buildings. Tampa, Orlando/International, and Tokyo/Narita have terminal designs with elevated guideways between a central terminal and satellite buildings. Some airside APMs serve long passenger buildings: Osaka/Kansai (on the roof), Hong Kong/Chek Lap Kok (in a tunnel under the building), and Detroit/Metro (inside the building on the mezzanine level). Table 17.9 lists airports with airside APMs.

Airside APMs enable the development of midfield passenger buildings that offer significant operational advantages to the airlines in terms of reduced aircraft taxi times (see Sec.14.3). These savings are particularly valuable to hub airports serving many banks of transfers. The size of the airports and terminal buildings serving the largest aircraft also motivates the use of airside APMs. The wingspan of the A380 means that the distance between gates approaches 90 m (300 ft). Moving sidewalks can deal adequately with distance of up to perhaps 500 m (1500 ft). Beyond that distance, APMs perform better in terms of capacity and service levels. So, for airports needing to move people great distance, people movers provide the attractive solution.

The design of airside APMs differs from that of landside APMs:

- *Speeds*: are slower [30–40 km/h (20–25 mi/h)] because distances are usually much shorter and, as more riders are standing, braking from a slower speed improves the comfort of the ride.
- *Seating*: usually minimal given the short trips and frequent need for great capacity.
- *Baggage*: limited to carry-ons, which increases vehicle capacity.
- *Vehicle capacity*: generally higher for APM vehicle used for airside than the same vehicle used for landside, due to fewer seats and only carry-on baggage.
- *System capacity*: requirements tend to be significantly higher due to the concentration of passengers particularly the peaking characteristics caused by airline hubs and large aircraft.

- *Frequency*: airside systems usually have shorter headways between trains (2–3 minutes) both to increase system capacity and reduce wait times, which are important to passengers hurrying to catch a flight.
- *Routing*: usually simpler, consisting of a shuttle or pinched loop directly between the terminal and concourses. Exceptionally, the Dallas/Ft. Worth APM has a double-loop routing with trains operating in opposite directions to provide the shortest trip from any station.
- *Technologies*: can include both self-propelled and, for shuttles, cable-propelled vehicles.
- *Passenger separation*: APMs at some international airports need to maintain separation between arriving (nonsecure) and departing (secure) passengers. Rather than building parallel routes at considerable expense, airports can have the trains serve one group in one direction and, after a quick screen at or after the last station on the route, serve the other group on the return trip. The Jeddah and Hong Kong/Chek Lap Kok airports use this approach.

17.8 Airport APM Planning Issues

This section has examples of airport APM planning that seemed appropriate initially but ultimately resulted in problems due to policy changes and changes in airline operations. The lesson is that a range of factors affect airport operations, and these can greatly impact the capacity and performance of APM systems. Flexible designs that can be adapted to future scenarios are best (see de Neufville and Scholtes, 2011).

Capacity and Function

Airport APM train capacity is a function of space per passenger and baggage per passenger. For urban transit, a typical space per passenger guideline is four passengers per square meter (one passenger per 2.7 square feet) of standing area, plus foot space for seated passengers. Airport APM users, however, have baggage: carry-ons for airside and all baggage for most landside systems. Thus for airside APMs the guideline is closer to three passengers per square meter (one passenger per four square feet) and for landside APMs two passengers per square meter (one passenger per five square feet) plus baggage room for seated passengers. The addition of baggage carts increases the space requirements significantly: essentially doubling the space per standing passenger.

At New York/Newark the initial planning decision was not to allow baggage carts on the trains, even though many passengers were going between the terminal buildings and remote parking or rental car facilities. Subsequently, a different airport management team decided that because many passengers had a number of large bags, baggage carts would be provided at stations and allowed on the trains to improve perceived passenger levels of service. This had several unintended consequences. Train, thus system, capacity decreased. The small cabins meant that only one cart could be readily accommodated in each cabin, and then it could hit other passengers and their bags. The narrow doors made cart entry and exit difficult, particularly if bags on the cart overlapped the sides of the cart. This has caused train departure delays from "door holds" and can damage the vehicle and station platform doors.

When designed and built, the primary function of the Newark SkyTrain was to serve passengers traveling among the three passenger buildings and remote parking and rental car facilities. Each building then served one airline (Continental in Terminal C) or set of airlines, so for the most part the only interterminal passenger transfers were between Terminal B international arrivals and Terminals A and C. With airline deregulation and the trend toward hubbing, Continental (now merged with United) ultimately had operations in all three terminals with online transfers that required passengers to leave the secure areas of the arrival terminal, ride the SkyTrain, and then go back through security in the departure terminal. Passengers were not happy about having to go through security again, and minimum online connection times were higher for interterminal transfers. Despite the guideway being on the airside of the terminal buildings, the stations and train operations were designed only for landside service and several attempts to develop terminal design changes to serve air-side passengers as well as landside did not develop an acceptable solution. The non-APM solution was to operate apron buses that connect the airside areas of the three terminals so that passengers do not leave the secure areas.

Changes in airport function, also because of hubbing, affected the APMs at Dallas/Ft. Worth. The initial unit terminals each had separate airlines with origin-destination operations. The original landside Airtrans system primarily served passengers and employees going to and from remote parking and rental car facilities. Relatively few passengers transferred to another terminal. As Braniff and Texas International ceased operations and American Airlines greatly expanded its operations, American Airlines occupied two, then three, then four terminals. The Airtrans system and stations in the terminals occupied by American were reconfigured so that some new routes served intra-American transferring passengers (staying on the secure airside) and some old or reconfigured routes served interterminal transferring passengers on the landside. Remote employee and passenger routes were replaced by bus service. Ultimately the capacity and trip times on the small vehicle, unidirectional Airtrans system was no longer providing the service required for airport, and particularly American Airlines, operations.

Because of the already massive investment in the existing terminals and roadways serving them, the Dallas/Ft. Worth airport decided not to change the entire airport configuration but instead to build a new, high-capacity, bidirectional APM, Skylink, that served the airsides of the existing five terminals and a future sixth terminal. Although this was very expensive, it was much less costly than changing the rest of the airport infrastructure to serve the changed airline and airport operating approach.

Network and Route Capacity

The practical capacity of simple shuttle systems, carrying passengers back and forth between only two points as at Tampa, is easy to estimate. The maximum capacity is simply the capacity of each vehicle or train times the number of departures per hour. As for any system that serves varying loads, the practical capacity over a sustained period is much less than this maximum (see Chap. 20).

Further, the capacity requirement is relatively sensitive to the numbers and sizes of the gates at the satellite or midfield concourse, and to the peaking of aircraft arrivals and departures. Changes in any of these are possible—indeed almost certain. This can necessitate changes to the APM: extending the stations and increasing the number of vehicles in the trains. This is not always possible, so planning for such an expansion is prudent. Sacramento provides a good example of flexible design for the APM: the stations and system elements are designed for two-car trains on each shuttle guideway, but initially the system is operated with one-car trains based on ridership estimates made during the planning phase of the terminal and APM.

Beyond the simple case of a shuttle, the effective capacity of an APM system can be difficult to determine. It is not only its ability to carry a number of passengers per hour. The system must also deliver reasonable service, with tolerable delays, to all points on the network. These requirements mean that the system should operate with substantial excess capacity so that users along the line will be able to board without excessive delays created by insufficient space on the trains due to upstream loads. This subtle issue is easy to overlook.

The essential problem is the likelihood of unequal service that can make the system a functional failure for many users. This difficulty can occur on any network with many stops, which is the typical arrangement for several of the larger airside APMs such as at Atlanta and Denver. In multistop systems, the users at the beginning of the line have the opportunity to take all the places. When they do this, users farther down the line will not be able to board the train. They will then have to wait for subsequent trains, which could be a long time until all the users at the start of the line have been served. This situation means that some users—those at the beginning of the system—may get good service with minimum delays. Meanwhile other users—those trying to use the system down the line—may face intolerable delays (see Example 17.4). Considering the overall average delay to all users, the service might not appear to be bad. In detail, however, an important group of passengers could see that the system was a failure.

This is exactly the situation at Denver where the APM connects the landside terminal with three airside concourses in a pinched-loop configuration. Initially, United Airlines, the dominant carrier at Denver, was primarily operating in Concourse B, with other airlines operating out of Concourses A and C. Trains arriving from Concourse C were not very full, so United's arriving passengers only rarely faced full trains. Subsequently, Southwest Airlines began significant operations in Concourse C and at certain times of the day, trains traveling from Concourse C to Concourse B to Concourse A to the terminal were full when they arrived at Concourse B. Arriving passengers at Concourse A had an alternative path: a bridge over the taxiway between the terminal and the concourse. Arriving passengers at Concourse B did not have an alternative, and many waited for several trains before they could board. During this wait, the station platform could fill with passengers, exacerbating the capacity and level of service problem and potentially creating a secondary problem with passengers using the escalators from the concourse into an increasingly crowded station. Initial APM planning included stations that could accommodate an "extra" vehicle in that trains, but the problem being discussed occurred with the full-length trains.

Denver's potential solution of a shorter "turnback" route that would not serve Concourse C could be possible as the tunnel guideway alignment between Concourses B and C has switches that allow it. However, outbound trains with departing passengers would have to have passenger information that kept Concourse C departing passengers off this short route, and well-known passenger behavior clearly shows that many passengers would ignore any such warnings and end up getting annoyed when they arrive at the wrong concourse. Because of the underconcourse existing station configuration, lengthening them to make longer trains has been deemed infeasible, particularly as the system must remain operational during any construction (and the volumes are much greater than a bus service could accommodate). Other potential solutions include having every other train operate empty between the Concourse C departure platform and the Concourse B arrivals platform, skipping the Concourse C arrivals platform, or reserving some cars in each train for Concourse B passengers by not opening the doors of those cars at the Concourse C arrivals platform. In either case, however, the total system capacity during coincident Southwest and United arrivals peak periods would be insufficient; the solutions merely help balance the lower level of service to both user groups.

If a system with many stations is operating near capacity, it is impractical to overcome this unequal service for passenger service. It is possible to imagine complicated operations that prevent passengers from boarding at the beginning of the service or involve trains skipping stops, as discussed for Denver. These are not always reasonable when dealing with people, however, and would require significant passenger information at a minimum. (With baggage systems, such arrangements might be possible because bags do not complain and resist delays. However, see the discussion in Sec. 17.9.)

Example 17.4 An APM serves a central passenger building (T) and three midfield concourses (A, B, and C), an arrangement similar those to at Atlanta or Denver. An APM with a capacity of 120 persons leaves C destined for T every 5 minutes. Suppose that demand surges over 10 minutes: 100 travelers arrive at A, B, and C each 5 minutes, wanting to go to T. To simplify calculations, imagine there is then no further traffic after this peak.

Table 17.10 traces how the APM carries this peak and shows the delays at the stations down the line. Passengers at the beginning of the run (at C) see an empty train and get immediate service. They leave a few spaces open for the next stop (B), so most of those passengers have to wait. Passengers at the third stop (A) face full trains until all passengers at the earlier stations have obtained service. The result is that passengers at C get immediate service; those at B have to wait for about two trains; and passengers at A have to wait for three to four trains and will perceive that the APM is totally inadequate. This demonstrates the potential problem caused on a network when certain users can seize priority. The distribution of delays is unequal, and it can make a system an operational failure for many users.

Minutes from Start	Persons in State	Midfield Concourse		
		C	B	A
5	Waiting	100	100	100
	Served	100	20	0
10	Waiting	100	180	200
	Served	100	20	0
15	Waiting	0	160	200
	Served	0	120	0
20	Waiting	0	40	200
	Served	0	40	80
25	Waiting	0	0	120
	Served	0	0	120

TABLE 17.10 Unequal Distribution of Delays Occurs When Some Users Can Seize Priority

The solution to providing adequate service to all users is to provide sufficient excess capacity. This is not always acceptable during the planning stage due to the financial constraints on the project. When excess capacity is available, users at the beginning of the line will not crowd out prospective users farther down the line. The right level of excess capacity is not easy to define, however. It depends strongly on the distribution of passengers along the line and their destinations. In Denver, planning efforts did reserve a third more system capacity, but did not envision the actual operating scenario that occurred at the airport. The right amount of excess capacity needed for an airport APM can change as airlines alter their schedules or move operations from one midfield concourse to another, as happens regularly and certainly has happened in Denver. Good design requires careful, expert examination of this issue, most likely including simulations of many different scenarios of terminal design and airline operations and the resultant APM requirements. Project, financial, airline, political, and other constraints might well limit the flexibility needed for an APM to be able to meet all conceivable contingencies. In any event, the practical capacity of a people-mover system with multiple stations can be as low as half its maximum capacity.

The capacity needed to ensure good service might not have to be provided at the start of the operations. Airport operators can size the stations to maximum capacity to give themselves the flexibility to expand the number and length of the APM trains when the need arises. In this case, paying for large stations in advance of need would give the airport operator the flexibility to add people-mover capacity as needed. This flexible approach to design provides assurance in a tangible form.

17.9 Within-Airport Distribution of Checked Bags

Baggage handling systems can be very large. Major systems involve miles of conveyor belts or thousands of baggage vehicles. The one for a Toronto/Pearson passenger building featured over 11 km (7 miles) of conveyors (Table 17.11). That of just for the international passenger building at San Francisco/International has 12,000 m (40,000 ft) of conveyors. Modern baggage systems are major projects by themselves.

Feature	Extent
Conveyors	11,600 m (7.25 mi)
Check-in counter belts	168
Sorting pushers	177
Conveyor levels	5
Screening system	All bags

Source: Jane's Airport Review.

Table 17.11 Design Characteristics of Baggage System for Toronto/Pearson Terminal 1

Furthermore, baggage handling can involve enormously complex networks. They merge, process, and separate varying streams of bags. They may assemble bags from over 100 check-in points, merge them, and then sort them out to at least as many destinations. Along the way, they will pass the bags through several levels of security screening. The layout of the system might remind one of noodles on a plate (Fig. 17.11). Some baggage systems can be simple, of course. At smaller airports, straightforward arrangements of carts and containers deal with the makeup of bags for flights and their subsequent distribution to baggage claim devices, as described in the FAA design guidelines (FAA, 1988). For significant airports, however, the baggage systems are definitely complicated. Designers need to pay special attention to these networks.

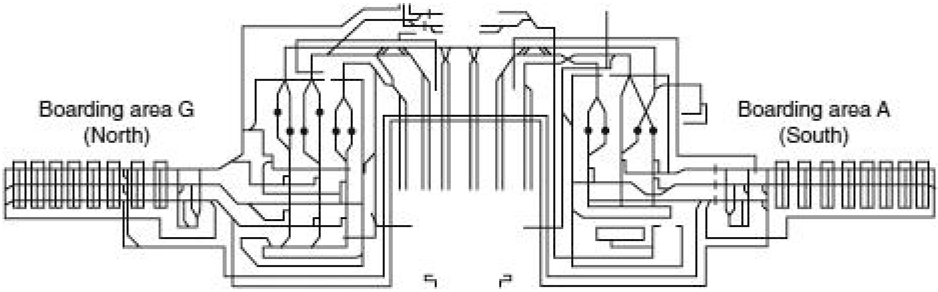


FIGURE 17.11 Outbound baggage handling system, International Terminal, San Francisco/International. (*Source*: Transportation Security Administration, 2011.)

Modern bag systems consist of three elements:

1. Security inspection of the checked bags. Governmental agencies determine the requirements for this process. In the United States, the Transportation Security Administration (TSA) defines the standards (TSA, 2011).
2. IT-based control systems that keep track of bags and that control their movement through the baggage system.
3. Mechanical systems that convey bags through the system, most obviously from check-in to the aircraft, but also "tail-to-tail" between connecting flights, and from the aircraft to baggage claim devices.

This section discusses each of these elements in turn and closes with a critical discussion of the concept of the capacity of baggage systems.

Security Inspection

Security systems currently involve at least two distinct processes:

1 *Screening of the bags* by x-ray machines and other devices

2 *Baggage reconciliation*, the process of making sure that each bag placed on an aircraft belongs to someone who is definitely on board

Governments are also considering and may introduce additional processes, for example some form of profiling for bags.

Baggage reconciliation is a way to discourage people from placing explosives on board an aircraft. As TSA (2011) describes, the process compares the list of passengers who have boarded with the bags checked in for a flight, in the effort to make sure no one avoids getting on a flight with a dangerous bag. This process must be performed on all international flights to and from the United States, the member countries of the European Union, and many other nations. Many countries also require it for their domestic flights. To implement this process efficiently, an airline must be able to identify rapidly, through electronic processing of boarding passes or by other means, all the passengers who have boarded a flight. The airline must also know the owner and location of each of the bags placed on board so that it can effect the baggage reconciliation and remove bags quickly if there is a discrepancy. Some airlines and airports are better prepared to accomplish this task than others, because of the different approaches they use to sort bags and place them on the aircraft. (People may miss flights for all kinds of harmless reasons: They fall asleep in a waiting room, get lost trying to make a transfer, stay too long in a shop or restaurant, etc., see Spake, 1998.) The next subsection describes the issues involved in creating information systems for bags.

Complete, 100 percent screening by x-ray and other devices is the standard as of 2012. Airports in the United States, Canada, and most of Europe have met this goal. The design of this process involves tradeoffs between accuracy, speed, and cost. A complete search of any bag requires some combination of highly sophisticated devices and manual inspection of the contents. Modern "checked baggage inspection systems" (CBIS) are expensive—and also heavy; the example in Fig. 17.12 weighs 8.5 tons. Manual searches are too long and too expensive to apply to every bag. As technology advances and threats evolve, the design of baggage inspection systems also changes. This fact, coupled with increases in the number of passengers and bags, means that baggage systems routinely have to be modified. A good design will provide the flexibility needed to accommodate this reality. Unfortunately, bag systems are often stuffed into insufficient space. For instance, any retrofit to the scanning and processing system for bags shown in Fig. 17.11 is likely to be extremely difficult. The U.S. TSA sets the evolving standards for the United States, as in TSA (2011).

FIGURE 17.12 Medium throughput checked baggage inspection system (CBIS), Model CTX-9400. (*Source*: Transportation Security Administration, 2011.)

The screening process is thus inevitably a triage involving several steps. Details of these operations vary from place to place and change over time. As of 2012, the United States and Canada were moving to a three-tier screening process consisting of

1. Autonomous screening by machine, which clears most of the bags
2. Resolution of possible issues by a person looking at the screened image
3. Manual search of bags by inspectors

IT-Based Control Systems

Most baggage handling systems rely on bar-code labels to identify the bags and their destinations, and thus to control their flow. These labels have the significant advantage of being inexpensive; they cost a few cents each. The baggage system reads these labels and uses the information to direct the bags down appropriate chutes. The basic reading device consists of laser scanners mounted on an apparatus over the baggage conveyors. As the reader may have experienced at a grocery store, lasers do not always read on the first try, so they require substantial redundancy. A typical arrangement for bags has several laser readers mounted around the path of the bags, and each laser attempts to read several times.

Overall, laser scanners work generally well—provided the tag is not hidden underneath the bag, dirty or torn, and that the lasers are cleaned regularly. As a practical matter, this means that they work best when bags are first checked in, and less well at transfer points, after mechanical devices and baggage handlers have pushed the bags around and thrown them in and out of containers. In a major baggage system that is performing well, about 3 percent of the bags may not be read correctly on their first circuit through the system. This error rate may be much greater after the bags have been handled many times, specifically at transfer points. Typical design specifications require 99 percent correct reads in testing, and 95 percent during operation. At one point British Airways reported that 10 percent of its bags at London/Heathrow could not be read automatically and had to be read by hand. This problem means that baggage systems may need substantial systems for sorting misread labels. In any case, the automated systems need substantial redundancy to minimize the probability of misreads. Despite these issues, bar-code systems have been more cost-effective than the alternative.

The other means to identify items is through *radio-frequency identification* (RFID). This technology uses radio antennas to read individual chips attached to bags or trays. This technology has not yet proven to be cost-effective for general airport use. A number of technologic difficulties need to be overcome: Obstacles may block or distort the communications, or the chip may detach from its item. Cost is also a major problem with RFID systems. The reading equipment may be less expensive than that for a bar-code system, because it consists of antennas rather than an array of lasers. As of 2012, however, an RFID chip costs around 20 cents, about five times more than a bar code label. If this cost decreases substantially, RFID systems may replace bar-code systems. Otherwise, they will be limited to special applications. For example, Vancouver and Barcelona airports are embedding RFID into their destination-coded vehicles (DCVs), which are containers that carry bags through a baggage system.

Management and design decisions can “encode” information and thus significantly enhance the performance of baggage handling systems. It is possible to sort prospective bags into containers destined for their ultimate destination that capture this information physically. They make it possible to transfer these bins of bags directly from one aircraft to another, thus bypassing the sorting system at the transfer hub. In a similar vein, designs that require passengers to bring their bags directly to the gate for check-in eliminate the need for a local sorting of outgoing bags. This feature determined the configuration of the first Aérogare 2 passenger buildings built for Air France at Paris/de Gaulle. Airport managers should always be on the lookout for similar arrangements that might make it easier to solve their baggage handling problems.

Mechanical Systems

Baggage systems need to separate bags from each other by about 1 m (3–4 ft), so that the sorting mechanism can distribute them correctly. The bags cannot be right next to each other, as they typically are at a bag claim device. Most commonly, the bag system simply spaces the bags out along conveyor belts. This arrangement allows mechanical diverters to divert a bag to its destination and retract before the next bag comes down the line (Fig. 17.13). Alternatively, the system consists of individual conveyances. These can be tilt-trays linked together that tip the bag into a chute or DCVs that operate independently (Fig. 17.14). Designers need to consult experts in baggage handling for current details of this rapidly progressing field.

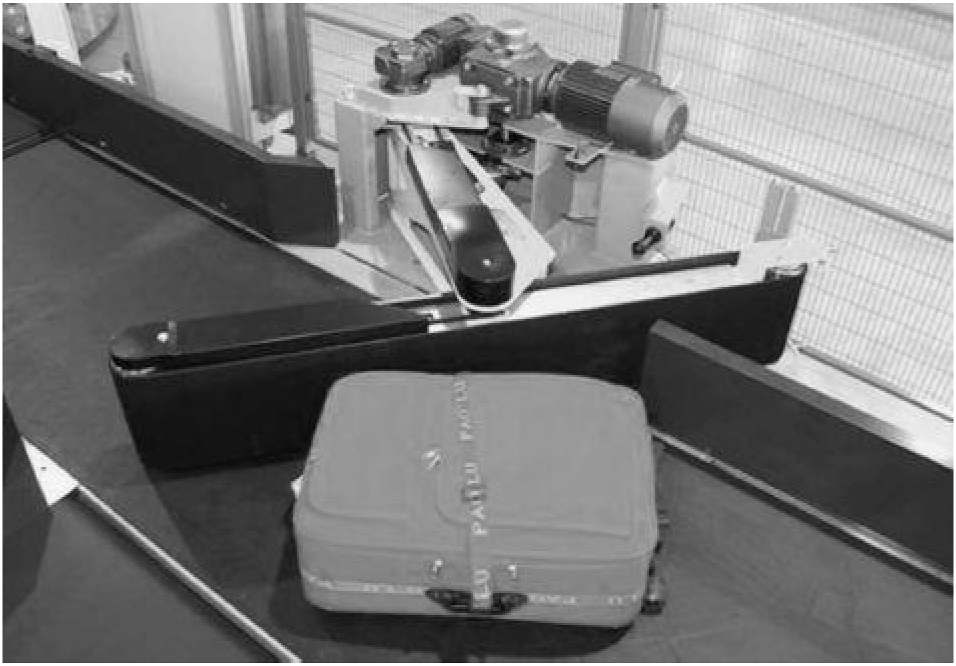


FIGURE 17.13 Tilt-tray baggage handling device. (*Source*: VanDerLande Industries.)

FIGURE 17.14 Destination-coded vehicles (DCVs). (*Source*: VanDerLande Industries.)

The requirement for individual spaces between the each bag has two important implications for the design and operation of an automated bag system. Most obviously, its mechanical devices require a lot of space. The separation standard limits the capacity of each baggage line and increases the total size of the system. More important, however, the need for individual spaces means that the baggage system is a highly complex system of queues. Note that the separation requirements apply to bags that must be sorted and screened; general in-bound bags destined for arriving customers can be placed very close to each other, as passengers see on bag reclaim devices.

The need to place bags in physical, spatial slots slows processing and can lead to operationally disastrous distributions of delay. It slows the check-in process because agents cannot feed another bag into the system until there is an empty space, for which they often have to wait. This kind of delay occurs every time baggage lines merge—when bags from check-in agents merge with the line behind a bank of check-in counters, when these lines merge with those of other banks, etc. Furthermore, unless the system has substantial excess capacity, certain users will have to wait excessively because others use up all the available spaces (see Example 17.5). Theoretically, the operation of the system might be able to allocate the spaces throughout the system very carefully and avoid excessive delays. In practice, this has not been possible because the pattern of flows changes unexpectedly as flights are delayed. This phenomenon was a principal reason the DCV system at Denver/International could not deliver bags within the time constraints associated with transfer operations (de Neufville, 1994).

Sorting bags at transfer hubs is a very challenging task. Bags must be taken off arriving aircraft, sorted, loaded onto potentially several connecting flights—all possibly within 30 to 60 minutes. To do this for all bags individually is not practical. The number of bags to be sorted within a limited time would require enormous capacity. It is also difficult to avoid excessive delays for many connections. Airlines solve this problem by presorting bags at the airports delivering passengers to the transfer hubs. For example, an airline might sort bags in Boston into containers destined, via connections at Miami/International, for Bogota, Santiago, Panama, etc. In Miami, the airline can then ship each of these containers directly from the Boston flight to the appropriate connection, in a "tail-to-tail" operation that bypasses the sorting system at Miami. [See Robusté and DaGanzo (1992) for an analysis of this situation.]

If connections are long, as they often may be at London/Heathrow and elsewhere, it may be necessary to store bags for many hours. Thus Amsterdam/Schiphol has a "cold bag" storage area consisting of lines of multilevel racks. This huge complex system is served robotically—and is obviously expensive (Fig. 17.15). At most airports, wages for labor to move and sort bags are not high enough to justify such automation.

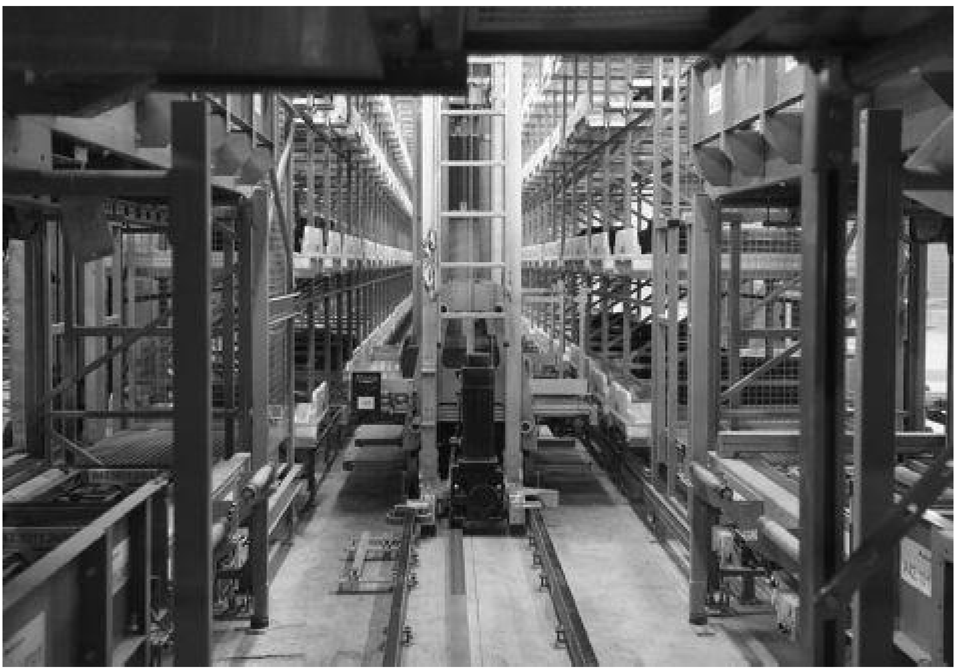


FIGURE 17.15 Early bag storage system in Amsterdam. (*Source*: VanDerLande Industries.)

Delivery of bags to passengers is, by contrast, simple. Normally, bags from any flight go directly to conveyor belts that feed claim devices. Passengers identify and pick up their own bags, as shown in Fig. 17.16.

FIGURE 17.16 A typical baggage claim device. (*Source*: VanDerLande Industries.)

Capacity

The practical capacity of a baggage handling system is much lower than the theoretical capacity. This is true even for the simplest arrangements. FAA guidelines emphatically stress that the practical capacity is one-third less than the theoretical amount obtained by multiplying the speed of the conveyor system times the density of bags per unit of length on the system (FAA, 1988). This is due to the inherent variability in the arrival of bags onto the system, due both to the waves of departures throughout the day and the random way in which agents process bags upon check-in.

For complex automated systems, the practical capacity is even less due to the need to wait for slots for each bag, and to provide reliable service to and from all points in the system. As for passengers on people-movers, bags need to wait for "slots" on the baggage handling system before they proceed through the sorting system. They can cause significant delays to some elements unless the system has ample extra capacity (see Example 17.5). Furthermore, when dealing with transfers, the reliability of the bag transfer system, in terms of delivering all connecting bags within a tight connection time, is paramount. In this connection, designers have to recognize the fact that the variability of queuing systems generally increases with congestion (see Chap. 20). This means that the need for reliable performance often limits the effective capacity of a baggage handling system. At Denver/International, for instance, the need for reliability apparently limited the automated baggage system to only about 40 percent of the notional capacity of the system.

Planners should note that mechanical baggage handling systems consume an enormous amount of energy. It takes a lot of effort to run the hundreds of machines dispersed over miles and kilometers of these systems. While detailed data are not available, industry professionals have estimated that bag systems may account for 30 percent of the energy used in airport facilities. Mechanical bag systems would seem to offer opportunities for new systems with low energy consumption.

17.10 Take-aways

One major point of this chapter is that special-purpose, high-speed rail projects for accessing airports are not cost-effective. This is because the airport traffic comes and goes from all across the region the airport serves. Cars, taxis, and buses will continue to dominate the airport access market. As a consequence, parking facilities are and will be major landside projects at major airports. Rail access to airports works well in those cities where it is part of a widely used, highly distributed network.

APMs—rail-like in that they run on their own guideways—have been the major innovation in the design of airport passenger buildings in the last generation, and consequently in the overall configuration of the airport. They make it possible to create gate positions at which the aircraft have plenty of room to maneuver freely and efficiently, while reducing the distances passengers have to walk. This technology has now become well-understood and highly reliable.

In parallel, baggage handling procedures and technology continue to evolve. Compared to what was standard until recently, these systems have become highly technological, complex, and expensive. Recent experience with new requirements underlines the need to plan the baggage handling spaces generously and flexibly.

Indeed, the bottom line is that the design of access systems can expect to face considerable uncertainty as airlines adopt low-cost practices, merge, and change their operational practices, as regulations change, and as the number of passengers increases. This reality should drive us to plan to develop these systems flexibly, with the ability to adapt them to new circumstances over their useful life.

Exercises

17.1. Investigate the traffic at a local airport or one in which you have contacts. As best you can, estimate the daily flow of passengers, employees, and commercial vehicles. The airport should have data on either these flows or their contributing factors. Compare these estimates with the ranges in the text and discuss the factors that account for differences.

17.2. Ask 10 or more persons around town about where they started and ended their last trips to airports. What percentage went to a downtown area? Used public transport? Taxis? Their own or someone else's private vehicle? To what extent do you think this informal survey is representative of passenger traffic?

17.3. Ask 10 or more persons (perhaps those selected for Exercise 2, perhaps at the same time) about the factors that matter most to them in the airport access trip. Cost? Speed? Reliability? Convenience? What do these responses indicate about the kinds of public transport that might be suitable for airport access?

17.4. Estimate the total door-to-door time and cost of your several modes of airport access. Which mode is most attractive to you? Do a similar estimate for your possible modes of travel from your last destination airport. What accounts for any differences in your choice? What factors influence these choices?

17.5. Go to some facility that has short-term parking—at an airport if possible, otherwise at some store such as a supermarket. Observe a set of about 20 spaces close to the main entrance of the passenger building or store for half an hour. How many cars parked in these spaces in that time? Estimate how long people stay on average and the capacity of these spaces, that is, the number of cars that might park in these spaces in a day.

17.6. Explore the distribution of delays on a network similar to that in Example 17.5. You can do this conveniently using a spreadsheet model organized similarly to Table 17.10. Look at short and longer surges of traffic, different patterns of arrivals and destinations, for various levels of line-haul capacity. Calculate how limits on delays at any station set the practical capacity of the network below its theoretical maximum.

References

Airport Cooperative Research Program (2009) *Guidebook for Evaluating Airport Parking Strategies and Supporting Technologies*, ACRP Report 24, Transportation Research Board, Washington, DC. www.trb.org.

Airport Cooperative Research Program (2010) *Airport Curbside and Terminal Area Roadway Operations*, ACRP Report 40, Transportation Research Board, Washington, DC. www.trb.org.

Airport Cooperative Research Program (2013) *Integrating Aviation and Passenger Rail Planning*, ACRP Report, Transportation Research Board, Washington, DC. www.trb.org.

de Neufville, R. (1994) "The Baggage System at Denver: Prospects and Lessons," *Journal of Air Transport Management, 1*(4), Dec., pp. 229–236.

de Neufville, R. (2006) "Planning Airport Access in Era of Low-Cost Airlines," *Journal of American Planning Association, 72*(3), pp. 337–346.

de Neufville, R. and Scholtes, S. (2011) *Flexibility in Engineering Design*, MIT Press, Cambridge, MA.

FAA, Federal Aviation Administration (1988) *Planning and Design Guidelines for Airport Terminal Facilities*, Advisory Circular 150/5360-13, U.S. Government Printing Office, Washington, DC.

Leigh Fisher Associates, Coogan, M., and MarketSense (2000) *Improving Public Transportation Access to Large Airports*, Transit Cooperative Research Program Report 62, Transportation Research Board, National Academy Press, Washington, DC.

Robusté, F. and DaGanzo, C. (1992) "Analysis of Baggage Sorting Schemes for Containerized Aircraft," *Transportation Research A, 26A*(1), pp. 75–92.

Shapiro, P. S., Katzman, M.G., Hughes, W.E., McGee, J., Coogan, M., Hoffman, M. A., Wagner, E. V., and Mandle, P. (1996) *Intermodal Ground Access to Airports: A Planning Guide*, U.S. Department of Transportation, Federal Aviation Administration, December 1996.

Spake, J. (1998) *Jeremy's Airport*, BBC Worldwide, London, U.K.

Transit Cooperative Research Program (2000) *Improving Public Transportation Access to Large Airports*, TCRP Report 62, Transportation Research Board, Washington, DC. www.trb.org.

Transit Cooperative Research Program (2002) *Strategies for Improving Public Transportation Access to Large Airports*, TCRP Report 83, Transportation Research Board, Washington, DC. www.trb.org.

U.S., Transportation Security Administration (2011) *Planning Guidelines and Design Standards for Checked Baggage Inspection Systems*, Version 4.1. Washington, DC.

[1]In a pinched loop layout, trains turn back at the end stations and use switches to cross to the other guideway of a pair.

PART V

Reference Material

CHAPTER 18

Data Validation

Data are often not what they seem to be. Any single series may contain different definitions, errors, and limitations that distort its meaning. Different series, from different contexts and sources, are unlikely to be fully comparable.

Analysts using historical data need to validate this information. To make forecasts and plan facilities correctly, they need consistent data that they understand. Consequently, they need to go through a process of examining their data to understand its definitions, to ensure that these definitions are consistent and to eliminate errors.

18.1 The Issue

The data available for airport planning and design may not be what they seem to be. Because of particular local definitions or needs, the data may not mean what the analyst thinks they imply. They may mean different things at different times, as definitions or methods of collection change what officials measure and report.

Errors

The data may simply contain significant errors. A few examples illustrate some of the possibilities.

Different, Changing Definitions

The government of Mexico employed a team to review long-term forecasts of passengers for Mexico City prepared for the World Bank. As part of its careful examination, the team found that the official data implied that the number of passengers per flight in domestic traffic had dropped significantly over a 6-year period. This was surprising, because airlines normally adjust their schedules to maintain consistent loads. What accounted for this pattern? Why did it persist for 6 years?

Discussions with the airlines indicated that from their perspective nothing special had happened during that period. Aviation officials then remembered that the Mexican government had changed the definition of international passengers during a specific presidential administration (a term of 6 years in Mexico). Originally, passengers leaving Mexico City for a final destination in the United States were counted as “international” from the moment they boarded the aircraft, even if it landed at a midway point such as Guadalajara. During this one administration, however, they were considered “domestic” passengers when they left Mexico City for the intermediate destination inside Mexico. This second definition is perfectly reasonable. However, the inconsistency of the definition over time led to confusing statistics. (It also provoked some imaginative explanations of the factors influencing traffic from the economists working for the World Bank.) The review team fixed the problem by adopting a consistent definition of the data. (For details, see de Neufville et al., 1980.)

Distorting Methods of Collection

In designing the new passenger building for Mombasa, Kenya, it was important to know the number of international passengers, because these would determine the size of the international departure lounges, the number of passport control booths, etc. The Kenyan government supplied data derived from their in-house aviation experts, the nationally owned Kenya Airways. After working in Kenya for a while, designers found that the actual international traffic was much less than shown in the official figures. What was the story?

It turned out that Kenya Airways reported data from its International Division, that is, from the crews and aircraft normally engaged in international flights. However, many of these aircraft flew regularly on domestic routes. An aircraft arriving in Nairobi from England early in the morning would make one or two round trips from Nairobi to Mombasa during the day, before it returned to England late at night. The traffic on these domestic flights was quite properly ascribed to the work of the International Division. Somewhere in the statistical reporting process, this distinction between the International Division and actual international passengers was lost. The designers worked with Kenya Airways to resolve this matter and, with the assistance of data from the passport control officials, derived a correct estimate of the international passengers.

Simple Errors

A master planning process for Los Angeles/Burbank airport used official forecasts from the U.S. Federal Aviation Administration (FAA). A careful reviewer of this plan looked at previous data and noticed a 25 percent jump in traffic in a single year—not only extraordinary in itself, but also very different from the normal low steady growth. What was going on here?

The reviewer went back to the airport officials and the original airline data they used to collect the annual statistics on enplanements. It turned out that, somewhere between California and the FAA headquarters in Washington, there had been a simple transposition of two similar numbers, a 3 switched for an 8 in an important column. The actual 2.3 million passengers had unintentionally grown to 2.8 million.

Incompleteness
In addition to statistical inconsistencies, some data series are inherently incomplete. A couple of examples illustrate this problem.

Systematic Undercounting
One example of incomplete data concerns the premier source of data for the international air transport industry, the World Air Transport Statistics published by the International Air Transport Association (IATA, annual). This publication presents information from airlines and air carriers that have chosen to pay to be members of the IATA. It is thus systematically incomplete, since many airlines, including some of the largest, have not been members of IATA. Furthermore, the membership normally changes from year to year, so that the series is not consistent over time.

Methodological Blindness
The statistics on air cargo for Canada presented another example of incompleteness. In that case, the responsible government agency, Transport Canada, developed data on air cargo from waybills on individual shipments. This procedure fails to capture data on all-cargo aircraft carrying freight for a single shipper, such as UPS or FedEx. Thus, the official statistics have shown virtually no cargo going through Toronto/Hamilton airport, when it was actually one of the Canadian hubs for integrated cargo carriers.

18.2 The Resolution

As the examples suggest, airport planners need to validate carefully the data they will use for design. However, no single procedure for checking statistics will resolve all the issues. The available statistics may be misleading for all kinds of reasons. Analysts should be looking constantly for potential problems. They should adopt two basic approaches to validating their data. They should

- Understand what the data mean
- Double-check the data using alternative sources

These guidelines are in fact the basis of all good investigations. Airport planners should first be clear about what the data mean. They should understand both the formal and practical definition of the statistics. Although the formal definition may seem obvious, it often is not. For instance, professionals might assume that there could be no confusion about what it means to be an "international passenger." However, this term represents different concepts in various situations. Thus, in the European Community many "international passengers” are indistinguishable from what planners elsewhere consider domestic passengers, in that they require no customs clearance or passport checks (this applies to travelers between the countries party to the Schengen agreement).[1]

Conversely, passengers arriving in the United States from Canada are generally not treated as international passengers, since they formally entered the country at special facilities in the Canadian airports. Planners often make significant errors when they assume they know what the categories of data mean. They should routinely check the local definitions.

Analysts should also check the practical definition of the data. How statisticians collect data gives this information an operational definition that can significantly modify the formal definition. The examples concerning the practical definition of international passengers in Kenya and Mexico illustrate how local practices give different meanings to equivalent formal concepts. Planners should find out how the local authorities collect data.

Analysts should also double-check their data to the extent possible. They need to obtain independent confirmation of the accuracy and completeness of the data. Any data series may contain biases and errors for many reasons, such as those cited in the introductory examples. It is easiest to corroborate data with alternative series that supposedly mean the same thing. Thus, in the examples of Los Angeles/Burbank and Mexico City, the analysts were able to identify airline statistics that should have agreed with the original information. The discrepancies between the parallel series of data helped identify what was really happening. Analysts should seek out complementary data series to verify their sources of data.

When independent data sources are not available, analysts may have to find more creative ways to validate their data. They can look at operational patterns, for instance, to see if the behavior implied by the official statistics agrees with what would be good practice. The original clue that the data on international passengers was inconsistent in the Mexico City example was that the implied ratios of passengers per flight were abnormal. Sometimes, the analysts will have to go beyond statistics and rely on interviews with local operators and observations in the field. Canadian planners may not find formal alternatives to the Transport Canada data on air cargo through Toronto/Hamilton because the shippers may not wish to release these figures. They can, however, verify that integrated cargo carriers have frequent flights to this airport and thus determine that the shipments are important. They can also estimate the tonnage shipped, knowing the number and size of the cargo aircraft. In short, airport analysts should also think “out of the box” when validating their data.

In summary, airport planners and designers should validate their data in four ways. They should

- Check the local meanings of the data, to establish the formal definitions
- Find out how the local authorities collect data, to determine the practical definitions

- Seek out complementary data series to verify their sources
- Think “out of the box” when validating their data

Exercises

18.1. Consider any series of airport data (e.g., passengers, aircraft operations, or cargo). Determine its definition. Then attempt to find out how the relevant authorities collect the data. Ask yourself what elements the data might leave out. Do the authorities count crew members flying on passes as passengers, for example? How do they count military or training operations? And so on.

18.2. For the same series or some other, think about how you could double-check these data in practice. What alternative independent sources could confirm their accuracy? What field observations or interviews could help you validate the formal series?

References

de Neufville, R., Zuniga, S., Kanafani, A., and Olivera, A. (1980) “Forecasting Airport Traffic: Mexico City as a Case Study,” *Transportation Research Record*, 732, pp. 24–29.

IATA, International Air Transport Statistics (annual) *World Air Transport Statistics*, IATA, Montreal, Canada.

[1]Some members of the European Union have agreed, originally at a meeting in the city of Schengen, that passengers between them do not have to go through passport control and customs. France and Germany have participated from the start, but the United Kingdom did not.

CHAPTER 19

Forecasting

Forecasting is an art, not a science. Any forecast of phenomena involving people is inherently unreliable and likely to be wrong. All forecasts of social activities involve many arbitrary assumptions and opinions. Users of forecasts need to understand and appreciate this reality.

The basic concept of forecasting is to estimate past trends and project them forward. The idea is simple in concept, questionable in execution. To estimate past trends, analysts have to make many assumptions. They have to select a span of data (over time or different circumstances), some principal drivers of the phenomenon being forecast, the form of these factors and of the mathematical model, and future values of their drivers. Different analysts choose these factors differently and obtain quite different results.

The mathematics involved in forecasting can look frightening, but all procedures call upon a simple idea. They attempt to create a formula that correlates well with past experience. Analysts can generally do this easily, whatever assumptions they make. They do this by adjusting the parameters of the formulas, the form of these formulas, and the variables included. Mathematics is not the decisive factor.

Ultimately, all forecasts about airport activities reflect judgment and opinion. In the short run, existing rates of change are likely to persist due to inertia. In the longer run, airport forecasts are based on the fallible opinions of analysts.

This chapter first focuses on the selections analysts have to make to create a formula for forecasting, using data from forecasts prepared for a master plan of Los Angeles Airports. It thereby underlines the important role of judgment in the creation of forecasts. It next proceeds to describe the fundamental mathematics of forecasting, principally regression analysis. This discussion shows how analysts can easily develop many models that fit past trends well, demonstrating that mathematics alone cannot validate a forecast. The discussion then indicates how experts can use their knowledge directly to develop scenarios of future developments. Finally, this chapter presents an overall approach to airport forecasting that blends analysis and expert judgment, giving more weight to trend extrapolation in the short run and judgment in the long run.

19.1 Forecasting Assumptions

Forecasters use the past to describe the future. They create some model of how things happened to project what may occur. They look at some factor in which they are interested (such as the number of passengers) and attempt to describe how some other secondary factors (such as airline fares and the level of local economic activity) made it change.[1] When they are satisfied with their model, they estimate the future level of the secondary factors and use them to project the future values of the activity in which they are interested. ACRP (2007) covers standard approaches to forecasting for airport aviation activities in the United States.

Alert readers will notice that this process is immediately questionable. Analysts start with a desire to forecast one factor and end up having to forecast several secondary factors in order to do so. If these secondary factors can be known in advance with some assurance, the process may be fine. If they cannot, as is often the case, the process actually makes a hard problem harder—instead of forecasting one factor, the analysts have to forecast several.

When analysts develop forecasts mathematically, they create formulas or models to describe the evolution of the factor in terms of other factors. For example, the consultants on the Los Angeles master plan wished to forecast the number of passengers for Los Angeles. Formally, they identified the variable they wish to forecast, Y, and many other factors, X_i, and tried to develop formulas linking the two. These were of the form

$$Y = f(X_1, X_2 \ldots X_n) \tag{19.1}$$

To describe a model numerically, analysts have to make the following four kinds of assumptions:

1. The *span of the data*, in terms of periods or different situations
2. The *variables* to be included in the formulas
3. The form of these variables
4. The form of the equation itself

Each of these choices affects the calculations of the forecasts, yet analysts have no firm scientific basis for making these choices. The forecasts inherently reflect opinion.

Analysts choose the span of their data. For a Los Angeles master plan, for example, the consultants considered the data in Table 19.1. Their choice of 30 years was reasonable, yet it was not the only possible choice. An alternative analysis might have dropped some of the earlier years. It might do this on the grounds that the eras before the first oil crisis (1973), or before the economic deregulation of the airlines in the United States and the second oil crisis (1978), were so different from the more recent situation that they represented different trends. Either of these alternatives is defensible; either gives different models.

Year	Traffic Type		Total Thousands
	Domestic	International	
1965	12,134	445	12,579
1966	14,691	561	15,251
1967	17,423	702	18,125
1968	19,449	897	20,346
1969	20,112	1,198	21,310
1970	19,388	1,392	20,781
1971	18,809	1,538	20,347
1972	20,196	1,882	22,078
1973	21,336	2,166	23,502
1974	21,241	2,344	23,585
1975	21,229	2,490	23,719
1976	22,997	2,987	25,983
1977	25,070	3,292	28,362
1978	28,746	4,155	32,901
1979	29,926	4,997	34,923
1980	27,386	5,652	33,038
1981	27,281	5,442	32,723
1982	27,647	4,736	32,383
1983	28,517	4,910	33,427
1984	28,978	5,383	34,362
1985	31,759	5,889	37,648
1986	34,968	6,450	41,418
1987	37,408	7,465	44,873
1988	36,340	8,059	44,399
1989	35,824	9,143	44,967
1990	35,969	9,841	45,810
1991	35,284	10,384	45,668
1992	35,509	11,456	46,965
1993	35,900	11,945	47,845
1994	38,371	12,679	51,050

Source: Landrum and Brown, 1996.

Table 19.1 Total Annual Passengers (in thousands) at Los Angeles/International

Aviation analysts typically work with time series, as in Table 19.1, and have to decide on the number of periods. However, they may also work with data on several different situations for a common period, using what is known as a cross-sectional approach. For example, in trying to model the number of passengers, researchers could look at the data for many cities in some year and would have had to choose the number of cities. Using either time series or cross-sectional data, analysts have to choose the span of the data.

Analysts next have to choose the factors they believe are the most important drivers of the factor they wish to forecast. For Los Angeles, the consultants focused on population, employment, per-capita personal income (PCPI) for the region, and national average yield for the airlines. This choice is reasonable but not the only one that they could have made. For a similar study of Miami, for example, other analysts considered the regional share of national traffic. They could also have looked at the national level of income, the economic growth of Latin America, etc. There is no definite set of variables on which all analysts agree.

Analysts also have to choose how they will represent the data. Normally, most traffic forecasts include some measure of cost (such as fares) as one of the factors to include in the analysis. Basic economics has taught us that prices strongly influence demand. Lower prices increase traffic, for example. However, there are many ways to represent this factor. Analysts can use figures on average price (as the consultants did for Los Angeles), on relative price (e.g., compared to road or rail), on price difference compared to competition, etc. For each of the basic concepts that might be involved, different analysts can and do justify different versions. These choices will affect the forecasts.

Finally, the analysts have to choose the form of the model. Most simply, it could be a linear model, such as

$$\text{Passengers} = (\text{population}) \cdot [A_0 + A_1\,(\text{income}) + A_2\,(\text{employment})] \tag{19.2}$$

Alternatively, the model could be a power relationship, which economic theory suggests more accurately represents the relationship between price and demand:

$$\text{Passengers} = B_0\,[\text{yield}]^{B_1} \tag{19.3}$$

It could also be exponential in time, a form that conveniently represents constant rate of growth C_1 over each period of time T.

$$\text{Passengers} = C_0(e)^{C_1 T} \tag{19.4}$$

For Los Angeles, the consultants tried all these three and others. Even with these, they had not exhausted the possibilities. Analysts can create all kinds of complicated models.

The point the reader should retain is that analysts have to choose between the models without any clear indication of which is best. This and the other subjective judgments are at the root of all forecasting analyses and influence the mathematical results. The forecasts have no firm scientific basis. All forecasts rest on opinions.

19.2 Fundamental Mathematics

To obtain a specific model of the behavior of the factor of interest, such as the number of passengers, analysts have to "fit" the model to the historical data. The standard approach is known as *linear regression analysis* (see, e.g., Allison, 1998; ICAO, 2006; Makridakis et al., 1997; Pindyck and Rubinfeld, 2000). Formally, it applies to linear models such as that in Eq. (19.2). In fact, it applies as well to exponential and power relationships [such as Eqs. (19.3) and (19.4)], because these can be made into linear equations once they are expressed in terms of logarithms. Thus

$$\begin{aligned} &\text{Passengers} = C_0\,(e)^{C_1 T} \\ &\rightarrow \log\,(\text{passengers}) = \log\,(C_0) + C_1 T\,(\log e) \end{aligned} \tag{19.5}$$

Linear regression analysis fits the model to the data by minimizing the sum of the squared differences between the actual observations of the data at different times, Y_t, and the values indicated by the model $\mathbf{Y}_t$

$$\text{Difference to be minimized for best fit} = \sum_t (Y_t - Y_t)^2 \tag{19.6}$$

The squared term has a double advantage:

- It ensures that positive and negative differences are taken in absolute terms and do not cancel out.
- It penalizes large deviations in favor of small ones so that the distribution of differences resembles a Normal, a distribution of random errors.

Regression analysis determines the constants of the equation that best fits the actual data. It minimizes the sum of squared differences by differentiating Eq. (19.6), with the model expressed in terms of the X_i replacing the $\mathbf{Y}_t$, with respect to the coefficients of the X_i. Setting these expressions equal to zero and solving gives the values of the coefficients that define the best-fit model.

Forecasters can ignore these details, because spreadsheet programs such as Excel do linear regression analyses automatically upon request. Spreadsheet programs also provide measures of how well the model fits the data. The most usual measure is R^2. Formally, R^2 represents the amount of variance in the Y_t accounted for by the model. $R^2 = 1$ indicates perfect fit. Any R^2 above 0.9 indicates good fit.

In practice, it is easy to get good fit to time-series data. This is because most statistics associated with humans change fairly constantly over time (e.g., population grows or shrinks at a few percent a year, as does employment, income, travel, and so on). These statistics can be expressed reasonably well in a form similar to Eq. (19.4). Therefore, any two time series (say, C_I and D_I) can in general be easily expressed in terms of each other, regardless of their relative rates of growth; for example,

$$C_0(e)^{C_1 T} = \left[\frac{C_0}{D_0}\cdot(e)^{\frac{C_1}{D_1 T}}\right]\cdot D_0 \cdot (e)^{D_1 T}$$

This means that it is easy to obtain good correlation, with high R^2, between factors that have nothing to do with each other. Past student exercises have shown good correlation between airport traffic and such diverse factors as the egg production of New Zealand or the prison population of the state of Texas. Correlation is not causality, however: good fit does not necessarily imply that a model is meaningful.

19.3 Forecasts

Once the analysts have developed a model to fit the past data, they can use it to forecast future levels of traffic. One of the models that the consultants fitted to the Los Angeles data was similar to Eq. (19.2):

$$\text{Domestic passengers} = (\text{population})[-3074.4917 + 0.1951\ (\text{income})]$$

(19.8)

They used this to project possible future levels of traffic for Los Angeles, by inserting forecasts of the future values of the population and per-capita income for the following 20 years.

As indicated at the start of this chapter, the standard forecasting process makes a simple problem more complex. To get the forecast of passengers, the process for Los Angeles had to predict two other variables. This is not easy, especially as the experts in those areas are not agreed. Thus, four reputable planning agencies had four different estimates of future levels of per-capita income. Moreover, they were not even agreed on past levels, as Table 19.2 demonstrates.

	LA Region PCPI in Constant 1987 Dollars			
Year	**NPA**	**W&P**	**REMI**	**SCAG**
1975	14,234	14,231	14,828	13,832
1980	16,528	16,528	16,563	15,200
1985	17,669	17,669	17,870	15,870
1990	17,984	17,984	17,658	16,540
1995	18,175	17,641	17,764	18,865
2000	19,362	19,276	19,206	21,190
2005	19,985	21,113	20,411	23,750
2010	20,988	23,082	21,180	26,310
2015	21,873	25,174	21,723	26,550

Abbreviations: NPA, National Planning Associates; PCPI, per-capita personal income; REMI, Regional Economic Models, Inc.; SCAG, Southern California Association of Governments; W&P, Woods and Poole Economics.

Source: Landrum and Brown, 1996.

TABLE 19.2 PCPI for Los Angeles as Reported and Projected by NPA, W&P, REMI, and SCAG

The bottom line is that an honest forecasting process leads to a wide range of results. This is because it is possible to fit several models [such as Eqs. (19.2–19.4)] to any specific span of data and because there are many forecasts of the secondary factors. For example, Table 19.3 provides some of the possible forecasts for domestic passengers between 1995 and 2015. The spread of forecasts in this case is between 89.2 and 52.7 million. The range is 36.5 million, or between 41 and 70 percent of the value. This is comparable to what one expects in practice, based on retrospective analyses of the accuracy of forecasts, as Chap. 4 indicates. These forecasts can now be compared to the reality that the 2010 traffic at Los Angeles served 42 million domestic passengers. To put the matter simply, the analysis did not provide confidence in the first digit, let alone the second or third decimal.

Year	Actual Data	Master Plan	No Employee Factor	Yield to Power	Exponential
1975	19,370				
1980	29,546				
1985	34,296				
1990	35,756				
1995	38,024	38,024	38,024	38,024	38,024
2000	49,927	50,875	49,372	45,077	48,513
2005	44,003	63,988	61,770	47,386	56,589
2010	42,085	78,544	75,328	49,921	66,009
2015		89,157	85,409	52,718	76,996

Source: Landrum and Brown, 1996; Los Angeles World Airports, 2000–2010.

Table 19.3 Forecasts for (thousands of) Domestic Passengers for Los Angeles/International Generated by Several Models Associated with the Master Planning Effort

The only way to develop a single forecast is to apply judgment. Mathematics or analysis by themselves cannot resolve the differences between the outcomes of the formulas. Judgment and opinion are essential parts of the process, both to select the mathematical forms and to reconcile the discrepancies between possible results. For Los Angeles and elsewhere, the selection of a single number as the forecast for 20 years ahead represents some effort to pick an acceptable value, a middle value, or an allegedly most likely value. It is an artful, not a scientific, decision.

The case of Miami/International illustrates how art rather than analysis ultimately drives forecasts in practice. A consulting team prepared a range of 30-year forecasts for domestic passengers (Table 19.4). They generated these forecasts by assuming

- Different spans of data (either Dade County alone, Dade augmented by Broward County, time series for Miami/International, and time series for the entire United States)
- Different variables (population, yields, and PCPI)
- Different forms of the variables (linear and nonlinear)
- Different forms of the equations (the five different ones listed)

Forecast Method and Variant		Forecast	Actual		
Method	**Data Used (form)**	**2020**	**1990**	**2000**	**2010**
Population	Dade County	13.96	9.92	17.4	19.5
	Dade and Broward	15.35			
	Dade and Broward (nonlinear)	17.74			
Yield and per-capita personal income	Dade County	19.87			
	Dade and Broward	19.69			
	Dade and Broward (nonlinear)	19.13			
Time series	Dade County	17.41			
	Dade and Broward	18.67			
	Dade and Broward (nonlinear)	40.05			
Per-capita personal income	Dade County	26.58			
	Dade and Broward	24.34			
	Dade and Broward (nonlinear)	42.4			
Share of the United States		23.48			
	Maximum	42.4			
	Average	22.97			
	Medium	19.69			
	Minimum	13.96			
	Preferred	**15.35**			

Source: Landrum and Brown, 1992; Miami Airport, 2001; ACI, 2010.

Table 19.4 Preferred Forecast Selected from Range of Forecasts, for Domestic Passengers at Miami/International

Once they had these 13 different forecasts displayed in Table 19.4, the forecasters decided to select one. In this case, their judgment led them to a "preferred" forecast for the year 2020 (a linear correlation with the regional population). Having to select a single forecast from the range of possibilities is not a scientific process.

In the case of Miami/International, note that the actual traffic in the year 2000 already surpassed the level of the preferred forecast for the year 2020. As seen by the data for 2010, the traffic is on course to be around 50 percent higher than forecast and not even the first figure will be correct! As stressed in Chap. 4 individual forecasts are "always wrong."

19.4 Scenarios

The mathematical approach to forecasting presented in the previous sections is not appropriate to long-term forecasts. Regression and other statistical analyses presume that past trends continue. In the short run, over a few years, the inertia in the system is certainly likely to continue existing trends. In the longer run, however, over 10 to 20 years, the assumption that past trends continue indefinitely is likely to be false. Newer trends eventually displace older ones. For longer-run forecasts, it is important to use judgment explicitly to develop forecasts.

A *scenario*, in terms of forecasting, is a concept of what might happen at some future time. Analysts develop scenarios to create a context that defines a long-term forecast. Scenarios can deal with all the factors relevant to airport development. For example, they might focus on the following:

- Macroeconomics, which might suggest the growth or fall of local industry
- Demographics, such as the stagnation and aging of the population and of the travelers
- Competitive airports, whose activity might lead to a shift in traffic
- Consolidation of the airline industry that might change the location of traffic hubs
- Environmental and resource constraints, which might inhibit growth
- Saturation of demand, in line with experience elsewhere

Analysts develop scenarios in consultation with experts in the several fields of interest. The scenarios represent judgments about the future.

Forecasters use scenarios to provide a rationale for bending old trends into new directions. For example, they might argue that a new airport such as Dubai will attract new airlines and traffic to a region because it provides more competitive, less congested facilities. Alternatively, they might suggest that recent rapid rates of growth would become lower, due to saturation of the market for air travel. Scenarios provide the means to override the continuation of past trends that seem obsolete.

19.5 Integrated Procedure

A responsible forecasting procedure recognizes that

- Careful analysis of data is important.
- Judgments are an integral part of the exercise.

It will not rely exclusively on either a mathematical analysis or pure judgment and opinion. It will combine both appropriately.

Good forecasting will also acknowledge that current trends are likely to dominate during the near term. Over the short term, trends persist due to the inertia in the air transportation system, as the ACI Airport Forecasting Manual indicates (Airports Council International, 2011). Passengers and shippers maintain their habits and relations. Airlines change their operations and fleets slowly. In the longer run, however, new trends will emerge. These new patterns can only be guessed at, because they do not constitute a major portion of the existing trends. Analysis should dominate short-term forecasts and judgment should drive the long-term estimates.

The recommended procedure balances analysis and judgment (de Neufville et al., 1980). It has the following five elements:

1. *Obtain and verify data on past traffic and relevant factors*. These observations should be carefully examined for consistency and correctness, along the lines indicated in Sec. 18.1.
2. *Do regression analyses to develop a model of traffic*. Examine several forms of models [such as Eqs. (19.2–19.4)] over different spans of time. Use judgment to decide which models are most suitable for local conditions.
3. *Project statistical models over the short term (5–10 years)*. It is a good idea to look at the range of forecasts implied by the most relevant models developed in step 2; the analyst can then estimate a middle forecast and a range of possible outcomes.

4. *Develop scenarios of future conditions* suitable to local region and situation.
5. *Estimate long-term (10–20 years) forecasts with wide ranges*. Do this by using the scenarios of step 4 to modify the short-term trends. In recognition of the unavoidable uncertainty in guesses at the future, be sure to associate wide ranges with the median long-term forecasts. Based on past experience, suitable ranges on 20-year forecasts are about plus or minus 30 percent (see Chap. 4).

Exercises

19.1. Obtain traffic data for an airport of interest. Use a spreadsheet program to establish a trend line using one or all of Eqs. (19.2–19.4). Project these trend lines forward—how do they differ?

19.2. Repeat Exercise 19.1, using a different span of data (e.g., 10 years back instead of 15 or 20). How do the results differ from those obtained in Exercise 1? Think about which is the appropriate span of data to consider, considering the relevant local circumstances.

19.3. Think about how you would obtain data on future prices, income levels, and other factors that might be relevant to a model of air traffic. How would you obtain reliable, credible forecasts of these factors?

19.4. What scenarios are relevant to the future levels of traffic for some airport of interest? How might these affect the forecasts? How would you develop consensus that these scenarios are appropriate for your airport and region?

References

ACRP Airport Cooperative Research Program (2007) "Airport Aviation Activity Forecasting: A Synthesis of Current Practice," ACRP Synthesis 2, Transportation Research Board, Washington, DC. www.trb.org.

Airports Council International (2011) *ACI Airport Traffic Forecasting Manual: A Practical Guide Addressing Best Practices*, Montreal, Canada. www.aci.aero.

Allison, P. (1998) *Multiple Regression: A Primer*, Pine Forge Press, Thousand Oaks, CA.

de Neufville, R., Zuniga, S., Kanafani, A., and Olivera, A. (1980) "Forecasting Airport Traffic: Mexico City as a Case Study," *Transportation Research Record, 732*, pp. 24–29.

ICAO, International Civil Aviation Organization (2006) *Manual on Air Traffic Forecasting*, 3d ed., Doc. 8991-AT/722/3, ICAO, Montreal, Canada.

Landrum and Brown (1996) *LAX Master Plan—Chapter III, Forecasts of Aviation Demand*, prepared for Los Angeles Department of Airports, Los Angeles, CA.

Landrum and Brown (1992) *Draft Master Plan for Miami International Airport*, Miami, FL.

Makridakis, S., Wheelwright, S., and Hyndman, R. (1997) *Forecasting: Methods and Applications*, 3d ed., Wiley, New York.

Pindyck, R. and Rubinfeld, D. (2000) *Econometric Models and Economic Forecasts*, 4th ed., McGraw-Hill, New York.

---

[1]The standard literature refers to the secondary factors as the "independent" or "explanatory" variables that "explain" the "dependent" variable that the analyst wishes to predict. These terms are not used here because they are misleading. Only in laboratory situations can the researcher control the situation and definitely know which variables cause a phenomenon (e.g., weights on the beam cause it to bend). In dealing with social situations, such as travel, we cannot assume which factors cause which events. In fact, many factors interact with each other and the direction of causality is ambiguous. For example, lower prices stimulate demand for a product, and higher demand can lead to lower costs and prices. Labeling one set of variables as explanatory or independent presumes that one knows what is happening, when actually the analyst is trying to that find out.

CHAPTER 20

Flows and Queues at Airports

Practically every airside and landside facility and service can be viewed as a queuing system. Prospective users (aircraft, passengers, bags, or other entities) form queues at these facilities and services and wait for their turn to be served. Flow analysis and queuing theory provide important tools for studying and optimizing these processes.

Queuing systems consist of three fundamental elements: a user source, a queue, and a service facility that contains one or more identical servers in parallel. Users arrive at the queuing system at instants described by the probability distribution of the demand interarrival times. Demand rates can be constant over time, but at airports they usually vary with the time of the day, the day of the week, and the season. The service process is described by the service rate and by the probability distribution of the length of service times.

Many measures of performance and of level of service (LOS) at airport queuing systems are of interest. Some of the principal ones include the intensity of utilization of the facility or service, the expected number of users in queue and the expected waiting time, the variability of queuing time, the reliability and predictability of the system, and the extent to which users perceive the system to be orderly and "fair."

Overloads exist whenever the demand rate exceeds the service rate. During overloads, the average delay per facility user increases linearly with the length of the overload period. Cumulative flow diagrams provide a convenient way of visualizing and analyzing what happens at a queuing system under overload conditions. The mathematical analysis of cumulative diagrams is simple and intuitive.

A queuing system cannot be operated, in the long run, with a demand rate that exceeds, on average, the service rate, as this will result in unacceptable delays. However, delays and congestion may also be present during periods when the demand rate is less than the service rate. Such delays are due to the probabilistic fluctuations of demand interarrival times and of service times. They are called "stochastic delays" to distinguish them from overload delays. When the demand rate is lower than but close to the service rate, stochastic delays can be very significant. They will increase nonlinearly as the demand rate approaches the service rate. Small changes in the demand rate or the service rate can thus have a large impact on the magnitude of delays and the length of queues. Moreover, the variability of delays and of queue lengths increases as the demand rate approaches the service rate. Queuing theory provides a number of important closed-form expressions for estimating stochastic delays under certain conditions. It also suggests some important guidelines for planning and designing airport facilities and services and managing airport operations.

20.1 Introduction

Queuing theory is the mathematical study of congestion. It explores the relationship between demand on a service system and the delays suffered by the users of that system. Because all airports—in their entirety or broken down into their individual elements—can be viewed as networks of queuing systems, queuing theory often plays a central role in the study and management of airport operations and in the planning and design of airport facilities and services. Those who wish to apply its results to airports should appreciate the kinds of questions that the theory can answer. They should also understand the nature of the assumptions behind the answers.

In working with queuing theory, one must study the particular airport facility of interest and specify a mathematical model to represent it. The analyst can either create this model or simply choose from a list of models that have already been studied. Through the model, one then computes the statistics that describe the behavior of the facility under the postulated conditions. Inherent to the process of creating and working with a mathematical model are the notions of *simplification* and *approximation*.

To make the analysis tractable, many details about the facility are necessarily disregarded as superfluous (or of minor importance) to the central points of interest. The details to be omitted from the model must be chosen carefully if the model is to resemble reality adequately. Data about the airport facility will also often be incomplete necessitating further assumptions and "educated guesses." Under the circumstances, the estimates of the quantities of interest obtained through a queuing analysis are, in most applications, only approximate indicators of the magnitude of these quantities in the real world. Consequently, the application of queuing theory is most useful in helping identify the inadequacies of existing facilities and services. It indicates the general directions in which to proceed for improving these facilities and services. It can also suggest the approximate values that some of the controllable variables must have if the queuing system is to achieve a satisfactory level of performance.

This chapter presents a general introduction to the application of flow analysis and queuing theory in the airport environment. The emphasis is on the fundamental concepts, on describing the behavior of queuing systems from a short- and a long-term perspective, and on the implications of this behavior for airport planners and managers. Many operations research textbooks (see, e.g., Hillier and Lieberman, 2009, or Larson and Odoni,

2007) offer more detailed introductions, whereas specialized books (e.g., Gross et al., 2008, or Wolff, 1989) provide advanced treatment.

20.2 Describing an Airport Queuing System

The generic model of a *queuing system* (Fig. 20.1) consists of three elements: a user source, a queue, and a service facility that contains one or more *identical servers in parallel*. Each user of a queuing system is "generated" by the user source, passes through the queue where (s)he may remain for a period of time (including possibly zero time), and is then processed by one of the parallel servers.

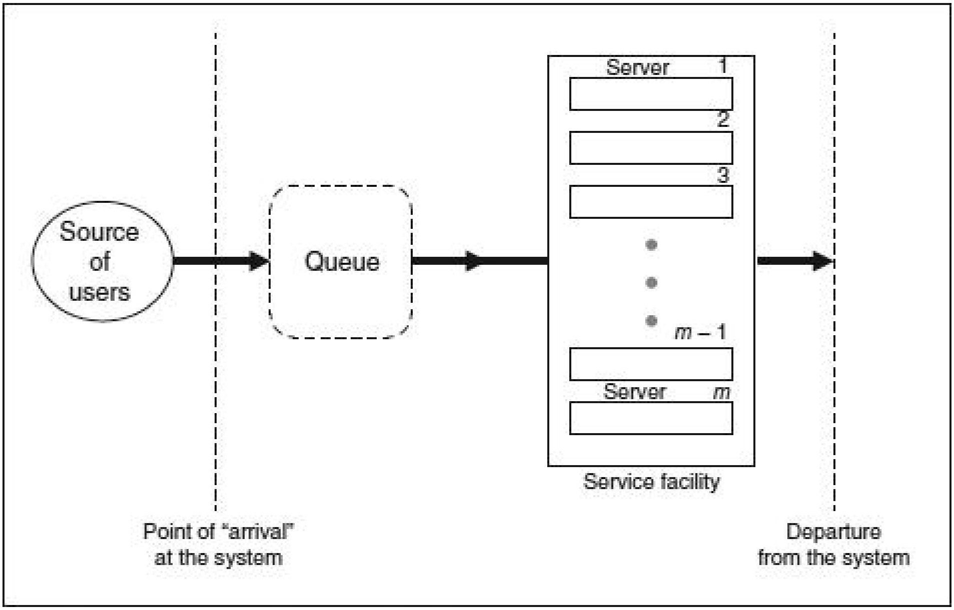


FIGURE 20.1 A generic queuing system.

A *queuing network* is a set of interconnected queuing systems (Fig. 20.2). In a queuing network, the user sources for some of the queuing systems may be other queuing systems in the network. For example, the output of a check-in desk (i.e., the passengers who have completed that process) is typically one of the sources for the queues that form in front of a security checkpoint. As noted previously, any airport or its elements can be viewed as a queuing network.

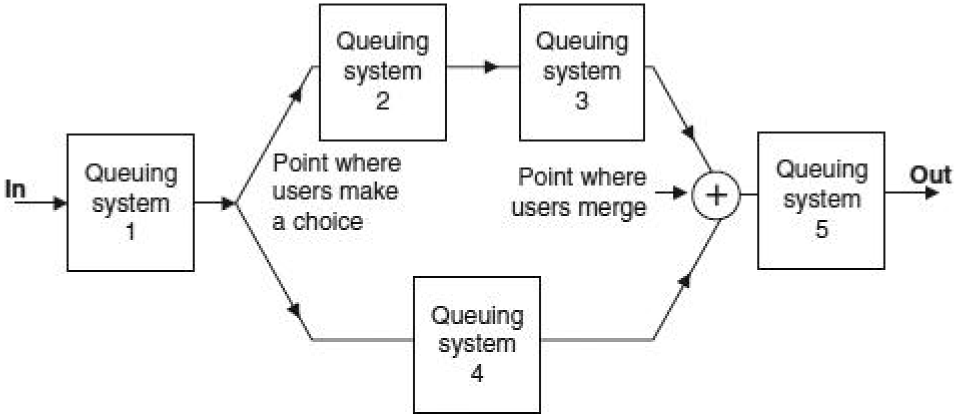


FIGURE 20.2 A queuing network consisting of five queuing systems.

To specify fully a queuing system, information must be supplied about its three generic elements: the user generating process, the queuing process, and the service process. The following discusses briefly these processes in the specific context of airports.

The User Generation Process

Flows of prospective users to airport facilities are described primarily by the following:

- The *rate* at which they occur over time (known as the *demand rate*), that is, the expected (average) number of demands per unit of time.
- The *probability distribution* of the time intervals between successive demands. These time intervals are often referred to as *demand interarrival times* or simply *interarrival times*.

The demand rate is typically denoted by the Greek letter λ (lambda).[1] For example,

λ = 40 aircraft departures per hour between 07:00 and 10:00

λ = 2000 arriving passengers per hour from 09:00 to 13:00 at a busy passenger building

If the demand rate is dynamic; that is, it varies as a function of time, this may be made explicit by using the notation $\lambda(t)$.

For any particular demand rate, the probability distribution for the length of the time intervals between successive demands is important in determining the performance of a queuing system. In general, the occurrence of demands at airport facilities is almost never a deterministic process; that is, demands do not appear at exactly their scheduled times or at perfectly spaced points in time (e.g., every 30 seconds). Instead, the time intervals between successive demands are almost always subject to some uncertainty. Random events and various levels of unpredictability, both airside and landside, are the rule rather than the exception in the airport environment. Thus, it is not surprising that even the Poisson process, which describes highly random behavior, often turns out to be a good approximate model for the generation of user demands at many airport facilities. At busy airports, for example, the instants when arriving airplanes come within a 100-km radius from the airport can often be approximated statistically as random events generated according to a Poisson process with a dynamic demand rate $\lambda(t)$, which varies according to the time of day. It is true that these instants are related to a preset flight schedule. However, daily deviations from this schedule are typically sufficiently large to make plausible the use of a (time-varying) Poisson demand model. In summary, the probabilistic behavior of user demands at airport facilities typically falls somewhere between perfectly deterministic and perfectly random (Poisson), that is, between the opposite poles represented by the constant and by the negative exponential demand interarrival times, respectively. Example 20.1 further explains this statement and introduces two fundamental types of distributions.

In addition to the variability of interarrival times, a further complication is that demands for airport facilities and services often appear in groups (*batch demands*) rather than individually. For example, departing passengers often show up at check-in desks in family groups of two or more people. More important, batch demands usually dominate when it comes to services and facilities for arriving passengers. These passengers typically come into the passenger building within a short interval of time, often in groups of 100 or more, following the arrival of an airplane at a gate. Batch demands are a major consideration in the analysis and planning of operations on the landside of airports, whereas they do not play a role on the airside, where one deals with the arrivals and departures of individual aircraft.

The Service Process

Entirely analogous ideas apply to the description of the service process. The *service rate* (or *capacity*), that is, the expected number of users that can be served per unit of time, can remain constant or vary dynamically over time. Service rates are usually denoted by μ (mu)—or $\mu(t)$ if time-varying. When a queuing system contains s identical parallel servers and the service rate for each server is μ, the total service rate for the queuing facility equals $s \cdot \mu$.

Example 20.1 Suppose the demand rate for some airport facility is equal to 60 per hour. At one extreme, demand requests at the queuing system could occur at intervals of exactly 1 minute. This is the case of *constant* (or *deterministic*) demand interarrival times at the rate of $\lambda = 60$ per hour.

The opposite extreme is the case in which demands, while *on average* occurring at the rate of 60 per hour, are *completely randomly distributed* over time. This naturally implies that some 1 hour intervals will have more than 60 demands, whereas others will have fewer. Moreover, within any interval (1 hour, 93 minutes, or whatever), the instants when the demands occur are distributed completely randomly in time and independently. For instance, suppose it is known that there were 86 demands over a 93-minute interval. Then the instants when these demands occur could be "simulated" in the statistical sense, by taking a 93-minute timeline and throwing randomly, as if blindfolded, 86 darts on it. Each dart would have equal probability of "landing" anywhere in the interval between $t = 0$ and $t = 93$ minutes, no matter where the other darts have landed. The process just described has a special mathematical meaning and is known as the *Poisson process*.

The Poisson process is described by demand interarrival times with the *negative exponential* probability density function shown in Fig. 20.3. Observe that, in qualitative terms, short demand interarrival times occur with high probability, while some very long interarrival times are also possible but with low probability. The expected (or "average") length of a demand interarrival interval is equal to $1/\lambda$, the inverse of the number of demands per unit of time. When the occurrence of demands at an airport facility is approximately Poisson, there is a significant probability of observing "bunches" of demands within relatively short intervals of time, interspersed between periods with low demand activity due to the presence of one or more long interarrival times. Because of the potential bunching of demands, users of queuing systems where the occurrence of demands can be approximated by the Poisson process are much more prone to experiencing delays than users of systems with approximately constant demand interarrival times.

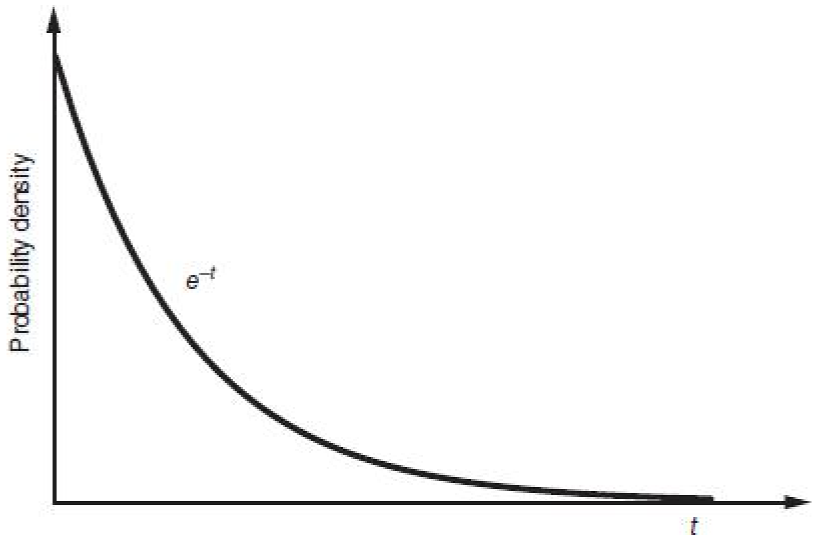


FIGURE 20.3 The negative exponential probability density function.

Obviously, any number of probability distributions, other than the two just described, could be more appropriate for modeling the demand interarrival times to a queuing facility, depending on the situation at hand.

Service rates at airport facilities may vary strongly over time. For example, the service rate of a runway system usually depends on weather conditions (see Chap. 10). Given a 24-hour weather forecast at the beginning of a day, one can specify a function $\mu(t)$ for the capacity expected to be available per runway during that time span.

The probability distribution that best describes the duration of *service times* varies from one airport facility to another. In some cases, service times may be more or less equal for all users—and can sometimes even be considered as approximately constant. In other cases there may be more variability, requiring probability distributions, such as the negative exponential or others, whose characteristics include a wide range of possible values for the service times and, consequently, a large variance. For example, service times at well-operated security (x-ray) checkpoints can be fairly constant. However, check-in counters often display wide variability, with average service times in the order of 1.5 to 2 minutes per passenger at many airports, but with some passengers taking as little as 1 minute and others as long as 5 minutes or more.

There are practical instances in which the service rate of a queuing system increases or decreases as a function of queue length. Several such examples can be found at airports. For instance, officers performing passport control often speed up service when a long line forms in front of their desks—sometimes limiting "control" to a nod to the passenger to move on. The reverse may also be true. Long queues at check-in counters sometimes confuse and distract check-in agents, resulting in reduced service rates.

The Queuing Process

To describe the queuing process, one must provide details on how prospective users line up for access to system service, as well as on their behavior while in queue. Numerous variations of queuing processes exist in practice and, interestingly, most of them can be encountered at airports.

The most obvious issue about a queue concerns the *priority discipline*. Most queues at airports operate on a first-come, first-served basis (FCFS). Under this regime, prospective users (passengers or aircraft) line up for service according to the order of their arrival at the relevant queue. However, at some airports, conditions in terminal buildings often become sufficiently chaotic that the next user to be served is chosen more or less randomly [service in random order (SIRO)]. The priority discipline for the retrieval of luggage for incoming flights is also nearly SIRO. In most cases, the order in which deplaning passengers arrive at baggage-claim carousels and other retrieval devices has little bearing on the order in which their bags arrive.[2]

Many queuing systems divide users into classes. These classes then receive different priorities for access to the system's servers. The typical example at airports is the subdivision of passengers into first, business, and economy classes. First- and business-class passengers, as a rule, have priority for check-in and possibly other services as well. They may either go to the head of queues (bypassing economy-class passengers) or, more often, have some service counters reserved for them. These counters may also serve economy-class passengers when no higher-priority passengers are present. Similarly, when a runway serves both landings and takeoffs, arrivals generally receive priority over departures, although the details of this practice vary considerably according to country and location.

Another important design parameter at airports is whether prospective users line up in a single queue or parallel queues. Check-in is again the prototypical example. At most airports, when several adjacent check-in desks serve a flight,[3] a separate queue forms in front of each desk. If there are s parallel counters, there are also s queues. However, at a growing number of airports, airlines utilize a *single* queue for such a group of counters. This is sometimes called the *snake queue* because of the twisting shape that it is usually forced to take. One important advantage of the snake queue is that it saves space in crowded terminal buildings, as the contours of the waiting line are clearly designated and more passengers are forced into a smaller area. Another advantage is that it gives passengers a sense of fairness, as everyone is served in a FCFS way, eliminating the possibility that the queue that one chooses to join will prove to be a "slow" one. On the negative side, snake queues sometimes grow to lengths of literally hundreds of people, causing anxiety and occasionally anger to those that must join it. The overall service rate, when all s servers are busy, equals $s\mu$, in either case, where μ is the service rate per desk. Note, however, that with snake queues it often happens that not all servers are busy: the person at the head of the queue may not recognize that a server is open and will, in any case, take some time to reach the server; either way, servers distant from the end of a snake queue may often be idle, thus degrading the overall system capacity. For this reason, some airports have the snake queue feed secondary queues in front of each server, thus ensuring that all servers are busy if there are customers waiting.

A related issue when parallel queues are allowed to form in front of servers is whether to allow those joining a particular queue to switch to another queue if their own queue moves slowly. The issue arises in passenger processing, for example, at check-in, passport control, or security control. Some airlines and/or airport operators try to prevent such switching by placing barriers between queues. The rationale is that this makes for a more orderly process. On the negative side, preventing queue switching may lead to underutilization of some service capacity. Some servers may sometimes be idle while persons queued in front of other servers are unable to take advantage of the presence of idle servers.

Airlines and airports can manage queues by monitoring queue lengths and adjusting the number of active servers accordingly. Whenever the number of passengers waiting for service exceeds a certain limit (exactly or approximately specified), an airline or the airport operator can activate one or more additional desks/counters/servers. When the number waiting falls back below the same or some other limit, the number of active servers can be reduced as well. This practice naturally requires the availability of idle counters and related equipment, as well as of standby employees. Typically, these employees can be primarily engaged in some other activity, but they are available to staff a service position (e.g., a check-in counter) as necessary.

A crucial parameter in describing and designing airport queuing systems is *queue capacity*. This is the maximum number of prospective facility users that the waiting line can accommodate at any single time. At airports, space limitations typically determine this capacity. Examples are the length of taxiway that departing aircraft may use to line up for takeoff, the area available for waiting passengers in front of check-in counters, and the volume of terminal airspace available for "stacking" arriving aircraft waiting to land. These examples also suggest that it is often difficult to pinpoint exactly the capacity of airport queues. After all, many people can be crammed into any particular part of a passenger terminal for a short time. Queue size limitations nevertheless do exist at airports and often present a real problem, as well as create severe side effects. For example, departing aircraft have to be held at their apron stands, because the taxiway system is saturated with other airplanes waiting for takeoff (see Chap. 10). This, in turn, may result in a shortage of available gates for arriving aircraft. Another example of a side effect of queue lengths is instances in which passengers waiting for check-in block circulation in a departures concourse (see Chap. 16). It is important to recognize the possibility of such side effects and anticipate them in planning and designing airport facilities and services.

20.3 Typical Measures of Performance and Level of Service

Many measures describe the performance of a queuing system and the resulting level of service (LOS). This section reviews those most relevant to airport facilities and services.

Utilization Ratio

The *utilization ratio* is quite possibly the most fundamental measure of LOS. It "drives" all other measures of a queuing system's performance. It is denoted as ρ (rho). For a single-server queuing system with demand rate λ and service rate (i.e., capacity) μ, it is given by

$$\rho = \frac{\lambda}{\mu} \tag{20.1}$$

When the queuing system has *s* parallel and identical servers,

$$\rho = \frac{\lambda}{(s \cdot \mu)} \tag{20.2}$$

Intuitively, ρ indicates the "intensity" of utilization of the queuing system. It is often referred to simply as the *demand-to-capacity ratio*. A queuing system with ρ greater than 1 is called "saturated," for obvious reasons. Values of ρ close to but less than 1 are desirable if one's objective is to make maximum use of the productive capacity of a facility or resource. However, this may also entail important inefficiencies, such as long waiting times for access to the facility or resource (see Sec. 20.6).

Expected Waiting Time and Expected Number in Queue

It is convenient to define two quantities central to the description of the performance of a queuing system. Consider a queuing system over a particular interval of time. Define

$W_q =$ the waiting time in queue experienced by a user selected randomly among all those who visited the system during that interval of time.

$N_q =$ the number of users waiting in the queue at a randomly selected instant during that interval

In a typical queuing system, both W_q and N_q are *random variables*. This is because both the demand interarrival times and the service times at the queuing system are, in general, probabilistic quantities as discussed in Sec. 20.2. Therefore, both the waiting time experienced by a system user and the length of the queue at the system will also vary probabilistically.

The characteristics of W_q and N_q are of great importance in describing the performance of any queuing system. The most obvious and by far most commonly used of these characteristics are their expected (or "average") values, denoted $E[W_q]$ and $E[N_q]$, respectively. The waiting time that a random user of the queuing system will experience, on average, at an airport service or facility is of vital interest. So is the average length of the queue of passengers or aircraft waiting for service. Many of the airport airside delay statistics that are often cited by airlines, government agencies, and the media involve $E[W_q]$, the expected waiting time.

Variability

Expected values tell only one part of the story, however. Almost equally important to passengers and airlines is the *variability* of W_q and N_q. For example, it is one thing for departing passengers to know that the total delay while being processed at an airport (check-in, security control, etc.) is 20 minutes on average (i.e., total $E[W_q] = 20$ minutes) with a typical range of 10 to 30 minutes. It is quite another to know that total $E[W_q]$ is equal to 20 minutes, but with a range of 5 to 90 minutes. Experience shows that departing passengers will behave differently in these two cases. They will get to the airport considerably earlier, relative to their scheduled flight departure time, in the second case than in the first. Similarly, high variability of delay from day to day means that airlines have to construct the daily itineraries of aircraft, cockpit crews, and cabin crews with scheduled flight durations (gate-to-gate times) and/or turnaround times on the ground between successive flights that include considerable "slack time."[4] If they do not, they will find it difficult to execute their schedule of flights reliably, discussed as follows. The most common measure of variability of delay is the *variance* of W_q, denoted here as $\sigma^2(W_q)$, or the variance's square root, the *standard deviation* $\sigma(W_q)$. A large variance or standard deviation indicates high variability of delay.

Reliability

Reliability and variability are closely interrelated. The more variable the behavior of a queuing system, the more difficult and costly it will be to ensure that it operates reliably. Airports are a prime example of this relationship. Airlines and airports measure reliability, especially when it comes to airside operations, primarily through statistics regarding the frequency with which long delays occur. In effect, they measure the probability that W_q will exceed certain threshold values (or "tail of the distribution" values) that are considered critical to maintaining a reliable schedule of flight operations. For example, both the U.S. Federal Aviation Administration (FAA) and EUROCONTROL regularly collect and report statistics on the percent of flights arriving or departing more than 15 minutes behind schedule at each of the major airports in the United States and in Europe. The 15-minute value has been chosen because it is considered critical to both the passengers and the airlines. Typically, a delay of less than 15 minutes on arrival will not prevent the aircraft involved from departing on time for its next flight. Turnaround times on the ground between flights—as scheduled by the airlines—usually include sufficient "slack" to absorb a 15-minute delay on arrival without affecting the scheduled time of departure. However, arrival delays of more than 15 to 20 minutes usually "propagate" to the subsequent departure and thus have a more disruptive effect. If the probability of long delays of this type is high, the on-time execution of an airline's overall schedule of flights becomes problematic.

In trying to design reliable flight schedules, the scheduling departments of major airlines examine carefully the probabilities of long airside delays at each airport they serve. This can all be put in quantitative terms. For instance, suppose that the estimated probability that W_q for a particular arriving flight will exceed 15 minutes is 10 percent during the course of a scheduling season. Then a ground turnaround time that includes a 15-minute slack will ensure 90 percent reliability of the departure time of the next flight to be performed by the same aircraft.[5] Analogously, a number of airports specify LOS standards for the design and operation of their passenger terminals (see Chap. 15) partly in terms of the probability of extreme delays. For example, they may specify that "80 percent of passengers at check-in counters should experience a waiting time of less than 12 minutes."

Maximum Queue Length

In much the same spirit, planners and designers of passenger building facilities and services are interested in estimates of the *maximum queue length*. This measure of performance is not well defined, because the length of any queue at a busy airport may, with very low probability, become extremely long under certain combinations of events for short periods of time. What planners and designers truly want to estimate is the amount of space they should provide at each terminal facility or service to run only a small risk that this space will prove inadequate. To answer this question, planners usually adopt one of the following two approaches. The first, and more correct one, is to compute a value of N_q that will be exceeded only with a small pre-specified probability, for example, 5 percent, and define this value of N_q as the maximum queue length to plan for. For example, if it is found from a queuing model or a simulation that the line in front of a set of check-in counters will have fewer than 40 waiting passengers on 98 percent of days (and if this is considered adequate), the space in front of the counters will be designed to hold up to 40 people. This is very similar conceptually to the approach described in the previous paragraph for estimating reliability. The second approach is to perform a detailed simulation of airport operations on the design peak day (DPD) or design peak hour (DPH)—see Chap. 21—and use for planning and design purposes the maximum queue length observed at the facility or facilities of interest in all the simulation runs (or a smaller, but comparable value). This approach may lead to excessive queuing space allocations.

The Psychology of Queues

When measuring and evaluating the performance of queuing systems, one should not underestimate the importance of *psychological factors* (Larson, 1988). This is a subject of growing interest among queuing specialists. The basic point is that psychological factors play a central role in user assessments of the severity of delays and congestion at any facility or service. Airport operators and airlines can and should ease the unavoidably negative reactions of air travelers to airport delays by taking appropriate steps to influence *perceptions* of the situation.

At the most obvious level, the *physical environment* is a central influence on perceptions about the severity of delays. The more comfortable the environment (area per occupant, ventilation and temperature, availability of seating, ambience of space, etc.), the more tolerant airport users are of delays. Ashford (1988) has presented results from interviews showing that passengers react less negatively to delays as the number of square meters per passenger in the area used for waiting increases.

The availability of *information* is also crucial in shaping perceptions. Airport users generally react less severely to delays if given reliable information on the reasons for the delays and/or some advance estimate of how long a delay will be. A number of airports, for example, now display electronically the estimated time until the bags of passengers on each individual incoming flight will start arriving on the bag retrieval carousels. In recent years, airlines have also been displaying increasingly detailed information on the reasons for flight delays.

A third important aspect of user perceptions regarding delays has to do with the notion of *fairness*. Larson (1987) suggests that perceptions of fairness (social justice) in a queuing system are strongly related to the number of "slips" and "skips" that take place per unit of time. A *slip* is an event in which a user receives service before another user who arrived at the queuing system (joined the queue) before him or her. A *skip* is the opposite. The higher the number of slips and skips, the more "unfair" the system is perceived to be—leading to increasingly negative user reactions to any delays they experience. The snake queue (see Sec. 20.2) is one way in which slips and skips can be controlled. Some airlines and airport operators place a high priority on preserving an image of orderliness and fairness at their facilities and services.

Finally and importantly, perceptions about the severity of delay usually increase nonlinearly with its length. Passengers typically perceive a 20-minute wait for check-in as being more than twice as "bad" as a 10-minute wait. Thus the *avoidance of extremely long waiting times* at individual facilities and services not only increases reliability but also contributes psychologically to a more positive assessment of LOS by airport users.

20.4 Short-Term Behavior of Queuing Systems

This section and the next provide brief discussions of how delays and queue lengths grow over time and as a function of the utilization ratio, ρ, the "intensity" of use of the queuing system. The presentation is mostly qualitative and conceptual. The detailed behavior of queuing systems within a specific short period (short-term behavior) is described first. Section 20.6 takes a more macroscopic point of view (long-term behavior). In the airport context, "short term" usually means periods ranging from 1 hour to a full day of operations, while "long term" typically involves entire seasons or years.

Consider an airport queuing system where, over a 24-hour period, the demand rate and the service rate (or capacity) undergo a number of changes. Suppose also that the demand interarrival times and the service times are random variables; that is, vary from instance to instance in accordance with their respective probability distributions.[6] Three general statements can be made about the behavior of W_q and N_q.

1. *Overload delays* will certainly occur (and queues will form) at times when the demand rate exceeds the service rate (e.g., a 2-hour period when 90 flights per hour are scheduled to arrive or depart, while the capacity of the runway system is 80 per hour). This is because prospective users will arrive on average at the

queuing system at a rate greater than the capacity of the system to serve them. Colloquially, "demand exceeds capacity" during such periods. In general, when there is a time interval during which the demand rate exceeds the service rate, both the expected queue length, $E[N_q]$, and the expected waiting time, $E[W_q]$, for users arriving at the queuing system during that interval will grow. The growth will be in direct proportion to the length of the interval, T, as the next section shows.

2. Delays may also occur when the demand rate is *less* than the service rate. As Example 20.1 notes, this is due to the probabilistic fluctuations in the demand interarrival times and/or the service times, that is, to the likely presence of time intervals with "clusters" of short interarrival times of demands and/or of long service times. This type of delay is often called *stochastic*[7] *delay* to distinguish it from *overload delay*. If the utilization ratio, ρ, is close to but less than 1 during a fairly long time, stochastic delays can be very significant, that is, both $E[N_q]$ and $E[W_q]$ may become large for users arriving at the queuing system during such periods. In general, the higher the variability (as measured by the variance) of the demand interarrival times and of the service times, the higher the stochastic delays will be. Section 20.6 returns to these points.

3. *The dynamic behavior of queues is complex*. The complex behavior of airside delay is described in Chap. 11. The same complexity can also be observed at landside facilities. The waiting times and queue lengths experienced during any particular time interval depend strongly on the waiting times and queue lengths during previous intervals. Consider, for example, two different hours of the day at an airport, one early in the morning (e.g., 06:00–06:59) and the other in late afternoon (e.g., 18:00–18:59). Assume that conditions within the two hours are identical: the demand rates, λ, are equal in the two hours—and so are the service rates, μ. However, the first hour is typically preceded by a period of little demand and the second by a period of high demand. The delays and queue lengths will then be far greater in the afternoon hour than in the morning. Moreover, the exact magnitude of these delays and queue lengths will depend on the time history of the queuing system before the hours of interest, the values of λ and μ, and the probability distributions of the demand interarrival times and of the service times.

   An interesting aspect of the dynamic behavior of queues is that a lag may exist between the times when demand peaks and when delay (and queue length) peaks. This phenomenon is sometimes called *hysteresis*. It is entirely analogous to a daily experience of people who drive home from work on urban road networks. Those leaving work during the peak demand hour (e.g., between 16:00

and 17:00) will usually experience less delay than drivers going home 1 or 2 hours later—when the number of those starting their commuting trip is smaller. The reason is that by 17:00 or 18:00, traffic congestion has already built up, so that those entering the traffic join "the end of the queue" and experience longer delays. In effect, they suffer the consequences of following on the heels of the earlier overload. Airside and passenger traffic and queues at busy airports typically display this same type of hysteresis.

20.5 Cumulative Diagrams

Cumulative flow diagrams (or simply *cumulative diagrams*) provide a convenient way of visualizing and analyzing what happens at a queuing system under overload conditions. When such conditions exist, cumulative diagrams can be very useful in obtaining approximate estimates of delays at both landside and airside facilities of busy airports. Their simplicity and intuitive appeal make it possible to present this graph-based approach in some detail in this section. The presentation is based on the analysis of a generic situation. The "users" of the facility described as follows can include passengers, aircraft, bags, or other relevant entities.

Consider Fig. 20.4, in which the functions $\lambda(t)$ and $\mu(t)$ denote, respectively, the demand rate and the service rate over time at a service facility in an airport. Note that the service rate indicates the "maximum throughput capacity" of the facility in the terminology of Chap. 10. The units for both the demand rate and the capacity are "users per unit of time."

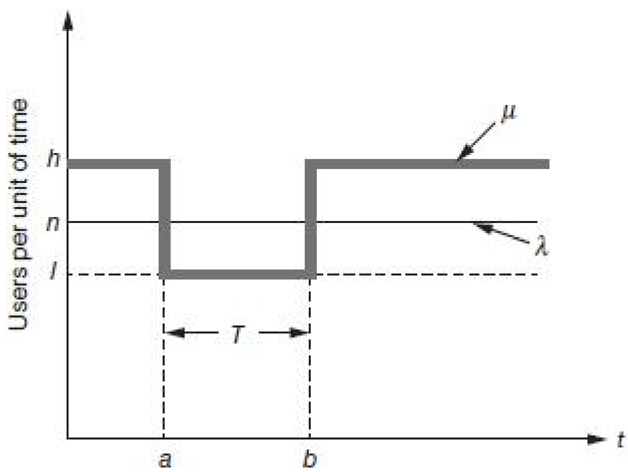


FIGURE 20.4 Demand and service rates at a queuing system.

Figure 20.4 approximates a situation that arises frequently at airports. The time axis shows the busy part of a day at the airport, for example, the origin may correspond to 07:00 local time. To simplify the analysis, it is assumed the demand rate is constant throughout the busy part of the day, that is, $\lambda(t) = n$ users per hour for all t. The normal service rate h (for "high") is greater than n. However, for the time interval between $t = a$ and $t = b$, the service capacity is reduced to l (for "low") users per unit of time. It will be convenient later on to denote with $T = b - a$, the interval of time during which the capacity is low.

The objective is to explore quantitatively the implications of the temporary overload on delay levels at this airport facility.

An important simplification in the analysis is achieved through the assumption that demands and service completions occur at evenly spaced intervals. For example, if $n = 60$, it is assumed that a user arrives at the facility to obtain service exactly every 60 seconds. Similarly, for any value of the service rate, h or l, the service times are assumed constant. If, for instance, $l = 30$ and the service facility is continually busy, a service is completed exactly every 2 minutes. This assumption, which is tantamount to adopting an entirely deterministic model, is used in practically all applications of cumulative diagrams.

Clearly, the situation shown in Fig. 20.4 will result in some delays to users during the time period between $t = a$ and $t = b$, at the very least. In fact, it is very easy to plot, as a function of time, the number of users waiting in the queue for access to the service facility. This is done in Fig. 20.5. There is no queue until $t = a$ because the facility's capacity is greater than the demand rate. Beginning at $t = a$, the queue builds up at a rate of $n - l$ users per unit of time, so that, by the time b, the queue will build up to a length of $(n - l) \cdot (b - a) = (n = l)\ T$ users. After time b, when the facility's capacity is "high" again, the queue length will decrease at the rate of $h - n$ users per unit of time, that is, at the rate at which users receive service minus the rate at which new users join the queue. Thus, it will take an amount of time equal to $(n - l) \cdot T/(h - n)$ for the queue to dissipate and get back to zero. Therefore,

$$\text{Total amount of time with a queue present} = T + \frac{(n-l) \times T}{(h-n)}$$

$$= T \times \frac{(h-l)}{(h-n)} \tag{20.3}$$

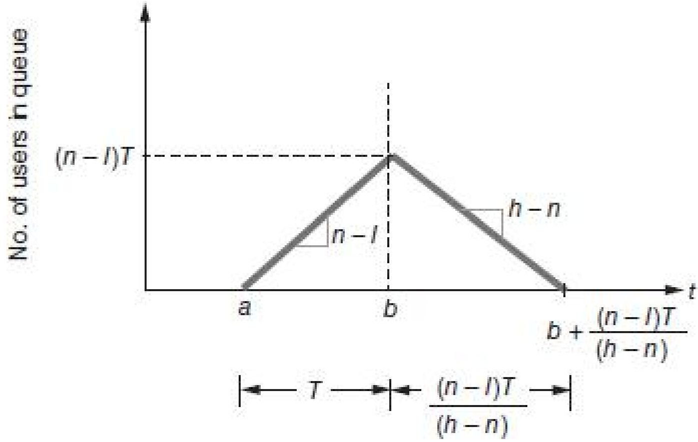


FIGURE 20.5 The number of users in the queue.

The area under the triangle in Fig. 20.5 represents the total amount of time that users will spend waiting in the queue, that is, the total delay time suffered by all delayed users due to the reduction in capacity between $t = a$ and $t = b$:

$$\text{Total delay time} = \frac{1}{2} \times (n-l) \times T \times \frac{(h-l)}{(h-n)} \times T$$

$$= \frac{1}{2} \times T^2 \times \frac{(n-l)(h-l)}{(h-n)} \tag{20.4}$$

Note that the total delay time increases with the *square* of the duration of the low-capacity interval. Equation (20.3) indicates that the number of users demanding service during the period when a queue exists is equal to $n \cdot [(h-l)/(h-n)] \cdot T$, the demand rate multiplied by the duration of the queue. This is the number of users that will suffer some delay. It is then possible to compute another quantity of interest by dividing the total delay time, given by Eq. (20.4), by the number of delayed users to obtain

$$\text{Expected delay per delayed facility user} = \frac{1}{2} \times T \times \frac{(n-l)}{n}$$

$$= \frac{1}{2} \times T \times \left(1 - \frac{l}{n}\right)$$

(20.5)

Note the meaning of Eq. (20.5): Given that a user was delayed, this is the delay that the user suffers "on average." It is remarkable that this expected delay (1) increases linearly with the duration, T, of the overload and (2) is a function only of the "low" capacity, l, and of the demand rate, n, and is independent of the "high" capacity, h.

An informative way to display the behavior of this queue is through the *cumulative flow diagrams* shown in Fig. 20.6. The cumulative flow diagrams typically show (1) the total (cumulative) number of user requests for service that have been made between $t = 0$ and the current time t, and (2) the total number of these requests that have been admitted for service up to the current time t. The former is the *cumulative demand diagram*[8] and the latter is the *cumulative admissions-to-service diagram*. Obviously, the difference between the cumulative demands and the cumulative admissions to service at any time t is the number of users *queued for admission* to the service facility. Figure 20.6 shows these two cumulative diagrams for the case at hand. Note that the vertical axis of Fig. 20.6 maintains a count of the cumulative number of demands and of admissions-to-service as they occur. For this example, the demand and admissions-to-service cumulative flow diagrams coincide up to $t = a$, because demands are admitted as soon as they show up at the service facility, given the availability of the high capacity h. At $t = a$, however, the number admitted begins lagging behind the demand, as it increases only with a slope of l users per unit of time. This lasts until $t = b$, when the number admitted starts increasing at the rate of h per unit of time, until it eventually "catches up" with the demand curve at time $(n - l) \cdot T/(h \cdot n)$ later. Thereafter the two cumulative flow diagrams coincide again and increase at the rate of n per unit of time, because capacity is higher than the demand rate.

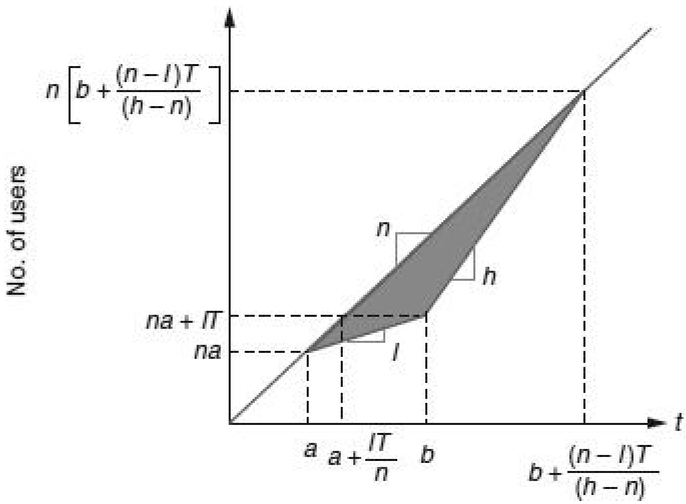


FIGURE 20.6 Cumulative demand and service diagrams

Note that the cumulative flow of admitted-for-service users can never exceed, by definition, the cumulative flow of the demands. At best, it can always be equal to the cumulative flow of demands, in cases where no delays occur due to the capacity being equal to or greater than the demand rate for all t. The (positive) vertical distance and the (positive) horizontal distance between the cumulative flow diagrams in Fig. 20.6 both have important physical interpretations. The *vertical distance* at any time t gives the *number of users in queue* at that time. It shows the difference between the number of users that have demanded service up to time t and the number admitted for service. Plotting that vertical distance as a function of time for Fig. 20.6 leads to Fig. 20.5. Similarly, the *horizontal distance* gives the *delay suffered by the i*-th demand, that is, the time that elapses between the instant when the i-th demand requests service and the time when that demand is admitted for service, assuming that demands are admitted for service in FCFS order.[9] Note that the total "area" between the two cumulative flow diagrams is equal to the total delay experienced by delayed users measured in units of user time (e.g., passenger-hours). Thus, this area is equal to the quantity shown in Eq. (20.4) (see Exercise 20.1).

The user that demands service at the instant when the horizontal distance between the cumulative flow diagrams is greatest is the one that will suffer the longest delay, assuming a FCFS priority discipline. From Fig. 20.6 it is clear that this is the user that will be *admitted* for service at $t = b$, because the horizontal distance between the two cumulative flow diagrams begins decreasing immediately thereafter. It can be seen that this is the $(n \cdot a + l \cdot T)$-th user to demand service (and to be admitted for service). This user demands service at the time t that satisfies $n \cdot t = n \cdot a + l \cdot T$, that is, at $t = a + (l \cdot T)/n$. Because it is already known that this user will be admitted for service at exactly $t = b$, the delay the user will suffer is given by

$$\text{Longest delay suffered by any user} = b - \left(a + \frac{l \times T}{n}\right) = T - \frac{l \times T}{n} = T\left(1 - \frac{l}{n}\right) \tag{20.6}$$

Note that the longest delay is twice the average delay shown in Eq. (20.5). From Fig. 20.6, taking advantage of the observations in the last paragraph, one can now prepare Fig.

20.7. This shows, for all values of t, the amount of delay that a demand *requesting service* at time t will suffer, under the FCFS assumption.

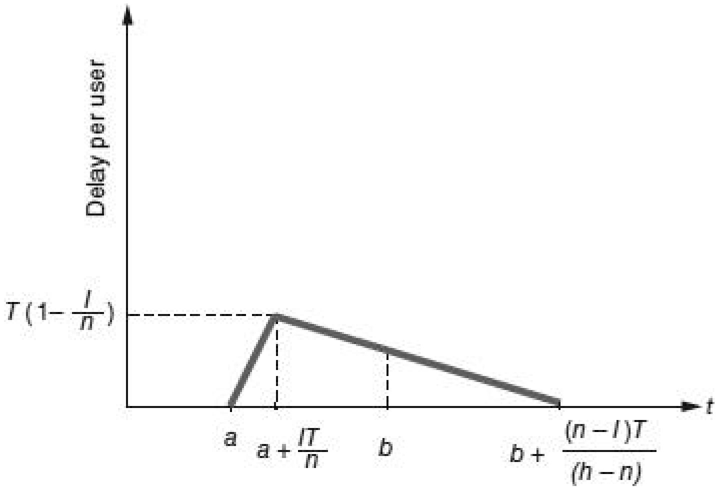


FIGURE 20.7 The amount of delay suffered by a user requesting service at time t.

Example 20.2 It is instructive to assign some realistic values to the various parameters of Fig. 20.4 and look at the implications. Assume that the situation examined is the queuing of aircraft for use of the runway system at a major European airport. This example uses values that may be typical of such an airport.

Let the time units be hours, and set $t = 0$ to correspond to 07:00 local time, $a = 1.0$ and $b = 4.0$, so that the duration, T, of the low-capacity period is 3 hours. The time a may thus correspond to the beginning of a period of fog lasting from 08:00 to 11:00. Let also $n = 60$ aircraft movements per hour, $h = 70$ aircraft movements per hour, and $l = 35$ aircraft movements per hour. The poor weather conditions reduce the capacity to one-half its normal value—not an unusual phenomenon in practice.

Equation (20.3) indicates that a queue will be present for a period of 10.5 hours beginning at 08:00 local time. The after-effects of the poor weather thus persist for 7.5 hours after the weather event ends at 11:00! With 60 aircraft movements scheduled per hour, a total of 630 movements will suffer some delay. The peak length of the queue is 75 movements (see Fig. 20.5 or 20.6) and occurs at 11:00 local time, but the peak delay time is suffered by the movement scheduled for $t = 2.75$ or 09:45 local time. This movement will suffer a delay equal to 1.25 hours, or 75 minutes [from Eq. (20.6)] and will actually reach the runway system at 11:00, exactly the instant when the weather event ends. The total amount of delay time incurred during the day, from Eq. (20.4), is equal to 393.75 aircraft-hours!

The economic cost of this delay depends primarily on the mix of aircraft at this airport and on whether the delays will be absorbed while the aircraft are airborne or on the ground (see Chap. 13). Just to indicate the order of magnitude of the cost, assume \$3600 as the direct operating cost to airlines of one aircraft-hour—a typical amount for aircraft using major airports. This gives approximately \$1,400,000 as the total cost of 393.75 hours of delay due to the 3-hour weather event, *not* including the cost of delay time to the passengers! From Eq. (20.5), the average delay per movement for the 630 movements delayed is 0.625 hour or 37.5 minutes, for a cost of \$2250 per aircraft at the assumed \$3600 per aircraft-hour.

The simple case examined in this section demonstrates, among other things, how cumulative flow diagrams should be drawn and interpreted. It is straightforward to extend this approach to the more general case in which (1) both the demand rate and the capacity vary over time and (2) there are several, not just one, instances during the day when the demand is higher than the capacity. Figure 20.8 shows a typical example of a cumulative flow diagram for such a more general case. The only differences from Fig. 20.5 are that (1) the "demand" and the "admitted-for-service" cumulative flow diagrams undergo several slope changes and (2) there may be several instances during the day when there is a backlog of demands waiting to be served, that is, when delays will occur. Note that the cumulative diagrams need not be piecewise linear functions, as in Fig. 20.8. They can have other shapes, depending on the functional forms used to describe the demand rate, $\lambda(t)$ and capacity, $\mu(t)$, during the time interval of interest, such as continuous nonlinear functions without breakpoints.

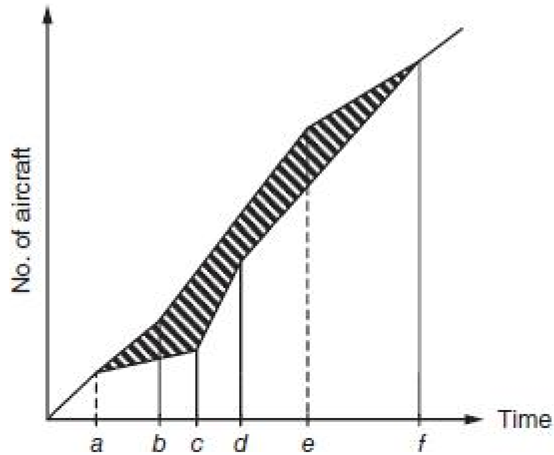


FIGURE 20.8 Cumulative demand and service diagrams for a more general case.

The deterministic model and approach described in this section have a fundamental deficiency: They capture *only* overload delay. For instance, in Example 20.2 it is assumed that there would be *no delay* at the runway system if the airport could operate all day at the high-capacity level of 70 movements per hour, which exceeds the demand of 60 movements per hour. In practice, however, one might see some very significant delays under such conditions, if the demands occurred sufficiently randomly in time to create periods of traffic surges and/or if the service times on the runway system exhibited significant variability. *Stochastic* queuing models and analysis can capture delays of this type.

20.6 Long-Term Behavior of Queuing Systems

This section discusses the behavior of queuing systems from a *long-term perspective*. It looks at the typical characteristics of a queuing system under *equilibrium conditions*, that is, over a span of time that is sufficiently long to allow the system to "settle down" to a statistically repetitive type of behavior. This does *not* mean that there will be no variability in W_q, N_q, and other related quantities over time, but that this variability involves probabilistic fluctuations around certain long-term average characteristics. These long-term average characteristics are precisely those that one wishes to observe from a macroscopic viewpoint. Stated more mathematically, the discussion focuses on the statistical description of the queuing system as the period, T, during which it operates tends to infinity.

A fundamental condition needs to be satisfied if a queuing system is to reach long-term equilibrium behavior, or *steady state* in the terminology of queuing theory. That condition illustrates in a simple way the meaning of "long-term equilibrium." The last section examined the behavior of a queue when, for a specific interval of time, T, the demand rate exceeds the service rate, that is the utilization ratio is greater than 1. From the "short-term" point of view, operating a queuing system in overload conditions ($\rho > 1$) is perfectly acceptable; queue length and delays will increase during the interval T and then will presumably decrease when the period of overload is finished. However, in the long term, one cannot operate a queuing system continuously in overload conditions and reach any type of equilibrium. If ρ is greater than 1 *on average* over time, for the very long period of observation, T, the length of the queue and the waiting time at the queuing system will grow without bound over time. This would happen even if demand exceeded capacity, on average, by only a minuscule amount per unit of time. The server(s) would then, on average, "fall behind" demand during every unit of time, so that more and more prospective users would accumulate in the queue. It follows from this argument that a queuing system can reach long-term equilibrium conditions only if $\rho < 1$, that is, only if the long-term demand rate is strictly less than the long-term service rate.[10]

Two fundamental observations from queuing theory can now be presented: Little's law and a description in general terms of the nonlinear behavior of queuing systems.

Little's Law

In addition to W_q and N_q, queuing theory is interested in the characteristics of two other random variables:

W = total amount of time spent by a user in a queuing system

N = total number of users in the queuing system

Note that W is simply equal to the sum of the amount of time, W_q, that a user spends waiting to be admitted to service and the time the user spends in service. Similarly, N is the sum of the number, N_q, of users waiting in queue and the number of users being served. (In a single-server queuing system the number of users being served is at most one.) When a queuing system is in steady state, the expected values of the four important random variables, W, W_q, N, and N_q, satisfy the following three relationships:

$$E[N] = \lambda \cdot E[W] \tag{20.7}$$

$$E[N_q] = \lambda \cdot E[W_q] \tag{20.8}$$

$$E[W] = E[W_q] + \frac{1}{\mu} \tag{20.9}$$

Equation (20.9) follows directly from the definition of W. If a server processes μ users per unit of time on average (the service rate), the expected service time is equal to 1/μ. The expected value of the total time in the system, $E[W]$, is then given by Eq. (20.9).

Equations (20.7) and (20.8) are both statements of Little's law (Little, 1961). The following argument provides an intuitive explanation (but not proof) of this law, as expressed by Eq. (20.7). Consider a queuing system operating with a FCFS priority discipline. In the steady state, that is, with the system in equilibrium, the average number of users that a random user finds at the queuing system upon arrival should be equal to the average number (s)he leaves behind upon departure, with both of these numbers equal to $E[N]$. However, the average number of users left behind is simply the demand rate per unit of time, λ, times the average amount of time, $E[W]$, that a random user stays in the system!

Relationship between Congestion and Utilization

Queuing theory has led to the discovery of the following relationship between congestion at a queuing system and the intensity with which such a system is utilized:

Under steady-state conditions, $E[W]$, $E[W_q]$, $E[N]$, *and* $E[N_q]$, *at any queuing system, increase nonlinearly with* ρ, *in proportion to the quantity* $1/(1 - \rho)$.

A graph of $E[W_q]$ for a particular queuing system is shown in Fig. 20.9, where the horizontal axis is ρ, the long-term demand rate as a fraction of the long-term service rate. This graph is typical of the behavior described by the above relationship. As ρ approaches 1, or as the demand rate approaches the service rate (or, colloquially, as "demand approaches capacity"), $E[Wq]$ increases nonlinearly in proportion to $1/(1-\rho)$. It reaches infinity at $\rho = 1$, as explained in the discussion on the necessary condition ($\rho < 1$) for equilibrium.

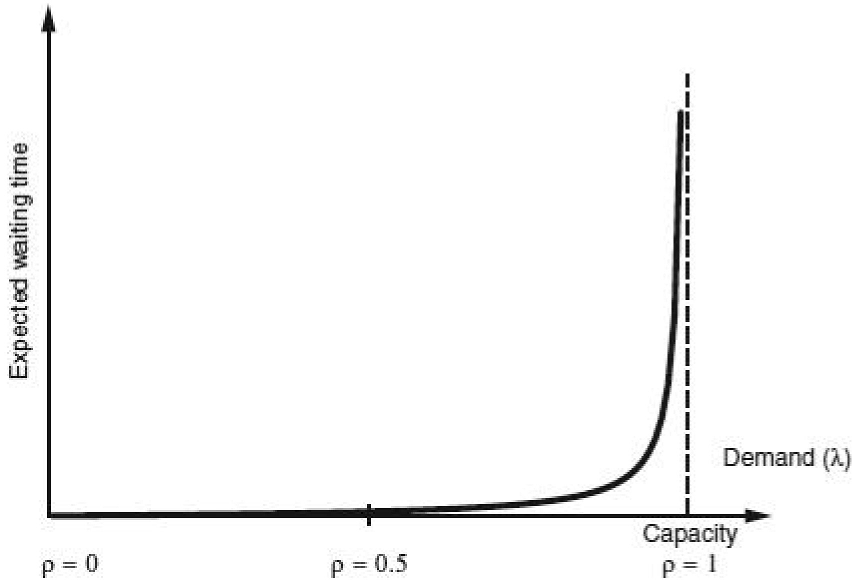


FIGURE 20.9 Expected time in queue as demand increases.

The exact mathematical expressions for $E[W]$, $E[W_q]$, $E[N]$, and $E[N_q]$ depend on the specifics of the queuing system under consideration. For example, consider a single-server system at which demands arrive at entirely random times according to a Poisson process. Interarrival times are thus described by the negative exponential probability distribution with parameter λ, the demand rate per unit of time. Suppose that the service rate is equal to μ and that the service time, S, has a variance equal to $\sigma^2(S)$. Finally, let this system have infinite queue capacity. The queuing system just described is known in queuing theory as an M/G/1 system.[11] It is important because of its many applications. For the M/G/1 system, it can be shown that

$$E[W_q] = \frac{\lambda \cdot [(\frac{1}{\mu})^2 + \sigma^2(S)]}{2 \cdot (1-\rho)} = \frac{\rho^2 + \lambda^2 \cdot \sigma^2\,(S)}{2\lambda \cdot (1-\rho)} \tag{20.10}$$

From Eq. (20.10), one can also compute $E[W]$, $E[N]$, and $E[N_q]$, using Eqs. (20.7 through 20.9), respectively. Note that all one needs to use Eq. (20.10) are the values of λ, μ, and the variance of the service time $\sigma^2(S)$. Note, as well the proportionality of $E[W_q]$ in Eq. (20.10) on $1/(1 - \rho)$. Equation (20.10) is valid only as long as $\lambda < \mu$, or equivalently, $(\lambda/\mu) = \rho < 1$. Example 20.3 illustrates through Eq. (20.10) the sensitivity of delay to even small changes in the demand rate when ρ is close to 1.

Figure 20.10 presents another example that compares the values of $E[N_q]$ for two different M/G/1 queuing systems, A and B. Both systems have $\mu = 60$ per hour, so that the expected service time $1/\mu$ is equal to 1 minute. However, system A has deterministic service times, meaning that $\sigma^2(S)$ is equal to zero, while system B exhibits significant variability of service times with a standard deviation equal to 0.9 minute [or $\sigma^2(S) = 0.81$ min^2]. While the overall shape of the two curves for $E[N_q]$ in Fig. 20.10 is dominated by the $1/(1 - \rho)$ term, the values that the two expressions take as a function of ρ differ, with $E[N_q]$ for the high-variability system B increasing faster. In general, the higher the variability (as measured by the variance) of the demand interarrival times and of the service times, the faster $E[W_q]$ and $E[N_q]$ increase as λ and ρ increase or as μ decreases.

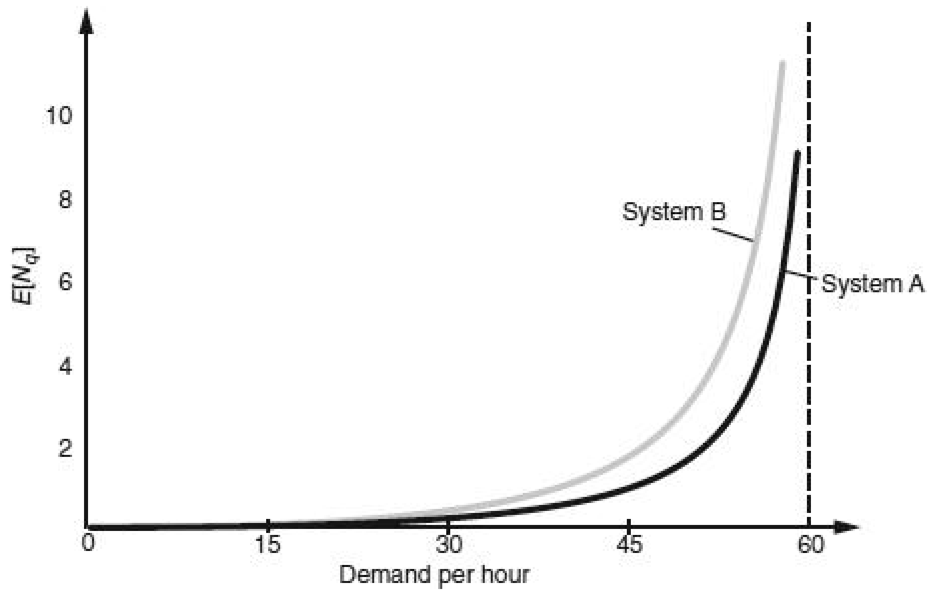


FIGURE 20.10 Expected number in queue as demand increases for two different queuing systems.

Equation (20.10) is valid only if the demand rate and the service rate are constant over time. Queuing theory provides few closed-form expressions, such as Eq. (20.10), for cases in which λ and μ vary with time, as usually happens at airports due to demand peaks, weather changes, etc. For these time-varying systems, $E[W]$, $E[W_q]$, $E[N]$, and $E[N_q]$ can only be estimated through numerical techniques or through computer-based simulation. However, as long as the queuing system reaches a long-term equilibrium, the observation that the long-term expected values $E[W]$, $E[W_q]$, $E[N]$, and $E[N_q]$ are proportional to $1/(1 - \rho)$ holds true, even if the demand rate and the service rate are functions of time. In such cases, one should be careful to interpret λ, μ, and ρ as the long-term averages of the demand rate,[12] the long-term average of the service rate (capacity), and the long-term average utilization of the system.

Example 20.3 Consider a single-server airport facility with a capacity of approximately 48 per hour (μ = 48 per hour). (This, e.g., could be a runway used for both arrivals and departures, where it takes 75 seconds on average between successive operations.) Assume that the service times at this facility can be approximated by a random variable whose expected value is 75 seconds and whose standard deviation is 25 seconds. Demands at this facility occur at a steady rate throughout the busy hours of the day (e.g., for 15 or 16 hours). Demand occurrences can be reasonably approximated as Poisson, so that demand interarrival times have an approximately negative exponential probability distribution.

These assumptions are consistent with the description of the M/G/1 queuing system. Therefore Eq. (20.10) can be used to compute $E[W_q]$. Consider the case where $\lambda = 36$ demands per hour. Applying Eq. (20.10) with[13] $\lambda = 36/3600 = 0.01$ demands per second, $1/\mu = 75$ seconds, $\rho = 36/48 = 0.75$, and $\sigma^2(S) = (25)^2$, gives $E[W_q] = 125$ seconds. One then obtains $E[W] = 125 + 75 = 200$ seconds [from Eq. (20.9)]; $E[N_q] = (0.01)(125) = 1.25$ aircraft in queue [from Eq. (20.8)]; and $E[N] = (0.01)(200) = 2$ aircraft in the system [from Eq. (20.7)].

Table 20.1 shows the values of the expected length of the queue and the waiting times, for several values of the demand rate. It also shows the changes resulting from a 1 percent increase in the demand rate, λ, at the 0.625, 0.75, 0.875, and 0.9375 levels of utilization.

		Expected Number in Queue		Expected Waiting Time	
Arrival Rate λ (per hour)	Service Rate ρ	$E[N_q]$	$E[N_q]$ (% change)	$E[W_q]$ (seconds)	$E[W_q]$ (% change)
30	0.625	0.58		69	
30.3	0.63125	0.60	3.4	71	2.9
36	0.75	1.25		125	
36.36	0.7575	1.31	4.8	130	4.0
42	0.875	3.40		292	
42.42	0.88375	3.73	9.7	317	8.6
45	0.9375	7.81		625	
45.45	0.946875	9.38	20.1	743	18.9

TABLE 20.1 The Indicative Approximate Values for the Expected Length of the Queue and the Waiting Times, as Derived from Queuing Theory

This model may approximate reality only roughly. Table 20.1 nonetheless underscores several points. Note, for example, that a 1 percent increase in demand at the 87.5 percent level of system utilization results in an almost 10 percent increase in expected queue length and 8.6 percent increase in delay. These percentages jump to roughly the 20 percent level when the system operates at about 94 percent of its capacity in the long run. Table 20.1 also suggests that the expected delay reaches the 4-minute level when the demand rate, λ, is equal to roughly 41 per hour. If this system were a runway, its practical hourly capacity (PHCAP) would then be about 41 aircraft per hour and it is reached at about the 85 percent utilization level (= 41/48)—see also Chap. 10.

20.7 Policy Implications

The observations of the last section and the "generic shape" of $E[W]$, $E[W_q]$, $E[N]$, and $E[N_q]$ shown in Figs. 20.9 and 20.10 have important implications for airports at the *policy level*. First, they provide a warning to airport operators, airlines, and civil aviation managers not to operate airport facilities and services at levels of utilization that are close to 1 over an extended period of time. Doing so risks having long delays, long waiting lines, and a poor LOS.

Moreover, queuing theory has also shown that not only do the expected values of W, W_q, N, and N_q increase in proportion to $1/(1-\rho)$, but so also do their standard deviations, $\sigma(W)$, $\sigma(W_q)$, $\sigma(N)$, and $\sigma(N_q)$. This means that, when ρ is close to 1, a queuing system not only experiences serious congestion, but it is also subject to great variability. Under the same set of *a priori* conditions (i.e., for the same λ, μ, and probability distributions for demand interarrival times and service times), delays on a particular day may be modest and tolerable and, on the following day, extremely long and unacceptable. This is a phenomenon observed very often at the busiest airports throughout the world.

It is difficult to make any precise general statements about the utilization ratio at which an airport facility or service should ideally be operated. The most appropriate value depends both on the particular operating characteristics of the system (probability distribution of service times and demand interarrival times, variability over time of the demand rate and the service rate, number of servers, etc.) and on the measures of performance considered most important (economic and other perceived costs of delay times, cost of the queuing system when idle, emphasis on avoiding extreme delays, etc.).

In most cases, it is fair to say that any facility or service operating at the range of 80 to 95 percent of its capacity for the duration of the consecutive active traffic hours of the day is near the "danger zone," or already in it, as far as serious delays are concerned. A long-term utilization ratio of more than 0.9 [when $1/(1-\rho)$ is ≥ 10] usually means long delays, low LOS on many days, and unstable conditions. These reflect the large expected value and standard deviation of waiting times and queue lengths. The reference to *active* traffic hours in the preceding text should be noted. The consecutive hours of truly active traffic in a day are considerably fewer than 24 at the great majority of airports. Moreover, certain facilities or services may be utilized for only some of those hours.

Another major point at the policy level is that when a queuing system operates at high levels of utilization, small changes in demand or capacity can cause large changes in delays and queue lengths. This simple practical observation is a direct consequence of the proportionality of both the expected value and the standard deviation of W_q and N_q to $1/(1 - \rho)$. It motivates much that is being done today at major airports around the world. Many initiatives are aimed at either managing/controlling demand (see Chap. 12) or at increasing the air-side and landside capacity of these airports (see Chaps. 13 and 16). Airport operators generally recognize that most of these initiatives will only produce small changes in demand or capacity. However, because many facilities and services at busy airports operate at very high utilization ratios, airport operators can reasonably expect that these small changes will produce significant reductions in delay. The reductions may be sufficient to maintain acceptable levels of service for a few additional years until more dramatic improvements in capacity might be achieved.

Exercises

20.1. Show that the area between the cumulative diagrams in Fig.20.6 is equal to the area of the triangle in Fig.20.5. Equation(20.4) gives both areas.

20.2. Demand for the runway system at an airport is 90 movements per hour throughout the busy hours of the day, except for the period 10:00–12:00 when it is 70 movements per hour. Suppose that, on a given day, the airport capacity was 100 movements per hour until 7a.m. However, due to a weather front, the capacity was only 60 movements per hour between 7 a.m. and 9 a.m. From 9 a.m. to 11 a.m. the capacity increased to 80 movements per hour and, finally, at 11 a.m., it went back to 100 movements per hour, where it stayed for the rest of the day.

Under the usual assumptions described in Sec. 20.5, draw carefully the cumulative diagram for the number of demands as a function of time and the number of aircraft “admitted” for service at the runway system as a function of time. Begin your picture at 6 a.m. What is the longest delay suffered by any aircraft during this day? What is the total delay suffered by all aircraft during that day?

20.3. Consider an airport with a runway used exclusively for landings during peak traffic hours. Under such peak conditions, the arrivals of airplanes at the vicinity of the airport can be assumed to be approximately Poisson with a rate $\lambda = 55$ aircraft per hour. Of these airplanes, 40 on average are commercial jets and 15 are small general aviation and commuter airplanes. The probability density function for the duration of the service time, S, to a random aircraft landing on the runway is uniformly distributed between 48 and 72 seconds. Peak traffic conditions occur during 1000 hours per year, and the average cost of 1 minute airborne waiting time (i.e., of time spent in the air while waiting to land) is $60 for commercial jets. (This accounts for additional fuel burn, extra flight crew time, and other variable operating costs.) Estimate the yearly costs to the airlines of peak traffic conditions. Assume that Eq. (20.10) for estimating waiting time is valid for this case.

References

Ashford, N. (1988) "Level of Service Design Concept for Airport Passenger Terminals—A European View," *Transportation Planning and Technology, 12*(1), pp. 5–21.

Gross, D., Shortle, J., Thompson, J., and Harris, C. (2008) *Fundamentals of Queuing Theory*, 4th ed.,Wiley, New York.

Hillier, F. and Lieberman, G. (2009) *Introduction to Operations Research*, 9th ed., McGraw-Hill, New York.

Larson, R. C. (1987) "Perspectives on Queues: Social Justice and the Psychology of Queuing," *Operations Research, 35*, pp. 885–905.

Larson, R. C. (1988) "There Is More to a Line than Its Wait," *Technology Review*, pp. 60–67.

Larson, R. C. and Odoni, A. R. (2007) *Urban Operations Research*, Dynamic Ideas, Belmont, MA.

Little, J. (1961) "A Proof of the Queuing Formula $L = \lambda W$," *Operations Research, 9*, pp. 383–387.

Wolff, R. (1989) *Stochastic Modeling and the Theory of Queues*, Prentice-Hall, Englewood Cliffs, NJ.

---

[1] The symbols in this chapter are consistent with standard notation used in queuing theory.

[2] Airlines try, not always successfully, to make sure that the bags of their prized first-class and business-class customers are the ones to appear first at the carousel.

[3] Each of the desks may, in fact, provide check-in service for more than one flight—and possibly all of the flights of an airline.

[4] Schedules of aircraft, flight crews, etc. are said to have a lot of "padding" in such cases.

[5] This assumes that the delay caused by airport congestion, W_q, is the only source of delay. Unfortunately, this is usually not true.

[6] At a more technical level we shall also assume that the lengths of successive demand interarrival times are (statistically) independent and so are the durations of successive service times.

[7] "Stochastic" is synonymous with "probabilistic"; we use the term here because it is also widely used in queuing theory.

[8]In fact, the cumulative demand diagram simply plots the value of $\int_0^t \lambda(x) \times dx$ for all values of *t*.

[9]If a FCFS priority discipline is not in effect, then the horizontal distance is the time that elapses between the *i*-th demand for service and the *i*-th admission for service; note that this may involve two different users.

[10]We have not discussed what happens when $\rho = 1$. It can be proved that queuing systems do not reach equilibrium (i.e., queues and waiting time go to infinity) if $\rho = 1$; the only exception is the very special case when both the demand interarrival times and the service times are constant with $\lambda = \mu$.

[11]The code "M/G/1" means that the demand is Poisson (indicated by the letter "M" for "memoryless"), the service times can have any probability distribution (indicated by the letter "G" for "general"), and the system has one server.

[12]More formally, we compute λ as $\left(\int_0^t \lambda(x) \cdot dx\right) / T, \ as\ T \rightarrow \infty$ and similarly for μ.

[13]Note the importance of using consistent units of time for all the parameters; in this case "seconds" is the time unit used.

CHAPTER 21

Peak-Day and Peak-Hour Analysis

Airport facilities are typically designed to accommodate loads during a design peak day (DPD) or a design peak hour (DPH), depending on the type of load at hand. Many alternative definitions of DPD and DPH are used in practice. All share a common characteristic: They specify a level of traffic that is exceeded only rarely during the target period. The intent is to ensure that airport facilities are designed with adequate capacity to handle demand at a desired level of service practically throughout the year, while not being overdesigned just to handle a few instances when extreme peaks may occur.

To estimate DPD and DPH loads, planners must review carefully historical data to understand seasonal, monthly, daily, and hourly peaking patterns at the airport under consideration. They must use judgment in assessing how these patterns will change in the future. Demand peaking at airports usually becomes less acute as traffic grows. Planners must exercise additional care in distinguishing between the peaking characteristics of passengers versus those of air traffic, as well as of arriving passengers versus departing passengers.

Annual demand forecasts can be converted into DPD and DPH traffic estimates through the application of conversion coefficients. These coefficients can be inferred from historical data and from experience with peaking patterns at airports of various sizes but always require judgmental inputs from planners. They usually provide good first-order approximations, but should never be used uncritically.

21.1 Introduction

Much of airport planning and design revolves around the notion of the DPD and, especially, of the DPH—also sometimes referred to as the *typical peak day* (TPD) or *typical peak hour* (TPH) or simply as the *design day* or *design hour*.

The capacity of airport facilities is often specified with reference to *annual* volumes of traffic. However, this is only because thinking in terms of annual totals is consistent with the way demand forecasts are specified—almost always in annual terms. One can readily compare, for instance, the capacity of a passenger building designed to handle 20 million passengers per year with a forecast of annual passenger demand in a particular year to determine whether the building will be adequate to handle traffic that year. The same is true when it comes to planning for other parts of the airport, such as cargo facilities or the airfield—in which case the relevant demand forecast is the annual number of aircraft movements.

However, any comparisons between estimated annual demand and annual capacity are very approximate (see Chap. 11 regarding estimating annual capacities) and useful only for purposes of "macroscopic" planning. Hourly and daily figures are typically far more important for the purpose of detailed design. It is short-term loads that determine the required size of a facility, the number of servers in each of its constituent parts, etc. For example, in designing an airfield, one would like to know the expected number of DPH movements. If, for instance, that number is 40, a single runway will probably suffice, but if it is 65, a two-runway system will be necessary (see Chap. 10). Similarly, as indicated in Chap. 15, space requirements in each part of a passenger building are determined with reference to the number of simultaneous occupants during the peak *hours* of the year. The number of required processing units (check-in desks, passport control desks, security-check machines, etc.) is also determined by the flows in the building during these hours. Likewise, when wishing to determine whether a group of contact and remote stands at an airport has adequate capacity, one typically works with a scenario involving a *daily* schedule of arriving and departing aircraft and must necessarily plan for a DPD when the number of movements is high.

This chapter reviews the estimation of flows during DPHs and DPDs and thus has a bearing on several other topics in this book. It first discusses the many existing alternative definitions of the DPD and the DPH along with a common condition that all these definitions should satisfy. It then describes a simple approximate process for converting annual forecasts of traffic into DPH and DPD forecasts. The final two sections address two specific questions that often come up in practice. The first concerns estimating the number of DPH aircraft movements; the second deals with the number of DPH arriving passengers and of DPH departing passengers.

21.2 Definitions of the DPD and DPH

Many alternative definitions of DPD and DPH are in use. The following is a partial list of possibilities:
For DPD:

1. The 10th, 15th, or 30th busiest day of the year
2. The average day of the peak month of the year (ADPM)[1]
3. The 90th or 95th percentile busiest day of the year, that is, a day whose traffic load is exceeded by only 36 or 18 other days in the year

For DPH:

1. The 20th, 30th, or 40th busiest hour of the year
2. The peak hour of the average day of the peak month (ADPM) of the year
3. The peak hour of the average day of the 2 peak months of the year
4. The peak hour of the 90th or 95th percentile busiest day of the year
5. The peak hour of the 7th or 15th busiest day of the year
6. The peak hour of the 2nd busiest day during the average week in a peak month
7. The "5 percent busy hour," that is, an hour selected so that all the hours of the year that are busier handle a cumulative total of 5 percent of annual traffic

All of these definitions have been used at times or been recommended by various organizations. For example, it has been standard practice to use Definition 1 for DPH in the United Kingdom, specifically, the 30th busiest hour of the year. The corresponding level of traffic is called the *standard busy rate*. In the United States, Definitions 2 for DPD and for DPH are often used (FAA, 1988), while the International Civil Aviation Organization (ICAO, 2000) has recommended Definition 3 for DPH.

For practical purposes, it makes little difference which definitions are used as long as they fulfill the following condition: The DPD and DPH should not be the day or hour, respectively, of the year with the highest traffic demand but one with a demand that is exceeded only on a reasonably small number of instances during the year. "Reasonably small" may mean something like 10 to 30 days or 20 to 50 hours, depending on the context of the planning exercise and on the intensity of demand peaking at the airport of interest. This condition is intended to ensure that airport facilities will have adequate capacity to handle demand practically throughout the year, while not being overdesigned just to handle a few instances when extreme peaks may occur. Such extreme peaks may be associated with a few days each year when traffic is exceptionally heavy (e.g., a holiday period or an annual religious pilgrimage) or with certain special events (e.g., the annual Super Bowl game in the United States or the Olympic games somewhere in the world every 4 years). This approach to selecting the DPD and DPH implicitly recognizes that some deterioration of the level of service during a few hours or days of extremely high demand should be tolerated in the interest of reducing overall capital and operating costs.

It can be seen that all the definitions listed at the beginning of this section satisfy this condition. Moreover, the differences among the estimates of DPD and DPH demand that are obtained from these alternative definitions are typically insignificant from the practical viewpoint. Arguing whether one should select the 30th or 40th busiest hour of the year is quite meaningless given (1) the uncertainty in forecasts (see Chaps. 4 and 19), especially when the target date for which the facility is being designed lies 10 or 20 years in the future, (2) the use of judgmental inputs in estimating DPD and DPH loads (see Sec. 21.3), and (3) the many simplifying assumptions and approximations used in all design methodologies.

Airport planners and designers should thus feel free (unless restricted to adhere to local practice, as is the case with the use of the "standard busy rate" in the United Kingdom) to select any one of the definitions above, as appropriate to the case at hand and the available data. Note that DPD Definition 2 and DPH Definition 2 are often the least demanding, in terms of data, as they require detailed hour-by-hour information only for the peak month of the year.

21.3 Conversion of Annual Forecasts into DPD and DPH Forecasts

It is usually the case that planners obtain DPD and DPH forecasts—whether for passengers or aircraft movements or some other measure of demand for an airport facility—in a "top-down" fashion from annual forecasts. In other words, one typically begins from a given annual forecast and obtains a DPD or DPH forecast by applying appropriate "conversion coefficients," based on historical data that are often adjusted judgmentally. The reason for starting with annual forecasts is that these are in most cases the only forecasts available. Airport forecasting methodologies (see Chap. 19) are, by nature, strongly oriented toward predicting annual figures of demand, as a function of various independent variables, or on the basis of trend analyses, or through a traffic-share analysis, etc.

The overall process can be summarized in the following steps:

1. Review traffic data from the airport of interest to determine the appropriate definition of DPD or DPH (or use a recommended or required definition, if specified).
2. Identify a day and an hour during a very recent year that roughly correspond to the definitions of DPD and DPH, respectively, that were selected in Step 1; for the DPD assemble detailed data on that day's operational profile.
3. Compute the ratio of the total traffic load of the DPD and the DPH in the selected year to the annual traffic loads of the airport during that year and call these the "current DPD conversion coefficient" and "current DPH conversion coefficient," respectively.
4. For the target future year, multiply the forecast annual traffic load by the "current DPD conversion coefficient" and "current DPH conversion coefficient" to obtain

preliminary estimates of the DPD and DPH traffic loads for that year; also, scale appropriately the DPD's operational profile obtained in Step 2.

5. Adjust the DPD and DPH loads estimated in Step 4 by using judgment and data from other airports whose current annual traffic loads are similar to the forecast annual load of the airport of interest during the target year. Likewise, adjust (e.g., through schedule "smoothing," also known as "peak spreading") the scaled operational profile computed in Step 4 for the DPD.

As can be seen, Steps 1 through 4 are quite mechanistic, as long as good historical data and a traffic forecast are available. In recent years, practically every major airport in the world has been developing databases with detailed historical information on traffic activity, such as scheduled and actual arrival and departure times of flights, number of arriving or departing passengers per flight, cargo volumes, etc. It is easy therefore to identify a DPD (e.g., fourth Friday of July) and a DPH (e.g., peak hour of ADPM, the average day of the peak month), as well as to develop operational profiles for the DPD, such as the ones shown in Fig. 21.1*c* for the number of movements per hour in February and August 2010 at New York/Newark.

To illustrate Step 3, consider a large airport that handled 370,000 arrivals and departures in 2009 and had 1250 scheduled aircraft movements on what has been defined as the DPD and 87 scheduled movements in what has been defined as the DPH. Then, the 2009 DPD conversion coefficient for this airport is 1250/370,000 ≈ 0.0034 (i.e., the traffic load, in units of scheduled aircraft movements, on the DPD amounted to about 0.34 percent of the annual traffic volume). Likewise, the 2009 DPH conversion coefficient is 87/370,000 ≈ 0.000235.

Step 4 is equally simple, given a forecast of future traffic loads. In the previous example, a forecast of 440,000 annual aircraft movements in 2019 would result in a DPD of about 1500 aircraft movements (≈ 440,000 × 0.0034) and a DPH of about 103 movements (≈ 440,000 × 0.000235). Moreover, the number of movements in each hour of the DPD in 2019 would be scaled up by about 19 percent (440,000/370,000 ≈ 1.19) from the number for the same hour in 2009.

Step 5, however, requires good judgment and deserves significant research and effort, as it is also very important from the practical viewpoint. In carrying out Step 5 it is important to bear in mind that, in general, peaking patterns at an airport are usually quite stable from year to year, but peaking tends to become less acute as traffic increases. This empirical observation is discussed further, after Example 21.1.

Example 21.1 Suppose an airport that today handles 12 million passengers per year is forecast to have a demand of 18 million 10 years from now, a 50 percent increase. Suppose also that local practice defines the 30th busiest hour of the year as the DPH and, based on the current year's traffic data, the number of arriving and departing passengers during that hour is 4500. Thus, the fraction of annual passenger traffic processed during the DPH, that is, the current DPH conversion coefficient, is equal to 0.000375 (= 4500/12,000,000) or 0.0375 percent. This would suggest a projected 6750 DPH passengers 10 years from now [= (0.000375) × (18 million) = (1.5) × (4500)]. However, planners should probably adjust this estimate. For example, by looking at the historical evolution of peaking patterns at the airport, comparing with other airports that now process about 18 million passengers per year, and speculating on the most likely ways peaking patterns might change, planners may decide to reduce the conversion coefficient to account for a reduction in peaking as the annual traffic increases. A value such as 0.00033 may be deemed more appropriate for use with the 18-million-passenger level, resulting in an estimate of about 6000 DPH passengers. Plausibility checks should also be performed. For instance, will the runway system be capable of handling aircraft operations consistent with a rate of 6750 (or 6000) passengers per hour, given the projected mix of aircraft types at the airport?

As noted previously, peaking patterns tend to be quite stable over the years at busy airports. This is illustrated in Figs. 21.1*a* through 21.1*d* for New York City's airports. Figures 21.1*a* and 21.1*b* refer to *monthly* peaking. Figure 21.1*a* shows the total number of aircraft movements in 2007 and 2010 at the three airports that serve the New York City metropolitan area, whereas Fig. 21.1*b* shows the number of passengers at New York/Kennedy in 2005 and 2011. Note that while the volume of traffic may have changed considerably, the peaking patterns remain remarkably consistent both for aircraft movements and for passengers. In a similar fashion, the *hourly* peaking patterns, as reflected in the daily operational profile for air traffic movements at New York/JFK in February 2010, the lowest month of the year, and in August 2010, the highest month, closely follow each other (see Fig. 21.1*c*), despite the fact that the total demand in August was over 30 percent greater than in February. Figure 21.1*d* provides a somewhat different perspective. It shows the daily operational profile of aircraft movements at New York/Newark in 2007 and 2010. Note that the vertical axis is now the percent of daily movements scheduled for any particular hour during an average day and that the two profiles are again very similar.

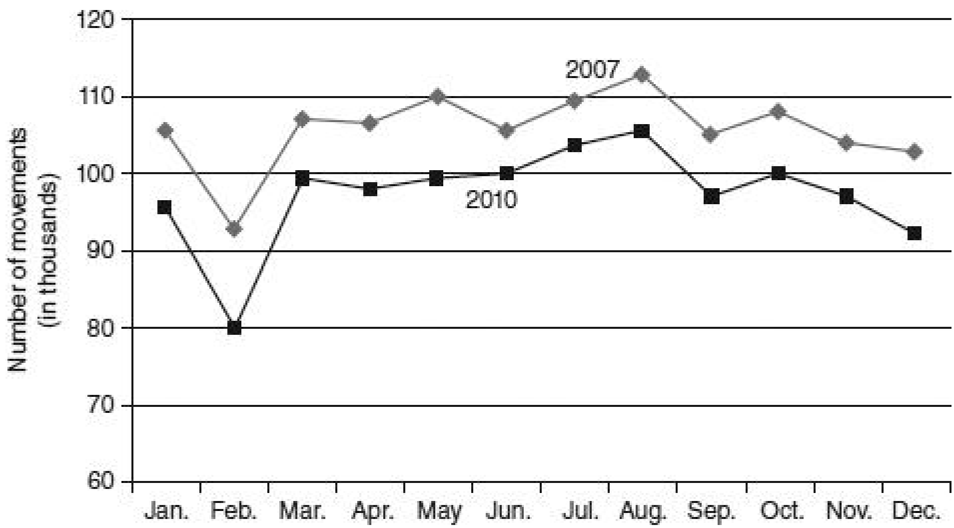


FIGURE 21.1*a* Total monthly runway traffic at the three New York City airports in 2007 and 2010; because of the financial crisis and other factors, total traffic in 2010 was lower than in 2007.

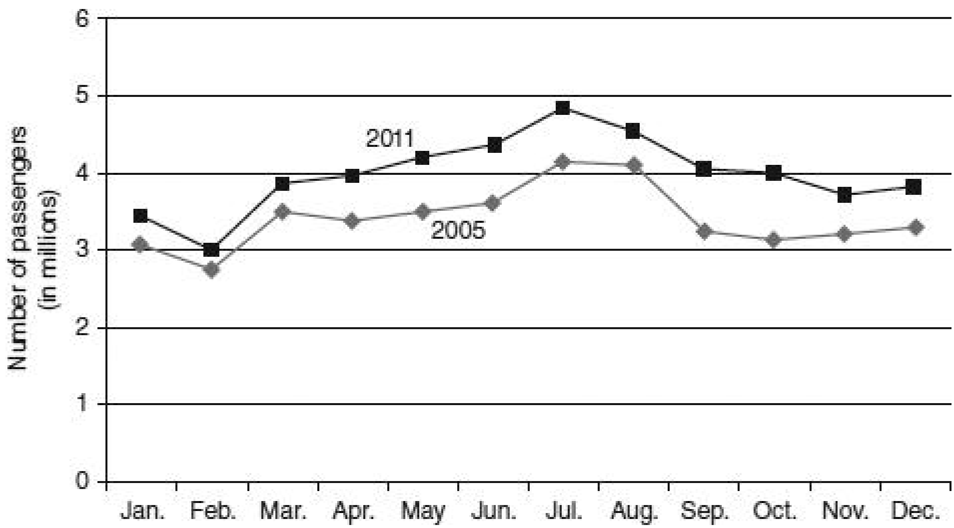


FIGURE 21.1*b* Monthly passenger traffic at New York/Kennedy in 2005 and 2011.

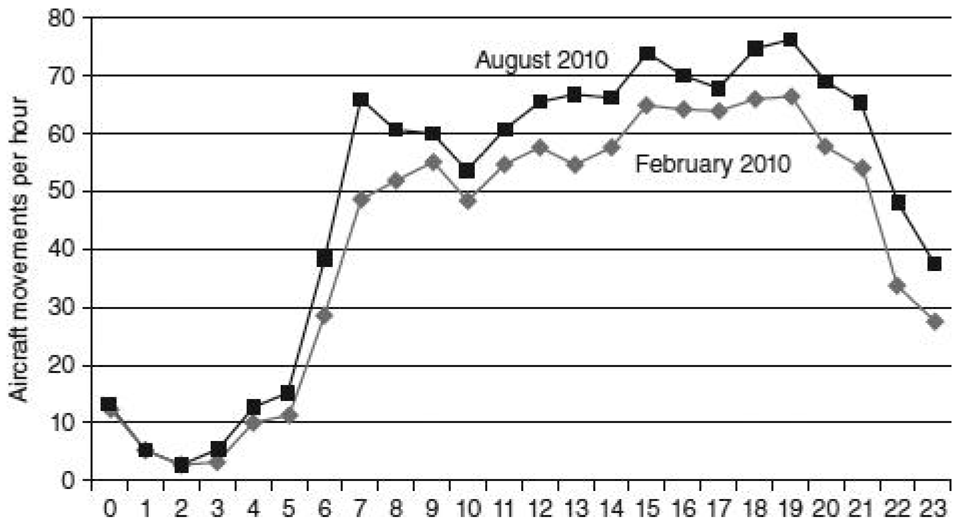


Figure 21.1*c* Similarity of 24-hour profiles of aircraft movements on an average weekday in February and in August 2010 at New York/Newark.

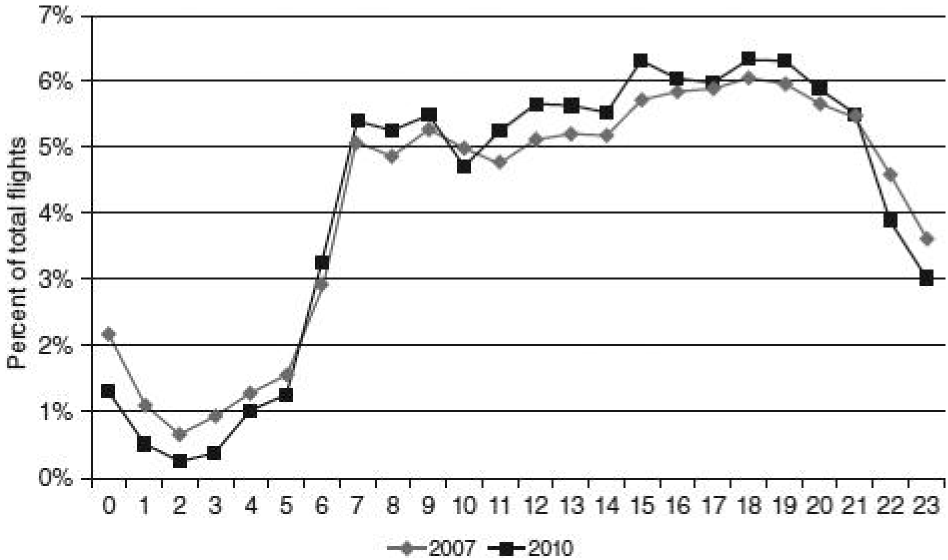


Figure 21.1*d* Hourly peaking patterns at New York/Newark in 2007 and in 2010; the vertical axis indicates the percent of total daily movements taking place in each hour.

However, as the annual totals grow, the peaking at any given airport may become *less acute* over the years, because the distribution of the number of aircraft movements and of passengers, across both the months of the year and the hours of the day, often tends to become progressively "flatter." Figure 21.1*d* is consistent with this observation: There were about 10 percent more movements per day, on average in 2007 than in 2010 at New York/Newark and, as the figure shows, the 2007 daily operational profile is somewhat flatter and less peaked than the one for 2010. There are several good reasons for this tendency. First, in competitive markets, especially short-and medium-haul ones, airlines use flight frequency as a competitive weapon. Typically, an airline would begin by scheduling a flight during the morning peak hour and another during the early evening peak hour for the markets it serves. However, as passenger volume increases, the airline is more likely to add a third flight in mid-day and even a late night flight to capture more passengers, instead of increasing the size of the aircraft that serve the peak hours. This, of course, results in a less peaked operational profile at the airport. A second reason is that traffic congestion during peak hours—on the runways, but also in terminal buildings—may increase the attractiveness, for scheduling purposes, of off-peak periods. In addition, slot unavailability at schedule-coordinated airports (see Chap. 12) may also force airlines to schedule flights outside peak hours and to even out the operational profile at some of these airports. For instance, the number of scheduled movements per hour is almost constant throughout a 15- or 16-hour stretch of each weekday at both London/Heathrow and Frankfurt/International. The FAA's scheduling limits have a similar effect at New York/LaGuardia (see Chap. 12).

Several other factors may also play a role vis-à-vis de-peaking (or more peaking) of airport operations. For example, New York/Kennedy used to be strongly oriented toward serving international flights to/from Europe and thus had few runway movements in the morning and a prolonged extremely busy period of international arrivals and departures between 2 p.m. and 10 p.m. in the evening. However, demand for runway operations became significantly more balanced throughout the day after the airport became the hub of JetBlue, a major domestic carrier. On the other hand, several possibilities may contribute to intensifying some traffic load peaks. For one, long-haul flights in some major intercontinental markets have "natural time-windows" within which they must operate. For instance, most flights from Europe to North America arrive between 2 p.m. and 8 p.m. and depart for Europe between 5 p.m. and 11 p.m., thus causing increased loads, especially on the passenger side. Another example is the introduction of very large aircraft, such as the A380, in certain markets. The arrival of such aircraft at an airport obviously generates sharp "spikes" in the loads of terminal buildings.

In summary, as total DPD loads increase, the peak period loads (as a percentage of total daily loads) are generally reduced and compensated for through increases in the off-peak loads. However, it is important to exercise caution when making such adjustments in DPD and DPH conversion coefficients. Two recent publications, ACRP Report 55 (ACRP, 2011) and, especially, ACRP Report 82 (ACRP, 2013) provide detailed guidance for preparing DPD and DPH estimates of facility loads. The latter also includes a set of spreadsheets that facilitate the five-step process we have outlined. The point that needs to be emphasized here is that these estimates inevitably require the application of judgmental considerations whenever they refer to time horizons of 5 or more years. They should therefore always be reviewed critically before they are used for planning and design.

Approximate "Default" Conversion Coefficients and Why They Work Well

The U.S. Federal Aviation Administration (FAA) long ago recommended a set of conversion coefficients, shown in Table 21.1, for estimating approximately the number of DPH passengers[2] from annual forecasts of demand (FAA, 1969). Experience indicates that these coefficients work very well in practice, *as first-order approximations*. Their use is recommended when a rough "figure of merit" is needed quickly or in the absence of detailed historical data on an airport's peaking patterns. Table 21.2 offers further insight into the peaking characteristics of airports. It summarizes the statistics on the peak-month traffic obtained from an Airports Council International (ACI, 1998) survey of 80 airports worldwide.[3] For example, for the 23 airports in the survey that had more than 20 million annual passengers, the *monthly peaking ratio*, that is, the "average number of passengers per day during the peak month of the year" divided by the "average number of passengers per day during the entire year," ranged between 1.09 and 1.43, with an average value of 1.18. Stated differently, average daily passenger traffic during the peak month was 9 to 43 percent greater than that during the entire year[4]—and was 18 percent greater when averaged over all 23 airports. In the case of 6 of these 23 airports (26 percent), the average daily traffic during the peak month was 20 percent greater than the average during the entire year.

Total Annual Passengers (millions)	DPH Passengers as % of Total Annual Passengers
20	0.03
10–20	0.035
1–10	0.04
0.5–1	0.05

Source: FAA, 1969.

TABLE 21.1 Conversion Coefficients for Estimating Number of DPH Passengers from Annual Figures

Total Annual Passengers (millions)	Sample Size	Monthly Peaking Ratio*		
		Average	Range	Greater than 1.2
> 20	23	1.18	1.09–1.43	6 of 23 (26%)
10–20	13	1.25	1.08–1.55	9 of 13 (69%)
1–10	44	1.35	1.11–1.89	34 of 44 (77%)

*Monthly peaking ratio = (average number of passengers per day during peak month)/(average number of passengers per day during entire year).

TABLE 21.2 Monthly Peaking Characteristics of the 80 Airports in the ACI Survey

The third and fifth columns of Table 21.2 suggest that what is true for individual airports ("monthly peaking tends to become less acute as traffic increases") is also generally true across airports. However, the wide range of values in the fourth column underlines the importance of utilizing, whenever possible, data specific to the airport of interest. In general, airports with primarily domestic flights and those serving large numbers of business passengers will have relatively low monthly peaking ratios, in contrast to those serving mostly vacation and pleasure travelers.

The three New York airports offer an excellent case in point. LaGuardia, with a large volume of domestic business traffic, had a monthly peaking ratio of only 1.082 in 2011. Newark, serving primarily domestic traffic and significant international traffic, had a higher monthly peaking ratio of 1.177. Finally, Kennedy, whose passenger traffic in 2011 was almost exactly equally divided between domestic and international had a 1.193 ratio. Interestingly, in 1998, the ratio for Kennedy was significantly greater at 1.252, consistent with the fact that the airport served only about 31 million passengers that year, as opposed to 47.7 million in 2011. The rapid growth of JetBlue's domestic traffic at the airport also contributed to a smoother seasonal (as well as daily) pattern of demand.

Analogously, Table 21.3 shows ranges of values for the *hourly* peaking of air traffic movements for a sample of U.S. airports. For instance, at airports with more than 20 million passengers a year, the number of (arriving or departing) aircraft seats during the peak hour of a typical day constitutes anywhere between 7 and 10 percent of all aircraft seats available at the airport during that day.

Total Annual Passengers (millions)	Peak-Hour Available Seats as % of Total Daily Number of Available Seats
>20	7–10
10–20	8–12
1–10	9–20

TABLE 21.3 Hourly Peaking Characteristics of Some U.S. Airports

Tables 21.2 and 21.3 suggest why the conversion coefficients in Table 21.1 usually provide reasonable DPH estimates. Consider, for example, an airport with more than 20 million annual passengers. Selecting the values of 1.18 from column 3 of Table 21.2, and 9 percent from the range of 7 to 10 percent in column 2 of Table 21.3 gives

$$\text{Conversion coefficient for airports with more than 20 million annual passengers} = \left(\frac{1}{365}\right) \times (1.18) \times (0.09) \approx 0.000291 \text{ or } 0.0291 \text{ percent}$$

This is within 4 percent of the 0.03 percent value suggested by Table 21.1. Similarly, for an airport with 10 to 20 million annual passengers, one obtains

$$\left(\frac{1}{365}\right) \times (1.25) \times (0.10) \approx 0.00034 \text{ or } 0.034 \text{ percent}$$

which is within just 3 percent of the 0.035 percent in Table 21.1. Finally, for 1 to 10 million annual passengers, a typical estimate might be

$$\left(\frac{1}{365}\right) \times (1.35) \times (0.12) \approx 0.000444 \text{ or } 0.0444 \text{ percent}$$

or within about 11 percent of 0.04 percent.

These observations do not absolve planners from responsibility for seeking, whenever possible, peaking data specific to the airport of concern. As Tables 21.2 and 21.3 indicate, local monthly and hourly peaking characteristics span a considerable range of values. The resulting conversion coefficients may differ significantly between airports.

21.4 DPH Estimates of Aircraft Movements versus Passengers

Peaking characteristics for aircraft movements are generally quite similar to those for passengers, but with some noteworthy differences. For one, whereas the peak month for aircraft movements almost always coincides with that for passengers, this does not necessarily hold true for the peak hours.

More important, the seasonal and daily peaking of aircraft movements is usually somewhat less acute than the peaking for passengers. For example, for the three New York airports, the monthly peaking ratios for the number of aircraft movements in 2011 were 1.047, 1.072, and 1.117 at LaGuardia, Newark and Kennedy, respectively, all lower than the corresponding ratios of 1.082, 1.177, and 1.193, respectively, for the number of passengers. The reason is that passenger load factors are generally higher during the peak season of the year, as well as during the peak traffic hours of the day. In most cases, however, the differences are small. Thus, the magnitude and ranges of the conversion coefficients in Tables 21.1, and of the peaking characteristics in Tables 21.2 and 21.3 are, for practical purposes, nearly as reasonable for estimating DPH aircraft movements as DPH passengers. Airport planners are nonetheless advised to check carefully local peaking data, when it comes to aircraft movements, especially at those airports whose runway systems impose serious capacity constraints. In fact, as already noted, major airports that operate near the limit of their airside capacity for much of the day (e.g., New York/LaGuardia, Chicago/O'Hare, London/Heathrow, Frankfurt/International) are characterized by essentially flat profiles of aircraft movements for 10 to 16 hours of the day, because the number of movements is "capped" by the runway system's capacity. At such airports, the number of aircraft movements during the peak hours of the day is of the order of only about 6 to 6.5 percent of the total movements in the day, that is, at the lowest end of the ranges shown in Table 21.3.[5] For such runway capacity-constrained airports, the simple FAA conversion coefficients of Table 21.1 may result in estimates of DPH aircraft movements considerably higher than the available runway capacity.

21.5 DPH Estimates of Flows of Arriving Passengers and of Departing Passengers

A second related issue concerns DPH estimates specific to the number of peak-hour *departing* passengers and peak-hour *arriving* passengers. Such estimates are necessary for the planning and design of passenger buildings and other landside areas, where many facilities and spaces are dedicated to either serving arrivals only or departures only. Two observations are useful in this respect.

First, the peak hours for departing passengers, for arriving passengers, and for total passengers need not coincide. In fact, it is not unusual that they occur at three different hours of the day. As well, peaking will probably be more acute for the arriving or the departing flows alone than for total passengers. For example, the scenario of Table 21.4 is entirely plausible. It is therefore necessary to collect data specific to the *hourly* peaking pattern of arriving passengers only and of departing passengers only.[6]

Passenger Flow	Peak Hour of Day	Peak-Hour Flow as % of Total Daily Flow
Total	16:00–16:59	8
Departing	07:00–07:59	10
Arriving	18:00–18:59	9

TABLE 21.4 A Plausible Hourly Peaking Pattern at a Major Airport

Second, the hourly peaking of arriving and departing passengers depends critically on the details of airline schedules. For this reason, arrival and departure peaking characteristics, considered separately, are less stable over the years than the peaking characteristics of overall traffic. Thus, predictions of DPH flows of departing passengers alone or of arriving passengers alone are less reliable, in general, than predictions of total DPH flows, especially when they concern target dates 5 or 10 years into the future. This is particularly true at hub airports, where the hourly peaking of arriving and of departing passengers can be expected to be very acute and may change rapidly as the hubbing airlines change their schedules and strategies.

Exercises

21.1. Explain why demand peaking at an airport usually becomes less acute as traffic grows. Consider seasonal peaking, day-of-the-week peaking, and hour-of-the-day peaking. How do airlines schedule additional flights on individual markets as traffic on these markets grows? Are there differences between short- and long-distance markets? What is the influence of airport capacity? What additional factors play a role in the gradual "de-peaking" of traffic?

21.2. For an airport with which you are familiar, prepare graphs and/or tables showing the evolution of traffic peaks on a monthly, day-of-the-week, and hour-of-the-day basis over a recent 10-year period. Compare, for example, the 1st, 5th, and 10th years of the period.

References

ACI, Airports Council International (1998) *Airport Capacity Demand Profiles: 1998 Edition*, ACI, Geneva, Switzerland.

ACRP, Airport Cooperative Research Program (2011) *Passenger Level of Service and Spatial Planning for Airport Terminals*, Report 55, Transportation Research Board, Washington, DC. www.trb.org.

ACRP, Airport Cooperative Research Program (2013) *Guidelines for Preparing Peak Period and Operational Profiles*, Report 82, Transportation Research Board, Washington, DC. www.trb.org.

FAA, Federal Aviation Administration (1969) *Aviation Demand and Airport Facility Requirement Forecasts for Medium Air Transportation Hubs through 1980*, FAA, Washington, DC.

FAA, Federal Aviation Administration (1988) *Planning and Design Guidelines for Airport Terminal Facilities*, Advisory Circular 150/5360-13, U.S. Government Printing Office, Washington, DC.

ICAO, International Civil Aviation Organization (2000) *Aerodrome Design Manual, Part 1, Runways*, 2d ed., 1984, reprinted July 2000 incorporating Amendment 1, Doc 9157, ICAO, Montreal, Canada.

---

[1]When this definition is used, it is often best to exclude days on which demand is consistently much lower than that on typical days. For example, many U.S. airports have much lower demand on Saturdays than all other days of the week. An alternative to ADPM could then be the average weekday of the peak month of the year (AWDPM).

[2]The FAA uses the term "typical peak hour passengers" (TPHP).

[3]This sample of airports provided seemingly reliable answers to the relevant question in the ACI survey. It is not clear how representative this sample was, as it included 38 European, 30 North American, 8 Asian, 2 African, 1 Australian, and 1 South American airports.

[4]The six airports in the sample with more than 35 million passengers (Atlanta, Denver/Metro, Frankfurt/International, London/Heathrow, Los Angeles/International, and San Francisco/International) all had very similar monthly peaking ratios (1.18, 1.19, 1.19, 1.14, 1.19, and 1.22, respectively).

[5]Note that if operations were exactly evenly distributed throughout the 24 hours of the day, each hour would have about 4.17 percent of the daily traffic. In the somewhat more likely event, that all traffic would be evenly distributed over an 18-hour span (with the other 6 hours having no traffic), each of the 18 hours would have 5.56 percent of daily traffic.

[6]Note, however, that the *monthly* peaking ratios for arriving and for departing passengers—see Table 21.2—will be virtually identical with the monthly peaking ratio for total passengers.

Acronyms and Symbols

Acronyms

AAR	Airport Acceptance Rate
AATF	Airport and Airway Trust Fund
ACA	Airport Carbon Accreditation
ACC	Area Control Centre (Europe)
ACCC	Australian Competition and Consumer Commission
ACES	Airspace Concept Evaluation System (a highly detailed network model)
ACI	Airport Council International
ACRP	Airport Cooperative Research Program (U.S.)
ACSA	Airports Company of South Africa
ADA	Americans with Disabilities Act
ADF	Anti-icing/De-icing Fluid *also* Airport Development Fund
ADL	Aggregate Demand List
ADMS	Atmospheric Dispersion Modelling System (U.K.)
AdP	Aéroports de Paris
ADPM	Average Day of the Peak Month of the year
ADS	Automatic Dependent Surveillance
AEDT	Aviation Environmental Design Tool (U.S. FAA)
AENA	Spanish Airports and Air Navigation (Aeropuertos Españoles y Navegación Aérea)
AFP	Airspace Flow Program
AIP	Airport Improvement Program (U.S.)
AIR-21	Aviation Investment and Reform Act for the 21st Century (U.S.)
AIRE	Atlantic Interoperability Initiative to Reduce Emissions
ALAQS	Airport Local Air Quality Studies (EUROCONTROL)

AMASS	Airport Movement Area Safety System
ANA	Airports company of Portugal (Aeroportos e Navegação Aérea)
ANCON	Aircraft Noise Contour model (U.K.)
AND	Airport Network Delays (a fast queuing network model)
ANSP	Air Navigation Service Provider
AOC	Airline Operations Center
APM	Automated People-Mover *also* Airport Performance Manual
APU	Auxiliary Power Unit
ARC	Airport Reference Code
ARSR	Air Route Surveillance Radar
ARTCC	Air Route Traffic Control Center
ARTS	Automated Radar Terminal System
ASD	Aircraft Situation Display *also* Accelerate-Stop Distance
ASDA	Accelerate-Stop Distance Available
ASDE	Airport Surface Detection Equipment
ASK	Available Seat-Kilometer
ASM	Available Seat-Mile
ASPIRE	Asia and South Pacific Initiative to Reduce Emissions
ASPM	Aviation System Performance Metrics
ASQP	Airline Service Quality Performance
ASR	Airport Surveillance Radar
ASSR	Airport Surface Surveillance Radar
ASV	Annual Service Volume
ATC	Air Traffic Control
ATCRBS	Air Traffic Control Radar Beacon System
ATCSCC	Air Traffic Control System Command Center
ATF	Aviation Trust Fund
ATFM	Air Traffic Flow Management
ATK	Available Ton-Kilometer
ATM	Available Ton-Mile *also* Air Traffic Management
ATM-NEMMO	Air Traffic Management macroscopic simulation model (France)
ATRS	Air Transportation Research Society

AWDPM	Average Week Day of the Peak Month of the year
BAA	British-based airports company (successor of British Airports Authority)
BART	San Francisco Bay Area Rapid Transit
BOT	Build, Operate, and Transfer
BT	Buffer Time
BTS	Bureau of Transportation Statistics (U.S.)
CAA	Civil Aviation Authority (U.K.)
CAD	Computer-Aided Design
CAEP	Committee for Aviation Environmental Protection (ICAO)
CASK	Cost per Available Seat-Kilometer
CASM	Cost per Available Seat-Mile
CBIS	Checked Baggage Inspection Systems
CCC	Capacity Coverage Chart
CCD	Continuous Climb Departure
CDA	Continuous Descent Approach
CDM	Collaborative Decision Making
CDTI	Cockpit Display of Traffic Information
CFC	Customer Facility Charge
CFMU	Central Flow Management Unit (Europe)
CH_4	Methane
CMAQ	Community Multiscale Air Quality (EPA)
CNG	Compressed Natural Gas
CNS	Communications, Navigation, and Surveillance
CO	Carbon Monoxide
CO_2	Carbon Dioxide
CODA	Central Office for Delay Analysis
COI	Cost Of Illness
Conracs	Consolidated rental car facilities
CPS	Constrained Position System
CR	Collaborative Routing
CRF	Concentration-Response Function

CTA	Controlled Time of Arrival
CTAS	Center TRACON Automation System
CTD	Control Time of Departure
CWY	Clearway
DA	Descent Advisor
DCC	Doppelmayr Cable Car company
DCV	Destination-Coded Vehicles
DDA	Delayed Deceleration Approach
DFS	German Air Navigation Service Provider (Deutsche Flugsicherung)
DME	Distance Measuring Equipment
DNL	Day-Night noise Level
DOT	Department of Transportation (U.S.)
DPD	Design Peak Day
DPH	Design Peak Hour
DSS	Decision-Support Systems
EA	Environmental Assessment
EAS	Essential Air Service (U.S. federal program)
EBRD	European Bank for Reconstruction and Development
ECAC	European Civil Aviation Conference
EDCT	Expected Departure Clearance Time
EDMS	Emissions and Dispersion Modeling System (U.S. FAA)
EIB	European Investment Bank
EIS	Environmental Impact Statement
EPA	Environmental Protection Agency (U.S.)
EPNL	Effective Perceived Noise Level
ETA	Estimated Time of Arrival
ETMS	Enhanced Traffic Management System
ETS	Emission Trading Scheme
EU	European Union
EUROCONTROL	The European Organisation for the Safety of Air Navigation
FAA	Federal Aviation Administration (U.S.)

FACET	Future ATM Concepts Evaluation Tool (a highly detailed network model)
FAF	Final Approach Fix
FAR	Federal Aviation Regulation
FAST	Final Approach Spacing Tool
FBO	Fixed-Base Operator
FCFS	First-Come, First-Served
FICON	Federal Interagency Committee on Noise (U.S.)
FIS	Federal Inspection Services (U.S.)
FMP	Flow Management Position
FSM	Flight Schedule Monitor
FTE	Full-Time Equivalent
FY	Fiscal Year
GAC	German Airport Coordinator
GAO	Government Accountability Office
GARB	General Airport Revenue Bonds (U.S.)
GBAS	Ground-Based Augmentation System
GDP	Gross Domestic Product *also* General Display Programs *also* Ground Delay Program
GHG	Greenhouse Gas
GNSS	Global Navigation Satellite System
GPS	Global Positioning Systems
GWP	Global Warming Potential
HDR	High Density Rule (U.S. FAA)
H_2O	Water
IATA	International Air Transportation Association
ICAO	International Civil Aviation Organization
IFR	Instrument Flight Rules
IHI	Ishikawajima-Harima Heavy Industries Co., Ltd.
ILS	Instrument Landing System
IMC	Instrument Meteorological Conditions. [Instrument Flight Rules (IFR) apply]

INM	Integrated Noise Model (U.S. FAA)
INS	Inertial Navigation Systems
IPCC	Intergovernmental Panel on Climate Change
IT	Information Technology
ITWS	Integrated Terminal Weather System
LAAS	Local Area Augmentation System
LAQ	Local Air Quality
LASPORT	Lagrangian Dispersion Model for Airports
LAWA	Los Angeles World Airports
LCC	Low-Cost Carrier
LDA	Landing Distance Available
LMINET2	A fast queuing network model
LOS	Level Of Service
LTO	Landing and Take-Off
MAGENTA	Model for Assessing Global Exposure to the Noise of Transport Aircraft (U.S. FAA)
MAP	Million Annual Passengers
MAR	Managed Arrival Reservoir
M/G/1	Poisson distribution that is Memory-less, General, and with 1 server
MHI	Mitsubishi Heavy Industries, Ltd.
MIT	Massachusetts Institute of Technology
MLW	Maximum Landing Weight
MTOW	Maximum Takeoff Weight
NAAQS	National Ambient Air Quality Standards (U.S.)
NADP	Noise Abatement Departure Procedure
NAP	Noise Abatement Procedure
NASA	National Aeronautics and Space Administration (U.S.)
NASPAC	National Airspace System Performance Analysis Capability (a simulation)
NDI	Noise Depreciation Index
NEPA	National Environmental Policy Act (U.S.)
NLC	Network Legacy Carrier

NO	Nitric Oxide
NO_x	Nitrogen Oxides
NO_2	Nitrogen Dioxide
NPD	Noise Power Distance
NPDES	National Pollution Discharge Elimination System
NPIAS	National Plan of Integrated Airport Systems
NRP	National Route Program (U.S.)
NU	Not Usable
O_3	Ozone
O-D	Origin-Destination
OECD	Organisation for Economic Co-operation and Development
OFZ	Obstacle-Free Zone
OPD	Optimized Profile Descent
OPSNET	Operations Network
PANCAP	Practical Annual Capacity
PANYNJ	Port Authority of New York and New Jersey
Pb	Lead
PCPI	Per Capita Personal Income
PFC	Passenger Facility Charge (U.S.)
PHCAP	Practical Hourly Capacity
PM	Particulate Matter
PMM	Persons per unit width per unit time (Persons/Meter/Minute)
PRT	Personal Rapid Transit system
PT	Positioning Time
QC	Quota Count
RBS	Ration By Schedule
RDC	Runway Design Code
RER	Réseau Express Régional (rapid rail system in Paris metropolitan area)
RF	Radiative Forcing
RFID	Radio Frequency Identification
RNAV	Area Navigation
RNP	Required Navigation Performance

ROFA	Runway Obstacle-Free Zone
RPI	Retail Price Index (U.K.)
RPZ	Runway Protection Zone
RSA	Runway Safety Area
RVR	Runway Visual Range
RWSL	Runway Status Lights system
SBAS	Satellite-Based Augmentation System
SBT	Standard Blocking Time
SCC	Schedule Coordination Conferences
SEL	Sound Exposure Level
SESAR	Single European Sky ATM Research
SFRB	Special Facility Revenue Bonds
SIMMOD	Simulation Model of Delays
SIRO	Service in Random Order
SO_x	Sulfur Oxides
SO_2	Sulfur Dioxide
SOT	Scheduled Occupancy Time
SSR	Secondary Surveillance Radar
STAPES	System for Airport Noise Exposure Studies (EUROCONTROL)
STARS	Standard Terminal Automation Replacement System
SWIM	System-Wide Information Management
SWOT	Strengths, Weaknesses, Opportunities, and Threats (strategic planning method)
SWY	Stopway
TAAM	Total Airspace and Airport Modeler (delay model)
TCA	Terminal Control Airspace
TCAS	Traffic alert and Collision Avoidance System
TDG	Taxiway Design Group
TGV	Train à Grande Vitesse (high-speed train, France)
TH	Tail Height
TMA	Traffic Management Advisor
TMU	Traffic Management Unit

TODA	Takeoff Distance Available
TORA	Takeoff Run Available
TPD	Typical Peak Day
TPH	Typical Peak Hour *also* Typical Peak Hour Passengers
TRACON	Terminal Radar Approach Control
TSA	Transportation Security Administration (U.S.)
TSS	Total Suspended Solids
UHBR	Ultra-High Bypass Ratio
UHC	Unburned Hydrocarbon
UHF	Ultra-High Frequency
U.K.	United Kingdom
U.S.	United States
USD	U.S. Dollar
V_1	Critical engine-failure speed
VAL	Véhicule Automatique Léger (French people mover)
VFR	Visual Flight Rules
VHF	Very High Frequency
VIP	Very Important People
VMC	Visual Meteorological Conditions [Visual Flight Rules (VFR) apply]
VOC	Volatile Organic Compound
VOR	VHF Omnidirectional Range
VSL	Value of a Statistical Life
WAAS	Wide Area Augmentation System
WS	Wing Span
WSG	Worldwide Slot Guidelines (IATA, 2012)
WTP	Willingness To Pay

Symbols

dB	decibel
dBA	decibel A-weighted
$E[X]$	Expected Value of X

ft	feet
h	hour
Hz, MHz	hertz, megahertz
I	Sound intensity
I_{ref}	Sound intensity reference (0 dB)
km	kilometers
L_{dB}	Sound level
m	meters
mi	miles
min	minute
nm	nautical mile
N_q	Number of users waiting in queue
S	Service time (queuing theory)
s	Number of servers (queuing theory)
T	Time of aircraft noise event
W_q	Time that a random user spends waiting in queue before being served by the system
λ	Demand rate, Greek lambda (queuing theory)
μ	Service rate, Greek mu (queuing theory)
ρ	Utilization ratio, Greek rho (used in queuing theory to denote average demand rate over a specified period of time divided by the average capacity over that time)
$\sigma(X)$	Standard deviation of X, Greek sigma
$\sigma_2(X)$	Variance of X, Greek sigma squared

Index

C

D

E

F

G

H

J

K

L

M

Q

T

U

V

W