Airport Demand and Capacity Modeling for
Flow Management Analysis

by

Guy A. Hocker

B.S., Electrical Engineering, United States Air Force Academy (1988)

Submitted to the Sloan School of Management
in Partial Fulfillment of the Requirements
of the Degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

at the

Massachusetts Institute of Technology

January 1994


Signature of Author

Sloan School of Management
Interdepartmental Program in Operations Research

Approved by

Stephan E. Kolitz
The Charles Stark Draper Laboratory, Inc.
Technical Supervisor

Certified by

Amedeo R. Odoni
Professor of Aeronautics and Astronautics and of Civil Engineering
Thesis Supervisor

Accepted by

Richard C. Larson
Operations Research Center, Codirector

MAR 23 1994
LIBRARIES
ARCHIVES
Airport Demand and Capacity Modeling for
Flow Management Analysis

by

Guy A. Hocker

B.S., Electrical Engineering, United States Air Force Academy (1988)

Submitted to the Sloan School of Management
in Partial Fulfillment of the Requirements
of the Degree of Master Of Science In Operations Research

Abstract

Air traffic flow management (FM) is the process of controlling the flow of aircraft through airports and airspace sectors for reasons of safety and efficiency. Because of increases in flight delays throughout the entire air traffic system, the FAA is encouraging research into alternative FM strategies. This thesis presents two models to support the development and evaluation of FM strategies. First, an approximate, airport demand scheduling model is presented. A heuristic algorithm produces hypothetical, daily airport flight schedules based on user-defined demand levels and flight connectivity parameters for a network of specified airports. With hypothetical schedules, air traffic experts can compare the effectiveness of FM strategies against projected future demand levels or changes in airline scheduling practices.

The second model is a weather simulation technique, developed by the United States Air Force, proposed for the purposes of airport capacity modeling. The simulation is based on a stochastic model that generates weather observations and forecasts for multiple sites with user-defined levels of spatial and temporal correlation. The mathematics involve superimposing multi-dimensional sawtooth waves in space and time. Testing suggests that both models could enhance FM research. This thesis also contains a detailed description of FM and air traffic resources.

Thesis Advisor: Professor Amedeo R. Odoni
Title: Professor of Aeronautics and Astronautics and of Civil Engineering

Technical Supervisor: Stephan E. Kolitz, PhD.
The Charles Stark Draper Laboratory, Inc.
Acknowledgements

I have many people to thank for their contributions to this thesis and my experience at MIT.

First, I thank Dr. Stephan Kolitz, who has been my mentor and friend from the beginning. I'm very grateful for his technical insight and support throughout the entire research process.

I would also like to express my appreciation to Professor Amedeo Odoni for his ideas and guidance. It has been an honor to work with a professional who is truly committed to research and people.

I am also grateful to Al Boehm of Hughes Corporation for his patience and time in explaining the sawtooth wave model; to the folks at FAA including JoAnne LeBreque, Bob LeBrie, Ken Thomas, and Gary Lavimoniere for introducing me to air traffic control; to Jeff Alexander for his help with the POAGG interface and general programming; and to Phylis Savari and Jonathan Robinson for their friendships and ground-breaking work on the ATC testbed and flight schedule generation.

I also thank my friends at the ORC for their diversity and enthusiasm. Thanks to them, taking classes at MIT and working on this thesis has been a great experience. Don't miss the grand opening of "Casa de Rodrigo y Guy", soon to be the best Mexican restaurant in Cambridge.

Special thanks go to my parents, John and Barbara, my sister, Connie, and her entire family. Without their unending love and support, this would not have been possible. "I thank my God in all my remembrance of you."

Finally, I would like to thank Draper, for its commitment to education and for funding this research.

This thesis was prepared at The Charles Stark Draper Laboratory. Publication of this thesis does not constitute approval by Draper or the Massachusetts Institute of Technology of the findings or conclusions contained herein. It is published solely for the exchange and stimulation of ideas.

I hereby assign copyright of this thesis to The Charles Stark Draper Laboratory, Cambridge, Massachusetts.

Guy A. Hocker
Captain, USAF

Permission is hereby granted by The Charles Stark Draper Laboratory to the Massachusetts Institute of Technology to reproduce any or all of this thesis.
# Table of Contents

Abstract .......................................................................................................................... 2  
Acknowledgements ......................................................................................................... 3  

## Chapter 1

1. **Introduction** ............................................................................................................. 8  
   1.1. Air Traffic Flow Management .............................................................................. 8  
   1.2. Thesis Objective and Content ............................................................................ 9  

## Chapter 2

2. **Air Traffic Management Primer** ........................................................................... 11  
   2.1. Air Traffic Resources ....................................................................................... 13  
   2.1.1. Airspace ....................................................................................................... 13  
   2.1.2. Airports ....................................................................................................... 27  
   2.1.3. Controller Duties, Responsibilities and Procedures ...................................... 32  
   2.1.4. Computer Resources .................................................................................... 35  
   2.2. Air Traffic Flow Management ............................................................................ 38  
   2.2.1. Air Traffic Congestion ................................................................................... 39  
   2.2.2. Flow Management ....................................................................................... 42  

## Chapter 3

3. **Airport Demand Schedule Modeling** .................................................................... 62  
   3.1. Motivation and Objectives ................................................................................ 62  
   3.1.1. Flow Management Analysis ......................................................................... 63  
   3.1.2. Commercial versus Total Air Traffic Demand ............................................ 63  
   3.1.3. Enhanced Traffic Management System ....................................................... 64  
   3.1.4. Modeling Objectives .................................................................................... 65  
   3.2. Problem and Approach .................................................................................... 67  
   3.2.1. Baseline Formulation ................................................................................... 68  
   3.2.2. Heuristic Solution Technique ..................................................................... 71  
   3.3. Implementation of the Model .......................................................................... 95  
   3.3.1. POAGG Functionality ................................................................................ 96  
   3.3.2. Feasibility Checks ....................................................................................... 98  
   3.4. Test Results ..................................................................................................... 100  
   3.4.1. Test Case 1: OAG Flight Parameters ......................................................... 104  
   3.4.2. Test Case 2: Connections .......................................................................... 107  
   3.4.3. Test Case 3: Shuttles .................................................................................. 113  
   3.4.4. Test Case 4: Airline Hubbing .................................................................... 114  
   3.5. Conclusion and Extensions .............................................................................. 115
Chapter 4

4. Airport Weather and Capacity Modeling ................................................... 116
   4.1. Background and Requirements ............................................................ 116
       4.1.1. Airport Weather ................................................................. 116
       4.1.2. Definitions ............................................................................. 118
       4.1.3. Requirements for an Airport Weather Model ............................ 120
       4.1.4. Alternative Weather Modeling Techniques ............................... 123
   4.2. Sawtooth Wave Model ........................................................................ 124
       4.2.1. Sawtooth Wave Geometry ....................................................... 125
       4.2.2. Correlation Curves ................................................................. 129
       4.2.3. Forecast Modeling ................................................................. 131
       4.2.4. Cross-Variable Correlation ...................................................... 132
       4.2.5. Inverse Transnormalization ...................................................... 133
       4.2.6. Basic Model Assumptions ......................................................... 134
   4.3. Implementation ................................................................................... 135
       4.3.1. Simulation Procedure .............................................................. 135
       4.3.2. Simulation Results ................................................................. 144
       4.3.3. Modeling Additional Variables ............................................... 152
   4.4. Discussion and Conclusion .................................................................. 153

Chapter 5

5. Conclusion ................................................................................................. 154
   5.1. ATM Primer ....................................................................................... 154
   5.2. Airport Demand Scheduling Model ...................................................... 154
   5.3. Airport Weather and Capacity Model ................................................. 155
   5.4. Potential Applications ....................................................................... 155
   5.5. Topics for Further Research ............................................................... 156

References .................................................................................................... 158
# List of Figures and Tables

| Figure 2.1: | The National Airspace System | 13 |
| Figure 2.2: | Terminal Area Airspace | 15 |
| Figure 2.3: | Side View of Terminal Area Airspace | 16 |
| Figure 2.4: | Terminal Area Sectors | 16 |
| Figure 2.5: | Terminal Area and En Route Airspace | 17 |
| Figure 2.6: | En Route Airspace Demand by Aviation Category | 18 |
| Figure 2.7: | Airport Demand by Aviation Category | 19 |
| Figure 2.8: | Major Airport Traffic | 20 |
| Figure 2.9: | Standard Routes | 24 |
| Figure 2.10: | Aircraft Separation | 26 |
| Figure 2.11: | Boston Logan International Airport | 29 |
| Figure 2.12: | IFR and VFR Thresholds | 31 |
| Figure 2.13: | IFR Flight Trajectory | 33 |
| Figure 2.14: | Enhanced Traffic Management System | 36 |
| Figure 2.15: | Resource Interdependencies - Sufficient Capacity | 41 |
| Figure 2.16: | Resource Interdependencies - Insufficient Capacity | 41 |
| Figure 2.17: | Airspace and ATM Hierarchy | 43 |
| Figure 2.18: | Ground Delay Program Implementation Cycle | 51 |
| Figure 2.19: | Ground Delay Program Timeline | 52 |
| Figure 2.20: | Display of Leadtimes for St Louis | 53 |
| Figure 2.21: | Follow-on Program Timeline | 59 |
| Figure 3.1: | Airport Demand by Aviation Category | 64 |
| Figure 3.2: | Relative Frequency Distribution of Arrivals at Airport $i$ | 73 |
| Figure 3.3: | Relationship Between $a_{i,t}$ and $c_{i,j,t}$ | 75 |
| Figure 3.4: | Arrival Distribution for Connection Assignment Example | 77 |
| Figure 3.5: | Originating and Connecting Flights | 79 |
| Figure 3.6: | Airport Network | 80 |
| Figure 3.7: | Connecting Flights versus Arrival Flights | 84 |
| Figure 3.8: | Probability Density Function for Layovers | 86 |
| Figure 3.9: | First Correction Technique | 87 |
| Figure 3.10: | Graphical View of Shuttle Service | 89 |
| Figure 3.11: | Three-Aircraft Shuttle Service | 90 |
| Figure 3.12: | Arrival and Departure Banks | 92 |
| Figure 3.13: | Arrival Times in Arrival Bank | 93 |
| Figure 3.14: | POAGG Flow Chart | 95 |
| Figure 3.15: | Test Cases - Flight Times | 101 |
| Figure 3.16: | Relative Frequency Distribution of Arrivals for Logan, January 1992 | 102 |
| Figure 3.17: | Relative Frequency Distribution of Arrivals for LaGuardia, January 1992 | 103 |
List of Figures and Tables (cont.)

Figure 3.18: Relative Frequency Distribution of Arrivals for National, January 1992 .........................103
Figure 3.19: Test Cases - Layover Time PDF .......................................................... 104
Figure 3.20: Test Case 1 - Connection Results ......................................................... 105
Figure 3.21: Test Case 1 - Shuttle Service 1 ............................................................ 107
Figure 3.22: Test Case 1 - Sample Output .................................................................. 107
Figure 3.23: Test Case 2 - Flights with Layovers Exceeding 2 Hours .................... 110
Figure 3.24: Test Case 2 - Proportion of Connecting Flights with Layovers > 2 Hours ......111
Figure 3.25: Test Case 2 - N=10, p”>50% ................................................................. 111
Figure 3.26: Test Case 2 - Total Flights ................................................................. 112
Figure 4.1: IFR and VFR Thresholds........................................................................ 117
Figure 4.2: Temporal Autocorrelation for Ceiling at DCA ................................. 121
Figure 4.3: Ceiling Spatial Correlation - Midwestern States ................................. 122
Figure 4.4: 1-D Sawtooth Wave ............................................................................ 125
Figure 4.5: 2-D Wave with Point p .......................................................................... 126
Figure 4.6: Earth-Centered Origin .......................................................................... 128
Figure 4.7: 4-D and 5-D Correlation Curves .......................................................... 130
Figure 4.8: Observation and Forecast Sites ............................................................ 132
Figure 4.9: Ceiling CDF for Site A .......................................................................... 134
Figure 4.10: Simulation Run Process ...................................................................... 136
Figure 4.11, 4.12, and 4.13: Lag s Spatial Correlation for Ceiling ....................... 145
Figure 4.14, 4.15, and 4.16: Lag s Spatial Correlation for Visibility ....................... 146
Figures 4.17 and 4.18: Temporal Autocorrelation for Ceiling and Visibility .......... 148
Figures 4.19, 4.20, and 4.21: CDFs for Ceiling ....................................................... 150
Figures 4.22, 4.23, and 4.24: CDFs for Visibility ..................................................... 151
Figure 4.25: Observed Windspeed CDF for BOS ................................................... 152

Table 2.1: Example of Delay Factors ..................................................................... 50
Table 2.2: Ground Delay Program Results By Departure Center ...................... 56
Table 2.3: Ground Delay Program Results By Airline ......................................... 57
Table 3.1: Rounding Process for a_1,i ..................................................................... 74
Table 3.2: Assignment of Initial Values for c_{1,2,i} .................................................. 78
Table 3.3: Test Case 1 - Airline Results ................................................................. 106
Table 3.4: Test Case 2 - Results for N=5, p”=50.0% .............................................. 108
Table 3.5: Test Case 2 - Results for Airport 4, N=5, p”=75%, Non-homogeneous p_{1,i} 113
Table 3.6: Test Case 2 - N=5, p”=75% .................................................................. 113
Table 4.1: Observed Lag s Spatial Correlation and Site Separation Distances ..........138
Table 4.2: Sample Ceiling Observation at BOS ...................................................... 143
Table 4.3: Cross-Variable Correlation Results ...................................................... 149
Chapter 1

1. Introduction

1.1. Air Traffic Flow Management

Although the U.S. operates the most sophisticated and largest system of air traffic resources in the world, flights are routinely experiencing unacceptable levels of delays. In 1990, 50 million hours of individual passenger time were spent in delays at the ten busiest airports alone [1]. Delays in the system are on the rise because of soaring demand for aviation resources and insufficient system capacity. It has been predicted that by the year 2000, 800 million passengers per year will participate in air travel [2] - nearly double the 1990 level. The total number of airline flights will continue to increase to keep pace with growing demand.

In addressing the problem, the Federal Aviation Administration (FAA) is encouraging research into alternative air traffic flow management (FM) strategies. Successful strategies will control the rates of traffic flow throughout the system, resulting in less congestion, reduced flight delays, and mitigated controller workloads. Furthermore, with improved efficiency more flights will be possible with the same resources.

The FAA manages air traffic and aviation resources through a hierarchical command and control structure known as Air Traffic Management (ATM). At the most strategic level, controllers in Washington D.C. manage a centralized flow facility that monitors all traffic and resolves national flow problems. It is conceivable that due to actions by controllers in the central facility, the departures of flights in Denver will be delayed due to congestion in Boston. At the most tactical level of this hierarchy, controllers in airport towers across the country facilitate the actual departures and arrivals of aircraft.

From an FM perspective, airports are the most important element in the air traffic system. The terminal area airspace surrounding airports experiences the highest overall density of traffic. If a major airport's capacity decreases, it is likely to become a bottleneck for the entire system, if FM is not properly implemented.

There are many factors influencing a traffic management specialist managing a constrained capacity scenario. For one, decisions must be based on future projections of both capacity and demand. Concerning capacity, the departure and arrival rates of airports are mostly affected by weather conditions, principally cloud ceiling height, visibility, wind
speed, wind direction and precipitation. Due to the uncertainty of weather conditions, airport capacity can be difficult to predict even a few hours in advance. Traffic management specialists work closely with weather personnel when projecting capacity prior to implementing FM actions.

Demand levels at an airport are also subject to change. High jet stream winds will cause planned traffic to arrive off-schedule, and unplanned traffic may appear in the system at any time. Moreover, airlines continuously make daily schedule adjustments.

In view of the complexities and uncertainty in air traffic, effective FM research requires a solid understanding of the ATM infrastructure and dynamic processes of FM. Moreover, the major components of the air traffic system, e.g., airport demand and capacity, must be accurately modeled in problem formulations. Without correct models, research experts may proceed under false assumptions resulting in inaccurate conclusions and recommendations.

1.2. Thesis Objective and Content

The research presented in this thesis is intended to set the stage for future FM research with system models. The thesis provides a primer on ATM and a model of airport demand scheduling. In addition, a weather simulation technique is proposed in support of airport capacity modeling. The thesis is divided into five chapters.

Chapter 2 presents a detailed description of air traffic resources and FM processes, in the form of an ATM primer. The first part of the chapter introduces the principal aviation resources and their roles with regard to FM. The second part describes FM issues and strategies and techniques available to controllers and traffic management specialists. Particular attention is given to the causes and dynamic nature of air traffic congestion. The primer was developed through an iterative review process with controllers and traffic management specialists at multiple FAA facilities. Their comments and suggestions were incorporated into the document.

Chapter 3 presents a multiple-airport demand scheduling model. The model contains an algorithm that produces hypothetical daily airport flight schedules that are statistically indistinguishable from real schedules and valid for FM research.

Schedules are generated based on user-defined demand levels of air traffic and flight connectivity parameters for a network of specified airports. Flight connectivity refers to the airline practice of scheduling a single aircraft to execute many flights throughout the day. Because of flight connectivity, flight delays will often ripple throughout the entire air traffic system.
Chapter 1: Introduction

The scheduling model provides air traffic research experts with a useful tool for the task of comparing the effectiveness of FM strategies with different demand scenarios. One interesting research area that may be addressed with the model is the performance of existing FM strategies given the projected increases in total demand and/or flight connectivity in the future. Existing flight schedules may be insufficient since they do not reflect future demand levels, nor do they contain flight connectivity information.

The airport demand scheduling problem is formulated as a mathematical program and solved using a heuristic algorithm that includes a model of airline shuttle services and hubbing operations. The solution is implemented in a computer application called the Pseudo-OAG Generator, or POAGG. The effectiveness of the algorithm is discussed in connection with the results of four test runs of POAGG.

In Chapter 4, this thesis assesses the suitability of an Air Force weather simulation technique for the purposes of airport capacity modeling for FM research. The simulation is based on a stochastic model that generates weather observations and forecasts for multiple sites. The model captures the interdependencies of weather observations at different sites, i.e., spatial correlation, and extends to produce data for multiple weather variables with cross-correlation amongst the variables. The mathematics involve superimposing multi-dimensional sawtooth waves in space and time, where temporal and spatial correlation is achieved as a function of the wavelengths. An analysis of sample, synthetically generated weather data suggests that the model preserves the statistical characteristics of weather that are important for FM research.

Finally, in Chapter 5 we present the conclusions from the modeling and experimental work performed in this thesis. In addition, future research questions are proposed.
Chapter 2

2. Air Traffic Management Primer

This chapter provides a detailed description of the current operational strategies of air traffic flow management (FM) by the Federal Aviation Administration (FAA). Air traffic FM is the management and control of aircraft operating through airports and airspace sectors within the United States. The principal purpose of FM is to achieve safe, orderly and efficient movement of traffic. The process of FM is implemented through a hierarchical command and control system, established by the FAA, known as Air Traffic Management (ATM).

From the perspective of an air traveler, ATM is an aviation service offered by the FAA. The two basic elements of ATM are 1) a series of rules and regulations established for the safety and efficiency of flight traffic and 2) FAA personnel, i.e., controllers and traffic management specialists, monitoring and directing traffic and resource utilization.

From the perspective of a controller or traffic management specialist, an air traveler places demand on air traffic resources. Air traffic demand often competes for resources, such as landing slots at an airport or access to a specific air route. If possible, FAA personnel seek to preserve a pilot’s prerogative for a flight, by minimally interfering with an intended trajectory. However, in a constrained resource scenario, e.g., too few landing slots at an airport, traffic management specialists must either address the capacity of the resource or manage the flow of traffic by adjusting flight itineraries, which is FM. Traffic management specialists are tasked to focus at least equal efforts on methods of expanding capacities, as are focused on FM strategies.

The primary objective of this chapter is to describe the operational functioning of FM strategies at different levels within the ATM hierarchy. The discussion will focus on managing actual and predicted congestion which materializes and can be resolved within the same working day. This excludes crisis FM strategies for extraordinary conditions, such as severely restrictive weather that disrupts traffic over multiple days, e.g., the East Coast blizzard in December 1992. There will be limited discussions on capacity enhancement efforts. The ultimate objective of this chapter is to support research in air traffic modeling and FM algorithm development.

The chapter is organized in two sections. The first section presents background information on aviation resources and air traffic demand. The emphasis is on major airport operations. Aviation resources will only be discussed functionally in terms of their roles with regard to FM. The second section presents a summary of FM strategies used by controllers and
traffic management specialists. Special attention will be given to the relationships among personnel and facilities involved in FM. This chapter will refer exclusively to FM within the 48 contiguous states known as the Continental United States (CONUS).

Research included a review of FAA literature and technical interviews with controllers, traffic management specialists and airline staff at the following locations:

- Boston Logan International Airport, ATCT and TRACON, Boston, MA
- Boston Air Route Traffic Control Center, Nashua, New Hampshire
- FAA ATC System Command Center, Washington D.C.
- United Airlines Control Center, Chicago, IL
- American Airlines Decision Technologies, Dallas-Ft. Worth, TX
2.1. Air Traffic Resources

The ATM system is a subset of the National Airspace System (NAS), the total collection of resources supporting aviation in the United States. Figure 2.1 depicts various components of the NAS.

![Diagram of NAS components: Air Traffic Management, radars, navigation equipment, controllers, aeronautical charts, computers, airspace, airports, other...]

Figure 2.1: The National Airspace System

The components of the NAS are interrelated. For example, controllers and traffic management specialists implement the ATM process, and a controller may be assigned authority over a specified region of airspace. This section will describe four components of the NAS from an ATM perspective: airspace, airports, controllers and FAA computers. This framework provides the necessary background for an understanding of FM.

2.1.1. Airspace

For management purposes, the airspace over the United States is divided into different regions known as sectors. Sectors are defined by geographical coordinates and altitude and resemble three-dimensional polygons. Air travelers move through sectors along a network of air routes that covers the United States. Among sectors, the size and level of traffic vary, depending upon the number of air routes traversing a sector and demand for travel along each route.

This section contains a description of sector division and control, air traffic demand and separation distances, fixes and air routes. Fixes are geographically referenced points along routes that assist in the control and FM process.

2.1.1.1. Sectors Classification and Control

Control of the airspace over the United States is achieved by delegating authority of groups of sectors to different FAA facilities across the country. Air traffic controllers at the these
facilities, monitor sector traffic on radar display screens and communicate control actions directly with aircraft pilots. Some facilities are equipped with graphics terminals, which receive flight information from a national FAA computer network, dedicated to ATM. A sector division is referred to as a control division of airspace.

At the same time, airspace is classified into regions based on regulatory divisions. Airspace regulatory divisions, independent of sector divisions, separate airspace regions into classes which contain specific rules for pilots operating in, and controllers managing the airspace. Aircraft equipment requirements varies between two different classes of airspace.

The difference between control and regulatory divisions is that control divisions are concerned with the management, or control, of aircraft, while regulatory divisions are concerned with the application of FAA rules. The subsequent descriptions of airspace divisions will cover both control and regulatory divisions. The two control divisions are terminal area and en route airspace. Regulatory divisions are Class A, B, C, D, E, and G\(^1\).

* Terminal Area Airspace

The airspace surrounding an airport, where most aircraft either maneuver in preparation for an airport landing or climb after takeoff to a higher altitude, is known as terminal area airspace. Airspace divisions and regulations for terminal area airspace reflect the heavy traffic environment surrounding an airport. Terminal area regulations become more restrictive as aircraft move closer to an airport. The basic dimensions and divisions of a typical terminal area are depicted Figure 2.2.

---

\(^1\)On September 16, 1993, the FAA implemented a revised nomenclature for airspace classes. This thesis will refer to the new airspace classes, i.e., A, B, C, D, E, and G, while providing the former airspace equivalent, when possible.
The outer bounds of terminal area airspace (the largest cylinder) is a control division separating the terminal area airspace from en route and uncontrolled airspace. The set of smaller cylinders within the terminal area constitute Class B airspace, a regulatory division and formerly the Terminal Control Area (TCA). Airports apply to the FAA for permission to establish Class B airspace based on high levels of air traffic at the airport. Airport controllers at a terminal facility have authority over all aircraft operating within Class B airspace. Any flight that desires to enter Class B airspace must receive prior FAA authorization. Most flights in Class B are either preparing to land or have recently departed the airport; however, aircraft may traverse Class B without landing at the airport, i.e., an “overflight”, upon receiving authorization.

The smallest cylinder centered on the airport in Figure 2.2 is Class D airspace, formerly the Airport Traffic Area (ATA), which extends up to 2,500 feet above the elevation of an airport and out to a horizontal radius of 5 nautical miles (nmi). Class D airspace is a subset of Class B, and both a regulatory and control division of airspace. There are specific controllers designated to manage Class D airspace, as well as specific regulations concerning aircraft operations therein. Only aircraft arriving at or departing the airport area authorized to enter Class D airspace. The remaining terminal area airspace is Class E, formerly General Controlled airspace.

Figure 2.3 shows a side view of terminal area airspace.
Terminal area airspace is further subdivided (for control purposes) into terminal area sectors, shown in Figure 2.4. Sectors are designed to balance controller workload and facilitate airport operations, namely aircraft approaches and departures.

The two FAA facilities which manage terminal area sectors are: Airport Traffic Control Towers (ATCTs) and Terminal Radar Approach Control Facilities (TRACONs). ATCTs are commonly referred to as towers. Approximately 400 airports in the United States are equipped with FAA operated towers. Class D airspace is managed by tower controllers, while the remaining terminal area sectors are managed by TRACON controllers. The TRACON and tower are collectively referred to as terminal area facilities. Often, the two facilities are collocated in the same building.

The dimensions and divisions of a terminal area are designed by FAA personnel, taking into account the local terrain surrounding an airport, alignment of runways, and density of airport traffic. The dimensions for the cylinders in Figure 2.2 are given as examples. The upper altitude of Class B airspace typically extends to a range of 7,000 to 10,000 feet above ground level. The altitude of the entire terminal area generally extends to an altitude of 10,000 to 14,000 feet above ground level. Obstructions near an airport, such as a mountain range, satellite airport, or a metropolitan area will impact the configuration of terminal area sectors, route structure, regulatory divisions and controller procedures.

---

2At some locations, TRACONs are referred to simply as Approach Control Facilities. Also, although the title implies control over approach flights, TRACONs also manage departing traffic and overflights.
In this thesis, any reference to an airport refers to a major airport equipped with Class B airspace and a TRACON.

- **En Route Airspace**

Most airspace in the United States away from airports and above 1,200 feet is en route airspace. Similar to terminal area airspace, en route airspace is divided into sectors to facilitate aircraft management and balance controller workload. En route sectors contain High (or Jet) and Low Altitude Airways. An airway (or route), defined by altitude and geographic coordinates, is approximately 8 nautical miles wide.

En route sectors vary in size depending upon the density of air traffic levels and the number of respective airways. Sectors near airports, where many routes converge, are smaller than sectors over less traveled areas. Sectors are often stratified between low and high altitudes. The separating altitude varies, although it is commonly at 18,000 ft above sea level, that being the threshold separating Low Altitude Airways from High Altitude Airways. Airspace above 18,000 ft is classified Class A, while airspace below 18,000 ft and away from terminal areas is Class E. Figure 2.5 is a graphic view of terminal area airspace surrounded by en route sectors.

![Figure 2.5: Terminal Area and En Route Airspace](image)

Aircraft operating in en route airspace are provided separation distances and other ATM services by controllers located in Air Route Traffic Control Centers (ARTCC), commonly known as Centers. There are 20 Centers across the United States monitoring all en route sectors. The Boston Center, located in Nashua, New Hampshire manages 34 sectors (20 Low and 14 High) that cover most of the New England states.

Technically, all airspace within a Center’s boundaries is the responsibility of the Center, including all terminal area airspace. In operation, responsibility for terminal area airspace is
delegated from Centers to respective TRACONs and airport towers through detailed Letters of Agreement which outline boundaries, control responsibilities and procedures for transitioning aircraft between facilities.

- **Uncontrolled Airspace**

  ATM services are not available in all airspace. Areas with a low density of air traffic, particularly in the western United States, at low altitudes (below 1,200 ft), are classified Category G or uncontrolled airspace. Pilots are permitted to operate aircraft in these areas, however, pilots will not receive controller assistance, and the aircraft may not be monitored on FAA radars. Uncontrolled airspace is indicated in Figure 2.5 with diagonal lines.

2.1.1.2. **Demand**

There are four categories of air traffic which place demand on air traffic resources:

- air carrier: airline jet operations between major cities
- air taxi: smaller commuter services with mostly propeller driven aircraft
- general aviation (GA): private or corporate owned aircraft
- military aviation

The following two charts illustrate the difference between en route airspace demand and airport demand. The percentages in Figure 2.6 provide a summary of the categories of total aircraft handled by en route facilities in 1992 [3]. Air carrier and air taxi services are collectively referred to as "commercial". Figure 2.6 suggests that commercial traffic places the most significant demand on airspace.

![Figure 2.6: En Route Airspace Demand by Aviation Category](image-url)
Figure 2.7 displays the 1992 airport demand percentages based on "total itinerant\(^3\) airport operations" for the ten busiest airports in the United States [3].

![Pie chart showing airport demand percentages]

**Figure 2.7: Airport Demand by Aviation Category**

The reduction in military demand to 1% at airports, as compared to 14% for en route airspace, is attributable to the fact that military flights typically operate out of military installations, with few operations originating or ending at commercial airports. The lower percentage of GA demand is due to the fact that a high proportion of general aviation traffic operate out of small to medium-sized airports, away from large metropolitan areas. The statistics in Figure 2.7 reflect demand at most major airports in the United States that experience congestion.

- **Flight Rules (IFR/VFR)**

  Flights are classified under flight rules according to the following two conditions:

  - scheduled highest altitude and
  - forecast weather conditions.

  If a flight is scheduled to enter Class A airspace by exceeding 18,000 feet at any time while airborne, or if a flight intends to navigate near or through instrument weather conditions, the flight is classified under *Instrument Flight Rules* (IFR). Instrument weather conditions refer to flying within specific minima in or around clouds, fog, or precipitation. In order to operate under IFR, an aircraft must be equipped with special instruments and the aircraft pilot must meet FAA

---

\(^3\)An itinerant operation refers to a flight that exceeds a 20-mile radius around a single airport.
IFR qualifications. While operating within controlled airspace, IFR traffic receives continuous controller services.

On the other hand, if a flight is scheduled to remain below 18,000 feet and weather conditions meet FAA standards for visual flying, the flight may be classified under *Visual Flight Rules* (VFR). VFR flights are permitted to enter controlled airspace below 18,000 ft, however, controllers are not required to provide VFR flights with air traffic and weather advisory services. VFR pilots must provide their own separation distances from all other flights and the terrain. Workload permitting, a controller may provide services to a VFR flight in controlled airspace upon request by the aircraft pilot.

Because of the busy nature of operations in and around airports, all flights, i.e., VFR and IFR, are under *positive control* when operating in Class B airspace. A flight under positive control is subject to the authority of air traffic controllers monitoring the respective sectors, regardless of classification. VFR traffic is referred to as *Controlled VFR* upon entering Class B airspace.

There is a third type of traffic known as *Tower En Route Control (TEC)* traffic. A TEC flight occurs when the terminal areas of two airports are adjacent due to the proximity of the airports and as a result, a transient flight between the airport pair never leaves terminal area airspace. Figure 2.8 summarizes the three basic types of traffic at major airports.

![Figure 2.8: Major Airport Traffic](image)

- **Flight Plans**

The FAA requires that flight plans are filed for all IFR flights. By filing a flight plan, a pilot registers the flight’s itinerary within the NAS computer system. A flight plan contains the estimated flight departure and arrival times, origin and destination airports, planned route of travel, and information about the flight aircraft. Once a flight plan is submitted to the FAA, the pilot must receive permission to deviate from the originally submitted route. The FAA encourages flight plan filing at least one hour prior to departure, however, flight plans and flight
plan amendments may be filed even after a flight is airborne. Filing a flight plan is optional for VFR traffic.

Most air carrier traffic is IFR. Since air carrier flights constitute the majority of air traffic, air carrier flight plan itineraries assist controllers in forecasting congestion at airports and in airspace sectors. Published airline schedules also assist FAA planners by providing an estimate of daily and monthly demand. Daily flight schedules are collectively published in a monthly listing known as the Official Airline Guide (OAG), which allow system planners to estimate demand, down to the 15 minute interval, at airports and airspace sectors.

The utility of schedules is limited by the dynamic nature of airline operations. Airlines, not bound by published schedules, react to the daily air traffic environment (weather, delays, etc.) and the air travel market. Airlines may adjust flight times or cancel flights based on excessive delays, low passenger demand, crew availability, equipment problems and/or scheduling constraints.

- **Flight Clearance**

Controllers authorize flights to proceed with an action by issuing flight clearances. During daytime traffic levels between major airports, clearances are issued for segments of a flight's itinerary. For example, if a flight is prepared for departure, the pilot contacts the tower and requests departure clearance. Once granted, the flight may proceed according to specific instructions.

A clearance specifies the limit of the authorization. In the departure example, the pilot will receive authorization to depart and proceed to a route, heading, and altitude. Although the pilot's ultimate objective is to travel to a destination airport, the clearance will most likely authorize the pilot to complete only a segment of the intended itinerary. Prior to completing the first segment, the pilot will receive clearance or alternative instructions for the succeeding segment. This process is repeated until the aircraft successfully reaches its destination. By controlling traffic at intermediate stages, controllers may tactically react to unforeseen weather and demand conditions. Due to low nighttime traffic levels, an overnight flight may receive clearance to complete its entire trajectory upon departure.

**2.1.1.3. Airspace Capacity**

Terminal area and en route sectors are capacitated. A sector's capacity is the total number of aircraft (aircraft count) that a sector can accommodate for a defined period of time. Sector capacity is defined by the FAA as the Operational Acceptable Level of Traffic (OALT). The
OALT is determined for a sector based on aircraft separation standards, route structures, weather, equipment availability, and controller staffing and experience levels. A sector’s OALT varies according to changes in the dynamic variables, such as weather. The motivation for setting maximum sector capacities is based upon the need to maintain operationally safe controller workloads and safe separation distances between aircraft. Although physical capacity within a sector may exist, a controller may deny a flight the permission to enter in order to maintain a reasonable workload.

Improving airway and sector capacity is an ongoing process under the oversight of Center, TRACON, and national FAA capacity offices. Sectors are continuously evaluated according to past traffic levels and congestion problems. Often a Center or TRACON will resectorate a region, by redefining the borders and altitudes of sectors, in order to improve efficiency or balance controller workload. Routes are also adjusted or modified for similar reasons.

2.1.1.4. Fixes

A fix is a geographic location marked by a visual reference, navigational beacon or other means. Fixes are located in both en route and terminal airspace along air routes. Controllers typically clear flights onto specific routes using a fix location as a heading for the aircraft. The alternative to fix navigation is vectoring an aircraft with radar headings, a more controller-intensive navigational process.

Fixes also serve as control points for traffic management specialists. At fixes, controllers monitor and adjust the rates of aircraft flow into and out of sectors or routes.

Some of the principal types of fixes are:

- **Arrival/Departure Fix**: Arrival fixes are the entry points for most flights entering a terminal area. Departure fixes are points along departure routes that mark the exit of a terminal area.

- **Metering Fix**: A metering fix is a point along a route at which en route controllers establish desired rates of flows of aircraft, typically into a terminal area. The metering process is either manual or automated. Automated metering involves an algorithm that regulates, over time, the arrivals of flights into a terminal area based on desired hourly rates of arrivals. Each arriving flight is assigned a fix crossing time.
- **Outer Fix**: Outer fixes are established along routes prior to metering fixes, as potential holding points for aircraft and to assist in the calculation of fix crossing times.

- **Coordination Fix**: A coordination fix is the point between sectors or routes, at which facilities will transfer control of an aircraft between controllers and facilities.

- **Instrument Approach Fixes**: Located within terminal area airspace, instrument approach fixes are used to assist pilots executing airport landings with instrument control systems. Instrument approaches are usually performed due to inclement weather conditions at an airport.

Some fixes serve multiple functions, e.g., an arrival fix may also be a metering fix.

To assist controllers in achieving a flow rate of aircraft through airspace, most fixes have an associated holding pattern where aircraft are delayed, if necessary, prior to crossing a fix. Delayed flights are cleared onto elliptical holding patterns at different altitudes. As capacity becomes available, flights will be cleared off the holding patterns and over the fix.

The capacity of a fix refers to the maximum allowable flow rate of aircraft. If a fix reaches its maximum capacity, it is referred to as **saturated or overloaded**. Fixes along routes entering terminal areas typically become saturated if airport capacity becomes constrained. In a saturated fix scenario, traffic management specialists at Centers are encouraged to execute **fix balancing**, the process of dispersing entering traffic over multiple arrival fixes in order to avoid heavy congestion in specific sectors.

The advent of Global Positioning Satellites (GPS) and GPS technology is predicted to impact fix navigation. The airlines and the FAA are in the process of evaluating the use of GPS satellites to provide more accurate and timely aircraft location and guidance information. GPS has the potential to enable controllers to clear pilots to points in airspace independent of fix locations. This technique, referred to as point-to-point navigation, could potentially alleviate fix overloading and sector congestion problems.

### 2.1.1.5. Routes

Similar to a network of highways, air routes consist major and minor airways, intersections, merge routes, one-way routes and areas of speed control. The FAA has established **standard and preferential** routes through routinely traveled regions of airspace. Standard routes are preplanned, published routes which assist pilots and controllers when transitioning between en route airways and terminal areas. Preferential routes provide an efficient trajectory for travel
between two airports or airspace sectors. Preferential routes are not mandatory, however, the FAA recommends pilots adhere to preferential routes whenever possible, in order to assist in the flight coordination process among controllers.

There are two types of standard routes: Standard Instrument Departure (SIDs) and Standard Terminal Arrival Routes (STARs); and four types of preferential routes: Preferred IFR, Preferential Departure, Preferential Arrival, and Preferential Departure and Arrival.

* **Standard Routes**

To facilitate the transition of aircraft between en route airways and terminal areas and vice versa, the FAA has established standard routes connecting terminal areas to high altitude routes. These routes are known as Standard Instrument Departure (SID) and Standard Terminal Arrival Routes (STAR). Similar to vehicular traffic, SIDs and STARs resemble one-way, non-intersecting streets. STARs contain holding patterns outside of a terminal area, where aircraft may be directed to execute airborne delays waiting for airport arrival slots. The diagram in Figure 2.9 provides an overhead view of two STARs and a SID, passing through low altitude-en route sectors towards, or away from in the case of the SID, a terminal area.

![Diagram of Standard Routes](image)

**Figure 2.9: Standard Routes**

SIDs and STARs originate and end at fixes. Considerable planning is given to the design of SIDs and STARs to reduce the need for realtime, pilot-controller coordination.
• **Preferred IFR Routes**

Preferred IFR routes provide pilots with efficient trajectories (from departure to arrival) between two airports. Normally extending through one or more Center areas, Preferred IFR routes are designed to achieve balanced flows among high density terminal areas. A portion of a Preferred IFR route may consist of a SID or STAR.

• **Preferential Routes**

Preferential routes are normally confined to one Center area. Preferential routes are recommended routes of travel to assist flow management through busy areas, particularly major airport terminal airspace. A Preferential Departure Route (PDR) describes a route from an airport to an en route point. A PDR may be a subcomponent of a SID or Preferred IFR route. A Preferential Arrival Route (PAR) describes a route from an en route point to a terminal area location, and similarly, may be included in a STAR or Preferred IFR route. Finally, a Preferential Departure and Arrival Route is similar to a Preferred IFR Route, however, PDARs connect two airports within one Center area or adjacent Center areas.

2.1.1.6. **Aircraft Separation**

The fundamental service offered by FAA personnel is safe separation distances among all aircraft within the system. Separation is considered to be the essence of ATM, because aircraft separation ensures safety. Separation intervals also determine throughput at airspace sectors and airports. Even though increased separation provides increased safety, high separation intervals reduces throughput. Traffic management specialists continuously study tradeoff methods of increasing throughput without compromising safe aircraft separation distances.

The diagrams in Figure 2.10 depict the three types of aircraft separation: longitudinal, lateral and vertical. Longitudinal refers to the distance between aircraft at the same altitude traveling along the same route in succession. Vertical separation refers to the difference in altitude between aircraft at the same geographic location. Lateral separation is the distance horizontally perpendicular to an aircraft’s direction of flight. Of the three, longitudinal separation has the greatest effect on throughput.

Maintaining safe separation distances is complicated by the fact that aircraft travel at different speeds and altitudes. Furthermore, aircraft on intersecting or merging routes require special attention.
Figure 2.10: Aircraft Separation

Fundamental to a controller's ability to provide aircraft with separation is the controller's knowledge of each aircraft's location in a sector. Since different monitoring techniques, e.g., radars, transponders, or direct communication with pilots, yield position data with different tracking resolution, FAA separation standards are largely dependent upon the manner in which an aircraft is being controlled.

Separation methods for IFR aircraft, and Controlled VFR traffic, are classified as radar and nonradar. Radar involves tracking an aircraft's position electronically on radar display screens. Nonradar implies inferring an aircraft's position through alternative means, e.g., communication with the pilot. Radar techniques and standards allow controllers to reduce the longitudinal separation between aircraft in the system. In most cases, use of radar over nonradar techniques increases airspace efficiency, reduces controller workload, and enhances safety.

The basic methods of achieving IFR separation are:

- Lateral: aircraft are cleared onto different airways whose widths and paths do not overlap.

- Vertical: aircraft are directed onto airways at altitudes that differ by at least 1,000 feet. Altitude assignments are based on direction of flight.

- Longitudinal: aircraft are sequenced into trails with controlled airspeeds. Controllers, using radar, maintain at least the following spacing intervals between any two aircraft:
Chapter 2: ATM Primer

- terminal area airspace: 3 miles
- terminal area on final approach within 10 miles of airport: 2.5 miles
- en route airspace: 5 miles

There are many exceptions and variations to the above minima. Special separation techniques are employed to maintain safe distances when merging traffic, managing flights in holding patterns, and separating multiple departing aircraft.

In the terminal area, visual separation is an alternative means of separating aircraft. Visual rules apply if a pilot can maintain continuous visual contact with the preceding aircraft. Under visual rules, the separation interval in the terminal area may be decreased to as low as 1/2 mile depending on the type of aircraft involved. Operations involving larger aircraft require increased intervals due the degree of turbulent air in the aircraft wakes.

If visual flying conditions prevail, controllers seek to clear flights for visual approaches and landings. If a terminal area controller clears a flight for visual approach and the pilot accepts the clearance, the pilot assumes the responsibility to maintain separation. Pilots may initiate the request for visual approaches as well. An IFR flight retains its IFR classification even though it may be cleared for a visual approach.

As a result of visual approaches, an airport’s arrival rate should increase, as arriving flights land with reduced separation distances. At the same time, controller workload is reduced, since pilots assume the responsibility of maintaining separation.

2.1.2. Airports

The capacity of an airport refers to the rate of aircraft operations, i.e., takeoffs and landings, an airport can sustain over a period of time. Airport capacity is synonymous with airport throughput. Projects designed to enhance an airport’s capacity can be classified according to the project’s planned impact on airport operations and implementation time. Major capacity enhancement projects for airports are typically long-term (10 to 15 years), or medium-term (2 to 3 year) efforts. Building a new, expanded airport, is an example of a long-term effort. Medium-term efforts include projects that enhance existing facilities, as in acquiring improved radar equipment or extending a runway to increase its operational utility.

Short-term capacity efforts exist, however, the capacity increases are limited. An example of a short-term project is aircraft sequencing procedures. Based on capacity studies, some airports have improved throughput by modifying the process of sequencing aircraft for landings, thus requiring no physical modifications.
Chapter 2: ATM Primer

The management of the active runways and terminal area airspace at an airport have a significant effect on throughput. Inefficient airport management, particularly at a major airport, will negatively impact traffic flows, and may affect national airspace operations, if delays are excessive.

2.1.2.1. Realtime Airport Capacity

Realtime airport capacity is a result of actual aircraft separation intervals implemented by controllers or pilots. For an airport to operate at full capacity, controllers and pilots must adhere to the minimum separation standards. However, pilots and controllers may choose to increase aircraft separation distances for various reasons. Controllers may delay clearance for aircraft departing an airport or entering terminal area airspace in order to maintain reasonable controller workloads. Also, since experience levels differ, inexperienced pilots or controllers may logically choose to maintain greater separation distances for safety reasons.

2.1.2.2. Runway Operations

A physical runway is in fact two independent runways corresponding to the two possible directions of aircraft operations. The numerical reference of a runway corresponds to the approach (or takeoff) direction of an arriving (or departing) aircraft. For example, at Boston Logan (Figure 2.11), Runway 27 and Runway 9 are the same physical runway. In practice, it is Runway 27 when operations are preceding from east to west (27 corresponds to magnetic compass heading of 270°) and Runway 9 when operations are preceding from west to east (i.e., a compass heading of 90°). The magnetic headings for runways allow pilots to reference on-board, magnetic directional equipment when confirming runway identity.
Runways vary in length and level of equipment. Some runways are equipped with electronic systems to allow pilots to land using onboard landing instrumentation. The current standard landing equipment is the Instrument Landing System (ILS). Non-ILS runways are only available during visual flying conditions. The length of the runway also impacts its utility. The FAA has established minimum length requirements for specific classes of aircraft, conditional on the altitude of the airport. For example, wide-body transport aircraft require a minimum of 8,000 feet of landing runway for landings at sea level.

An airport's runways are grouped together into different FAA authorized, operational configurations. The active runway configuration for an airport refers to:

- the current runways in use,

- the operation(s) to be conducted on each runway (i.e., departures, landings or both), and

- the type and category of aircraft dedicated to each runway (i.e., propeller and/or jet engine aircraft ; IFR and/or VFR).

Boston Logan has over 26 different approved configurations. Most configurations are rated by the FAA with an hourly acceptance rate and an hourly departure rate, known as the Engineered Performance Standards (EPS). EPS figures are mathematically derived rates for each
airport on an individual basis and reflect an airport's aircraft mix (jet versus prop), operating procedures, runway layout, and specific weather conditions.

2.1.2.3. Configuration Selection Process

Airport towers are managed by tower supervisors. There are multiple factors and tradeoffs leading to a tower supervisor's selection of an active configuration. The overall objective is to select the configuration that best meets the realtime and projected needs of the airport.

The factors contributing to the tower supervisor's decision are described below. The first three are weather related.

• Wind Speed and Direction

An airport tower supervisor selects active runways from those that are oriented for aircraft landings and takeoffs with headwind. Headwind reduces the speed of aircraft relative to the ground, thus requiring less runway length and time for landings and departures. Since the FAA requires that an aircraft safely exits a runway before a subsequent aircraft may enter a runway, as aircraft move off runways more quickly the overall capacity of an airport improves.

Landing or taking off with headwind is also safer. Reduced takeoff and landing distances provide pilots with additional runway length that may be necessary if a pilot error or equipment failure occurs.

Concerning crosswind, generally the maximum allowable crosswind for an active runway is 15 knots. While controllers may defer to a pilot's discretion concerning crosswinds, the tower supervisor determines the operational configuration only considering runways with crosswind components of less than 15 knots. It is conceivable, therefore, that wind speed and direction may dictate an active configuration with less than optimal capacity.

• Cloud Ceiling and Visibility

Cloud ceiling refers to the altitude of the lowest cloud in the airport area. Visibility refers to the horizontal visibility at ground level. These two aviation weather variables classify an airport as IFR or VFR, which are separate from instrument and visual flight rules that apply to flight classification. An airport operating in IFR may only conduct landings on adequately equipped runways. Under IFR controllers may not authorize visual approaches. Figure 2.12 provides the visibility and ceiling thresholds for airport classifications. LIFR refers to Low IFR. MVFR is Marginal VFR.
The tower supervisor disqualifies inadequately equipped runways during IFR weather conditions, when selecting the active configuration.

- **Precipitation**
  Precipitation impacts airport operations by increasing the required runway length for operations. Aircraft maneuver on and off runways with slower speeds if precipitation is present, reducing throughput.

Other non-weather factors considered by tower supervisors when selecting configurations are:

- **Arrival/Departure Ratio**
  Each configuration is rated with two capacity rates for each aviation weather condition, i.e., the acceptance and departure rates. In general, departure rates are either equivalent or greater than acceptance rates. Certain airports consider the projected demand ratio of arrivals to departures when selecting the active configuration, if there is a biased need for one type of capacity over the other.

- **Noise Abatement**
  Local communities can exert enormous pressure on airport authorities to disperse the amount of air traffic in the terminal area. For instance, Logan Airport in Boston is constrained from using certain runways for extended periods of time due to noise complaints from
surrounding residential areas. Configuration selection may include determining the approach and departure patterns for air traffic that minimizes the impact of noise on the local community.

• Maintenance

Most major airports are owned and maintained by regional authorities that will close runways for standard maintenance or emergency repairs. The schedule for routine maintenance is normally a coordinated process between the FAA and the local authority. The result, however, is a reduction in usable runways.

The tower supervisor attempts to maximize the efficiency and capacity of an airport, given the status of the above influencing factors. Due to the dynamic nature of many of the factors involved in the process, particularly weather conditions, a tower supervisor may be required to change configurations several times throughout the day. The realtime impact of a configuration change on capacity is substantial. Although necessary at times, a change often results in the need to adjust the positions many arriving, airborne aircraft, and redirect ground traffic of departing aircraft. A configuration may initially result in reduced overall capacity and additional controller workload. At Logan between May 1991 to May 1992, the average number of daily configuration changes was four.

2.1.3. Controller Duties, Responsibilities and Procedures

The pilot of an IFR flight typically interacts with many controllers during the flight trajectory. As an aircraft leaves one sector and enters another, control is transferred from one controller to another. When transitioning between terminal and en route airspace, control of an aircraft is transferred from one facility to another. The trajectory depicted in Figure 2.13 demonstrates the different phases and events of an IFR flight.
Figure 2.13: IFR Flight Trajectory

This section will describe an example IFR flight’s trajectory, in order to introduce the different air traffic controller duties, responsibilities, and procedures.

2.1.3.1. Airport Traffic Control Tower

Airport towers are the first FAA facility encountered by a pilot. Tower controllers maintain control of a departing aircraft from the departure gate until the aircraft departs Class D airspace. When clearing an IFR flight for departure, clearance is coordinated with terminal area controllers at the TRACON. The tower controller also coordinates with the Center managing the airspace surrounding the terminal area to ensure adequate capacity exists in the immediate en route sectors.

At the departure airport (Airport A in Figure 2.13), the pilot receives clearance to pushback from the gate from Ground Control. Ground Control usually consists of one to two tower controllers directing taxi and ground operations. From gate departure until an aircraft taxis onto the active departure runway for takeoff, a flight is under the direction of Ground Control. If there is excessive departure demand at an airport, Ground Control sequences and manages flights in departure queues. Once cleared onto the departure runway, control of the flight is transferred to Local Control, which consists of one controller per active runway. Local Control tracks and advises flights in Class D airspace up to and including the active runway.

At Logan, as at many airports, Ground Control and Local Control are collocated in the tower facility. The tower facility provides controllers with full visual (weather permitting) and radar monitoring capabilities of flights at the airport and in the immediate airspace. Due to ground movement of vehicles and taxiing aircraft, runway operations are highly interactive, requiring extensive coordination between Local and Ground controllers.
2.1.3.2. Terminal Radar Approach Control

Once airborne a flight is transferred to a TRACON controller, for the departure leg of the flight. TRACON controllers will accept departing flights from Local Control, and direct flights safely through the remaining terminal area airspace, usually via a PDR to a coordination fix. At this point, responsibility for the flight is transferred to an en route controller. Concerning arrival flights, TRACON controllers accept arriving IFR traffic from en route facilities, directing flights through terminal area sectors to the Class D airspace immediately surrounding the airport.

The TRACON Traffic Management Unit (TMU) is staffed with traffic management specialists monitoring the efficiency of the terminal area airspace. TMU personnel are concerned with both strategic modifications to the standard routes and tactical congestion problems that occur on a daily basis.

2.1.3.3. En Route Control

Aircraft enter an ascent transition phase upon departing the terminal area airspace, which extends out to approximately 250 miles from an airport. Under the supervision of en route controllers, a flight climbs at a controlled rate to a cruising altitude. Under normal traffic conditions, flights are successively cleared onto SID's, in order to maintain an orderly flow. In low traffic periods, most flights are quickly cleared to fuel-efficient altitudes.

Upon attaining a cruising altitude, the departure phase of an IFR flight is complete. Control is transferred from controller to controller as flights traverse sectors during the cruise phase. An interesting scenario arises if a low altitude controlled flight passes through an airport's terminal area airspace as an overflight. In this situation, control of the aircraft is temporarily transitioned to terminal area controllers until re-entering en route airspace.

Within approximately 200 miles of the destination airport, the pilot begins the descent transition phase. The descent transition phase is the reverse of the ascent. If arrival capacity is insufficient to accommodate demand, an arriving flight may be directed into an airborne holding pattern along an arrival route near the destination airport. For short delays, i.e., less than 15 minutes, controllers may apply speed restrictions or other informal FM strategies. Flights are eventually cleared into the terminal area as airport capacity becomes available.

Once inside the terminal area, the IFR flight follows an arrival route, usually a PAR, under the guidance of an approach controller in the TRACON. As the flight reaches Class D airspace, control is transferred to a Local Controller, and cleared to land. Typically, after landing, the flight taxis under Ground Control supervision, to the arrival gate, completing the trajectory.
In addition to en route sector control, Centers operate Traffic Management Units (TMU) which are responsible for FM within Center boundaries. Center TMUs are normally staffed with two to three controllers and a shift supervisor.

A sample of the functions performed by Center TMUs is:

- acts as a central agency for all airports within the Center boundaries on weather advisories and anticipated congestion
- negotiates the forecasted capacity for major airports within the Center boundaries with the airport tower supervisors
- coordinates traffic flow with neighboring Centers and TRACON TMUs
- ensures the rate of aircraft entering a terminal area matches, or remains below, the negotiated capacity rates
- coordinates and implements national flow management initiatives
- issues long-range clearances to aircraft departing an airport, if necessary
- receives and processes all flight plans filed within the Center boundaries

2.1.3.4. Air Traffic Control System Command Center

The Air Traffic Control System Command Center (ATCSCC) oversees all air traffic in the United States. Staffed with controllers and personnel from various FAA facilities across the country, ATCSCC coordinates all national FM initiatives and some regional air traffic activities. ATCSCC controllers rarely interact directly with flights, but rather exert influence on the aggregate flow of aircraft in the system. All national FM strategies are initiated at ATCSCC and transmitted to the appropriate Centers, TRACONs or towers. The FM responsibilities of ATCSCC and Center TMUs will be discussed in the second section of the chapter.

2.1.4. Computer Resources

The FAA relies heavily on automation and data systems to handle the large volume of flight, weather, and NAS data required for the effective control of flights. The Enhanced Traffic Management System (ETMS) and Arrival Spacing Program (ASP) contribute to the FM process.
2.1.4.1. Enhanced Traffic Management System

A vital component of the command and control of air traffic is the ability to continuously collect, distribute, and project information on the state of the NAS. Traffic management specialists require realtime information on aircraft locations and airport capacities in order to manage traffic flows. This information is made available to ATCSCC and Center and TRACON TMUs, through the Enhanced Traffic Management System (ETMS).

ETMS is a network of computers that maintains and monitors national air traffic data and displays the data to traffic management specialists in a variety of formats. Realtime flight control data and status is received from the 20 Centers across the country and fused at the Inter-Facility Flow Control Network (IFCN), in Atlantic City, NJ. Data from IFCN and other sources are inputs into ETMS processors at the Volpe Transportation Center in Cambridge, MA. Volpe is responsible for the day-to-day effectiveness of ETMS and future upgrades of the system. The processed data are transmitted every 3 minutes from Volpe to ATCSCC and TMUs. An overview of the ETMS system is depicted in Figure 2.14, followed by a description of the main components.

![Diagram of Enhanced Traffic Management System]

*Figure 2.14: Enhanced Traffic Management System*

*Flight Data*

FAA planners project daily demand using the Official Airline Guide (OAG). Flights from the OAG are identified as scheduled until a flight plan is submitted by an airline scheduling office to one of the Centers. The Center inputs the flight plan data into the NAS computer (NAS computers are located at every Center), via a NAS message. At this point, a flight is upgraded to proposed. Once airborne the flight becomes active. Some scheduled flights are never upgraded to
proposed, if canceled by an airline. VFR flights only appear in the system when pilots chose to submit flight plans, therefore, the initial OAG list is an approximate estimate of daily demand.

- **Geographic and Weather Data**

ETMS fuses multiple sources of geographic data to generate graphic overlays. Geographic data define airport and fix locations, sector and Center boundaries, and airway routes. The FAA acquires weather data from the Weather Message Switching Center (WMSC) consisting of jet stream wind reports, terminal (airport) forecasts, and hourly, airport weather observations, known as terminal surface observations.

- **Aircraft Situation Display (ASD)**

Traffic management specialists access data through an integrated, multi-functional display platform, referred to as the Aircraft Situation Display (ASD). The utility of ETMS is greatly enhanced by the selectivity and charting capability of ASD. A traffic management specialist has the ability to request that ASD display all active flights, or highlight specific flights with certain characteristics, such as all flights over a certain altitude, bound for a specific destination airport, or flights of a certain airline. Selectivity also applies to geographic features. The normal use of ASD is for displaying real-time flight positions.

- **Monitor/Alert**

ETMS allows a traffic management specialist to project traffic demand levels at airports, sectors, and fixes. Upon request, the Monitor/Alert feature of ETMS automatically alerts a specialist if the projected demand for a sector or airport exceeds a user-defined threshold. The accuracy of Monitor/Alert forecasts is conditional upon the tracking of individual flights. Alerts are graphically displayed through the ASD interface.

**2.1.4.2. Arrival Spacing Program**

The Arrival Spacing Program (ASP) is an automated computer program located at selected Centers. ASP is used to meter traffic into a terminal area during periods of excess demand or restricted capacity. The program evaluates all incoming IFR traffic into a terminal area and calculates a crossing time for each aircraft at a metering fix, working from a predetermined touchdown time at the destination airport. The program estimates flight times from fix crossing to runway touchdown given the approach route in use by the TRACON. The ASP output consists of fix crossing times and anticipated airborne holds for flights bound for a capacity constrained airport. The ASP is not a part of ETMS.
2.2. Air Traffic Flow Management

This section describes the basic air traffic flow management (FM) strategies available to traffic management specialists. No attempt is made to cover all methods of alleviating or avoiding congestion. Instead, we address a representative set of standard FM strategies that are designed for routine congestion. Although the implementation of FM occurs at the controller level, initiatives and command decisions are coordinated and directed at the Traffic Management Unit (TMU) level and higher. This section focuses on processes through which traffic management specialists initiate FM strategies. The principal entities in this discussion are the TMUs at TRACONs and ARTCCs, and national-level command management as directed by ATCSCC.

Reviewing the functions of the principle entities, TRACONs (Terminal Area Control Facilities) are responsible for the management of terminal area airspace. TRACON controllers direct departing and arriving flights through the heavy traffic sectors surrounding major airports. ARTCCs (Air Route Traffic Control Centers, or simply Centers) manage all en route airspace sectors. In order to reduce overcrowding within terminal area airspace, many FM initiatives are conducted at the ARTCC level in en route sectors. Both Centers and TRACONs are staffed with traffic management specialists at Traffic Management Units (TMUs). TMU personnel are concerned with both the strategic resource utilization and tactical FM strategies to relieve realtime congestion. Finally, the Air Traffic Control System Command Center (ATCSCC), located in Washington D.C., oversees all air traffic flow in the United States. ATCSCC coordinates national FM initiatives and some severe regional air traffic activities as well.

As expected, there are numerous FAA manuals which detail the responsibilities of the TMUs and ATCSCC. This paper differs from FAA literature in the following ways:

- it discusses at a general level the nature and causes of congestion,

- it reviews the FM responsibilities and strategies as they are applied within the hierarchical framework of the ATM system, including the inter-relationships among facilities, and

- it describes FM strategies on the basis of level of severity, i.e., the paper will cover the different strategies as they are employed relative to the severity of the actual or anticipated congestion within the system.

38
2.2.1. Air Traffic Congestion

Three air traffic resources experience congestion: airports, airspace sectors, and fixes. Congestion occurs when demand for a resource exceeds capacity. A constrained resource implies that the resource is operating at maximum capacity and unable to accept further demand.

There are multiple negative effects of congestion. Among them are high controller workloads, excessive flight delays, safety concerns, and costly, unnecessary fuel consumption. Traffic management specialists attempt to avoid congestion by implementing FM in order to maintain safe and efficient flow rates throughout the system. An understanding of the dynamic nature of congestion, requires insight into the capacity and interdependencies of air traffic resources, and characteristics of air traffic demand.

2.2.1.1. Resource Capacity

Airports, sectors and fixes are capacitated. An airport’s capacity refers to the rate of aircraft operations, i.e., takeoffs and landings, the airport can sustain over a period of time. Realtime airport capacity is determined by the separation intervals of arriving and departing flights implemented by controllers or pilots. For an airport to operate at full capacity, controllers and pilots must adhere to the minimum separation standards given weather conditions and runway layouts. However, pilots and controllers may choose to increase aircraft separation distances for concerns of safety or workload.

Airports experience both departure and arrival congestion. Departure congestion occurs if departure demand exceeds the maximum departure rate, resulting in departure queues on the airport taxiways. Arrival congestion occurs if the arrival rate is insufficient for arrival demand. Typically, arrival congestion will be translated to en route sectors outside of a terminal area, in order to maintain safe traffic levels in the immediate airport area.

A sector’s capacity is the total number of aircraft (aircraft count) that a sector can accommodate for a defined period of time. The maximum capacity for a sector is determined by FAA capacity specialists based on aircraft separation standards, route structures, weather, equipment availability (e.g., radars), and controller staffing and experience levels. A sector’s maximum capacity varies according to changes in the dynamic variables, e.g., weather. The motivation for determining maximum capacities is based upon the need to maintain operationally safe controller workloads and safe separation distances between aircraft. Although physical capacity within a sector may exist, a controller may deny a flight the permission to enter in order to maintain a reasonable workload.
Chapter 2: ATM Primer

The capacity of a fix refers to the maximum allowable flow rate of aircraft over a fix. If a fix reaches its full capacity, it is referred to as saturated or overloaded. Fixes along routes entering terminal areas typically become saturated if airport capacity becomes constrained. Traffic management specialists at Centers are encouraged to attempt fix balancing, the process of dispersing arriving traffic over multiple arrival fixes in order to avoid heavy congestion in specific sectors.

To assist controllers in achieving a flow rate over a fix, most fixes have an associated holding pattern, where aircraft are delayed, if necessary, prior to crossing the fix. Delayed flights are cleared onto elliptical holding patterns at different altitudes. As capacity becomes available, flights will be cleared off the holding patterns and over the fix.

2.2.1.2. Capacity Uncertainty

A traffic management specialist considers both predicted and realtime capacity of a resource. Predicted capacity is the basis for the FM planning process. The accuracy and leadtimes of predictions are critical elements of the planning process.

The uncertainty in predicting a resources capacity is primarily due to the influence of weather on air traffic operations. Weather conditions such as cloud ceiling height, visibility, wind speed, wind direction and precipitation are the principal factors affecting an airport’s ability to support demand. Low visibility in the terminal area, requires increased aircraft spacing intervals, which decreases capacity. As a result of the high degree of uncertainty in forecasting weather conditions, resource capacities can be difficult to predict even a few hours in advance. Traffic management specialists work closely with weather personnel when projecting capacity forecasts prior to implementing FM initiatives.

If a prediction underestimates a resource’s capacity, the resource will possibly be “starved” of aircraft in the future. This will occur due to unnecessary actions initiated to avoid exceeding the inaccurately predicted, low level of capacity. Conversely, if a resource’s capacity is overestimated, the resource may become constrained requiring realtime actions, usually airborne delays.

A tradeoff exists between the accuracy of a weather forecast and a reasonable leadtime. Effective FM requires that traffic managers have some advance warning of potential resource overcrowding. Perfect knowledge of the status of the system even one hour in advance may be insufficient leadtime for some airports or sectors to implement an FM action to relieve congestion.
2.2.1.3. Resource Interdependencies

Even with perfect predictions, congestion may be unavoidable due to resource interdependencies. If one resource is identified as facing potential congestion, a traffic management specialist must identify an alternative resource, or resources, to hold or absorb demand. If alternative capacity does not readily exist, congestion may ripple through the system. The following simple, three sector example in Figures 2.15 and 2.16 illustrates this point.

![Diagram showing resource interdependencies](image)

**Figure 2.15: Resource Interdependencies - Sufficient Capacity**

*Sector count* refers to the current number of aircraft within a sector. The arrows represent the desired trajectories of a group of aircraft. In Figure 2.15, Sector 3 can only receive 2 of the four flights from Sector 2. Sector 2 can hold the remaining two flights, and, at the same time, accept the five flights inbound from Sector 1. Although Sector 3 becomes constrained, there is sufficient capacity in Sector 2 to absorb demand.

Figure 2.16 represents a different scenario.

![Diagram showing resource interdependencies](image)

**Figure 2.16: Resource Interdependencies - Insufficient Capacity**

In Figure 2.16, there is a count of 14 aircraft in Sector 2 instead of 10. As before, Sector 2 must hold two of the four flights bound for Sector 3. At this point, insufficient capacity exists in Sector 2 to manage the arriving traffic of five flights. Constrained capacity in Sector 3 has caused congestion to ripple into Sector 1. In actuality, this effect may occur on a larger scale among airports, fixes, and sectors.

Another interdependency within the system is the relationship between the arrival capacity and departure capacity at an airport. Within a given runway configuration, there may
be a tradeoff between the number of arrivals an airport can accommodate versus the number of departures. If departure congestion at an airport becomes excessive, a tower may choose to favor departures over arrivals to “clear out” a ground departure queue.

2.2.1.4. Demand Uncertainty

Demand on airports, airspace sectors and fixes is also characterized by uncertainty. Unannounced VFR traffic is one contributing factor. The first indication of a VFR flight’s intention to enter a terminal area, may be an in-flight contact by the pilot requesting clearance. Towers may deny clearances to VFR traffic during periods of heavy IFR traffic. General aviation traffic under IFR may also arrive unannounced into a congested area (a “pop-up”) if the flight plan was not introduced into the system with sufficient leadtime.

Uncertainty, however, is not restricted to VFR or General Aviation traffic. The components of an IFR flight’s trajectory can each experience degrees of uncertainty which affects the future demand the flight will place on resources. For example, departure congestion may result in a delayed flight departure, shifting the schedule of a flight for its entire trajectory. The aircraft will reach fixes, sectors and the arrival airport at unscheduled times. The en route flight time is uncertain as well, being influenced by the prevailing jet stream conditions. A high velocity jet stream may cause an early arrival for flights traveling west to east, or a late arrival for flights traveling east to west. In both scenarios, demand projections become uncertain.

Further uncertainty is introduced into demand by realtime airline practices. Airlines, not bound by published schedules, react to the daily air traffic environment (weather, delays, etc.) and the air travel market. Airlines may adjust flight times or cancel flights based on excessive delays, low passenger demand, crew availability, equipment problems and/or scheduling constraints.

2.2.2. Flow Management

To counter overcrowding and air traffic congestion, the FAA supports air travelers with FM. The goal of FM is to maintain safe and efficient rates of aircraft flow throughout the entire system. The FAA approach to managing demand and implementing FM is through the Air Traffic Management (ATM) hierarchy. The roles of the FAA facilities in the hierarchy are discussed in the first section of this chapter. Air traffic is managed at the appropriate level of hierarchical authority, progressing from decentralized control to centralized management as problems increase in scope and/or region. ATM deviates from a traditional hierarchy in that
often, guidance and directives from higher levels are detailed and flight specific, instead of general.

Figure 2.17 illustrates the relationships among the different resources and the responsible agency or person for traffic flow at each level. Within the hierarchical structure, flow management problems are handled at the lowest level possible. If a flow problem can be isolated to a specific region and handled by the region’s respective controlling agency (on the right-hand side of the chart), appropriate action is taken at that level. A problem will be elevated to a higher level if the regional resources are insufficient or if a traffic problem within a region will ripple into other resources.

\[ \text{CONUS Airspace} \]

\[ \text{ARTCC} \]

\[ \text{ARTCC Area} \]

\[ \text{ARTCC Area} \]

\[ \text{Terminal Area} \]

\[ \text{Terminal Area} \]

\[ \text{Airport Traffic Area} \]

\[ \text{Airfield} \]

\[ \text{Local Area} \]

\[ \text{Terminal Area Sectors} \]

\[ \text{En Route Sectors} \]

\[ \text{Controllers} \]

\[ \text{TRACON TMU} \]

\[ \text{ARTCC TMU} \]

\[ \text{ATCSCC} \]

Figure 2.17: Airspace and ATM Hierarchy

In Figure 2.17, horizontal lines connecting same-level resources indicate provisions or procedures which allow the entities responsible for two or more same-level resources to handle a traffic flow problem without elevating control to the next higher level. The term airfield refers to
the physical airport and surrounding ground facilities. The Local Area refers to the airspace surrounding an airport, extending out to 5 nautical miles and 2,500 feet in altitude, i.e., Class D airspace.

As problems are elevated, the facility or controller at the lower level(s) remains in the loop of the FM decision process. This is possible through the extensive communication and data availability among TMUs, TRACONs, and ATCSCC. The capability exists to establish immediate conference sessions with multiple parties for fast decision making.

### 2.2.2.1. Terminal Area

Within the terminal area, there are very few opportunities for long-lead forecasting and strategic flow management initiatives. Under normal circumstances, once inside a terminal area, an arriving flight will touch down within 20 minutes. Tower and TRACON efforts are focused primarily on maximizing the use of the terminal area airspace and runways, in view of realtime demand.

Strategically, TRACON TMUs explore alternative methods of prioritizing flights arriving into the terminal area and designing patterns that promote efficient traffic flow. Efficiency is gained by dedicating specific runways and queues to certain types of aircraft, implementing timely runway configuration changes, and clearing flights for visual approaches, when possible.

### 2.2.2.2. ARTCC

The principal FM concerns of Center TMUs are the traffic flows into and out of terminal areas. If at a given airport, arrival capacity is deemed sufficient to handle the projected demand, there is no anticipated congestion. In this situation, the Center TMU monitors traffic into the airport terminal area with minimal adjustments to flight itineraries. However, if an airport's capacity is forecast to be constrained, the respective Center TMU becomes proactive. Assuming the airport is operating at maximum capacity for the given weather conditions, the TMU’s objective is to adjust arriving traffic so that flights enter the airport terminal area at a rate that matches capacity. The FM strategies applied to avoid or relieve airport congestion are also applicable to fix and sector congestion.

The following options are available to Center TMUs which mostly apply to constrained arrival capacity.
• **Miles-in-Trail**

Miles-in-Trail (MIT) refers to a minimum longitudinal distance between two aircraft along routes bound for a constrained airport or congested airspace. TMU personnel may restrict aircraft to MIT in order to increase the interarrival times of aircraft, thus reducing demand. The advantage of MIT is that it requires no additional manpower since it is implemented by en route sector controllers already on duty. When applied to an airport experiencing moderately constrained, yet stable capacity, MIT is very effective.

However, if an airport were suddenly to experience an increase in capacity while under MIT restrictions, controllers would not be able to take advantage of the available capacity, due to the aircraft spacing intervals. Moreover, the MIT arrival queue may eventually extend geographically and cause disruptions and delays at departing airports. Aircraft may be held at a departing airport to meet the MIT restrictions of a destination airport hundreds of miles away.

MIT is a common strategy used to control inter-Center traffic. Rather than elevating a problem to the national level, one Center TMU may request a neighboring Center TMU to initiate MIT restrictions for all traffic bound for the problem Center. ATCSCC personnel in Washington D.C. are nonetheless informed if such an action is initiated.

• **Minutes-in-Trail**

Minutes-in-Trail also refers to the longitudinal separation among aircraft destined for a constrained resource. Minutes-in-Trail is the most common method of implementing longitudinal separation. It is a safe method of controlling a queue with various types of aircraft traveling at different speeds. The benefits and drawbacks of Minutes-in-Trail are similar to Miles-in-Trail.

• **Rerouting**

Rerouting is primarily a method of redirecting en route traffic onto alternate en route airways around severe weather conditions or overloaded en route sectors. Rerouting requires aircraft to deviate from preferred routes between airports or fixes. TMUs attempt to determine in advance, the full impact of rerouting as well as the time when the system is expected to return to normal.

Rerouting due to severe weather conditions is more complicated than rerouting for sector workload problems. The difficulty lies in the uncertainty of weather predictions and the movement of an inclement weather pattern. TMUs working with weather specialists must determine which routes will be available for use during severe weather conditions, accounting for meteorological changes. TMUs seek to reopen weather-closed routes as soon as it is safely
possible. This entails sending *pathfinder* flights on previously closed airways in order to re-establish preferred routing. Severe en route weather problems affecting multiple Center areas are usually elevated to ATCSCC.

- **Arrival Metering**

   Arrival metering is the process of regulating the flow of traffic into a terminal area. The process involves clearing flights over arrival fixes at a rate that matches the capacity of the destination airport. As flights reach outer fixes, controllers calculate cross times for arrival fixes that will ensure desired rates of flow. The calculations may be performed manually or with the Arrival Spacing Program (ASP) available at selected Centers. Controllers delay flights as necessary to meet the desired rate, using holding patterns, or other speed-control techniques.

   In ATM, any delay above 15 minutes is reportable. This threshold allows controllers and traffic management specialists to tolerate delays below 15 minutes. From an en route controllers perspective, it is possible to informally initiate limited airborne holding on flights entering a terminal area, if delays can be maintained below 15 minutes. If delays begin to exceed the 15 minute level, alternative formal FM strategies are initiated in connection with arrival metering.

   Typically, arrival metering does not involve all arriving traffic. At Boston, for example, certain runways are too short for larger aircraft, and only available for smaller commuter type aircraft. For this reason, aircraft capable of executing short landings may be exempt from metering, if a shorter runway is active.

   Finally, arrival spacing provides a TMU with an "inventory" of aircraft in airborne holds. If capacity were to become suddenly available, the TMU may fill the available landing slots with aircraft in holds. Conversely, if capacity drops without warning, the TMU may adequately manage the demand by assigning limited airborne holds in conjunction with arrival spacing. A drawback for arrival spacing is that it requires a full-time traffic specialist, the Arrival Spacing Coordinator, on duty at the TMU.

- **Controlled Departure Times (Local)**

   Controlled Departure Times (CDTs) at the Center-level are assigned using local Ground Delay Programs (GDPs) which adjust flight itineraries prior to departure. In order to control demand at a constrained airport, a Center TMU contacts the airport towers located in the Center area, and instructs tower controllers to delay the departure times of specific flights destined for the low capacity airport. For example, at the Boston Center in Nashua, NH, if arrival spacing cannot adequately disperse demand with holds of less than 15 minutes, the normal progression is to initiate a GDP.
GDPs require projecting airport capacity and demand at the 15 minute interval. All flights are ordered and assigned a first scheduled-first served priority within 15 minute intervals. Beginning with the first constrained interval, the GDP identifies the arrival flights that cause demand to exceed capacity. These flights are assigned departure ground delays that will shift their arrival times into the succeeding 15 minute interval. Obviously, a GDP only considers flights that have yet to depart. The process is repeated for the next interval, considering both scheduled flights, and flights that were shifted from the previous interval with ground delays. Previously shifted flights are prioritized over scheduled flights. The GDP continues until capacity is no longer constrained. The scheduled departure time of a flight combined with a GDP-assigned ground delay produces a CDT.

CDTs are most effective when traffic management specialists issue ground delays to towers in advance of the constrained time intervals. Otherwise, flights are already airborne and bound for the problem airport, thus unavailable for ground delays. For Boston Center, 2 hours is sufficient leadtime to implement CDTs for flights departing within the Boston Center boundaries. The major benefit of ground delay programs is that delays are transferred from the air to the ground, avoiding potential airborne congestion and reducing en route controller workload. Assigning CDTs is the safest FM strategy, although it requires continuous monitoring and adjustment, if actual capacity or demand varies from predicted levels.

- **Traffic Stops**

  Traffic stops indefinitely hold all flights bound for a problem airport on the ground, hence a more severe form of CDTs. Traffic stops are initiated when critical events occur at an airport. For example, a severe thunder squall with high winds may reduce airport capacity to zero.

- **Departure Spacing**

  Departure spacing involves separating departing aircraft at fixed intervals, usually due to congestion in the terminal area or immediate en route airspace. The FAA has specific guidelines for departure spacing based on weather conditions and runway layouts.

  CDTs and Traffic Stops are the two most severe options available to Center TMUs when managing a constrained resource. The major limiting factor for CDTs and Traffic Stops is that Centers only control the departure times (and hence arrival times) of flights originating within their respective boundaries. If CDTs or Traffic Stops are insufficient in relieving congestion a Center will elevate the problem to ATCSCC.
Combinations of two or more of the options are possible. The process of selecting the appropriate strategy depends on the characteristics of the problem, i.e., long-term or temporary problem and severe or mild. Two other influencing factors are the on-duty staff level at the TMU and controller experience levels.

2.2.2.3. ATCSCC

The Air Traffic Control System Command Center (ATCSCC) oversees all air traffic in the United States. Staffed with controllers and personnel from various sites across the country, ATCSCC coordinates national flow management initiatives and general air traffic activities. ATCSCC controllers rarely interact directly with flights, but rather exert influence on the aggregate flow of aircraft in the system. National flow management actions are initiated at ATCSCC and transmitted to the appropriate Centers, TRACONs or Towers.

This section will discuss the organization of ATCSCC and the principal FM initiative of ATCSCC, Ground Delay Programs. The conclusion of the section discusses the effect of FM on airline operations.

2.2.2.3.1. Organization

Within ATCSCC, located in the FAA building in Washington D.C., there are five operating areas:

- Eastern Complex
- Central Complex
- Western Complex
- Severe Weather/National Route Management
- Special Traffic Management

The function of the Eastern, Central, and Western complexes is to monitor and assist flow management in the eastern, central, and western United States. The complexes are manned by controllers, recruited from various FAA facilities across the country. Using the ETMS and FM applications, controllers track capacity and demand levels at air traffic resources, coordinate flow initiatives among Center TMUs, and, if necessary, administer national FM procedures. The procedures are fundamentally extensions of Center-level processes, i.e., in-trail restrictions, reroutes, ground delay programs and traffic stops.

The Severe Weather/National Route Management division manages severe national weather problems and other large-scale crisis scenarios that disrupt the flow of traffic, requiring adjustments in demand and resource utilization. The principal process which reroutes traffic
around adverse weather is the Severe Weather Avoidance Program (SWAP). Although implemented at the tactical level in response to real-time adverse weather conditions, SWAP is a strategically designed program. Traffic management specialists continuously research and develop alternative routes for reoccurring weather problems.

The National Route Management division is staffed with personnel that coordinate and review requests by aircraft operators to deviate from preferred routes between airport pairs. Requests are typically received from airlines seeking routes at altitudes with more favorable prevailing wind speeds. National Route Management personnel review and transmit proposed flight trajectories to Centers managing the sectors affected by the changed itineraries. If any Center indicates that a specific flight change will adversely affect sector or airport traffic levels, ATCSCC disapproves the request.

Finally, the Special Traffic Management Office handles flight plan coordination and approval at the national level. Much of this function is automated.

2.2.2.3.2. Ground Delay Programs

A Ground Delay Program (GDP) is an automated, air traffic flow management process available to controllers at ATCSCC. GDPs coordinate, on the national level, the departure times of flights destined for an airport projected to have insufficient arrival capacity (the "problem airport"). The purpose is to maintain operationally acceptable traffic levels, reducing airborne holding in the airspace surrounding the problem airport, if possible. In principle, it resembles local (Center level) GDPs, however, the scope of a national GDP extends to all flights in the CONUS, if necessary. Whenever possible, efforts are made to distribute ground delays in an equitable fashion amongst the airlines and other traffic.

The process leading to the administering of an ATCSCC GDP usually involves attempts at the Center level to manage anticipated or actual arrival congestion. If a Center implements a local GDP or another inter-Center metering strategy, the action is coordinated with ATCSCC. The problem is elevated to ATCSCC, if it is determined that capacity constraints or demand levels at an airport warrant a national GDP. On occasion, it is immediately recognized that a problem will be too large to be managed at the Center level, at which point, ATCSCC may become involved at the onset. The discussion to follow on GDPs will cover the two types of GDPs (select and general), the process of implementing a program, and follow-on programs.
• **Select Program**

A select GDP identifies specific flights for delays based on a unique condition at the problem airport. Select programs are employed when sufficient leadtime exists to design the program input parameters around the problem airport condition. The ground delays assigned by a select GDP combined with the scheduled departure times produce new release times for the selected flights, known as Controlled Departure Times (CDTs).

Select programs are rarely used because they require analyzing congestion and flight delays at the individual flight level. The most common use of select programs involves the Atlanta airport. Atlanta routinely experiences fix saturation at the northeast and southeast arrival fixes. In reaction to this recurring problem, ATCSCC runs select programs that target flights that intend to enter the Atlanta terminal area via these fixes, while exempting other arriving flights.

• **General Program**

General programs are more common. A general GDP may consider all national flights destined for the problem airport as candidates for ground delays. However, if ATCSCC controllers desire to limit the range of flights affected by GDP delays, general programs may be tailored to exclude flights from certain regions or airlines.

The output of a general GDP relates to groups of flights by arrival time at the problem airport. Flights are grouped together by 15 minute time intervals and each time interval is assigned a delay factor by the ground delay program. Delay factors, when applied to individual flights, produce Expected Departure Clearance Times (EDCT) for each flight. An output for a one-hour GDP (9:00 to 10:00 AM) may resemble the results in Table 2.1.

<table>
<thead>
<tr>
<th>Estimated Time of Arrival (ETA)</th>
<th>Delay Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>0900/0914</td>
<td>16</td>
</tr>
<tr>
<td>0915/0929</td>
<td>21</td>
</tr>
<tr>
<td>0930/0944</td>
<td>22</td>
</tr>
<tr>
<td>0945/0959</td>
<td>27</td>
</tr>
</tbody>
</table>

In Table 2.1, all flights scheduled to arrive at the problem airport between 9:15 and 9:29 AM would receive a 21 minute ground delay prior to departure.
• **Implementation Process**

The implementation process for both types of delay programs, select and general, are similar. Figure 2.18 displays the ground delay program process. Following the diagram is a description of the different steps and decisions involved in the implementation.

![Figure 2.18: Ground Delay Program Implementation Cycle](image)

**Step 1.** Future airport capacity is projected by forecasting an acceptance rate for the problem airport. This function, as well as demand forecasting, is performed through conference sessions among ATCSCC and Center, TRACON, and/or tower personnel at the problem airport. Capacity estimates reflect the anticipated severity of the constraining problem. The forecasted flow rate may vary hourly.

Future demand is predicted based on OAG schedule data in ETMS for the problem airport. The OAG flight schedules only contain commercial air traffic data, therefore, VFR and GA traffic must be estimated from historical experience. Moreover, the OAG data are somewhat uncertain. Airlines adjust schedules in realtime in order to react to daily weather conditions, maintenance problems or passenger demand levels. When an airline submits a flight plan for a proposed flight, the schedule information is updated in ETMS. However, if a flight plan is not submitted early enough, the updated information may not be available to FM planners.

**Step 2.** The **time interval** specifies the range of time over which the problem airport is expected to experience constrained capacity. The interval is determined by a **start time** and a **stop time**, and reflects the anticipated persistence of the constraining problem, e.g., weather or radar
failure. Controllers often pad the start time in advance of the actual predicted capacity problem. This strategy provides en route and terminal area controllers managing the airspace surrounding the problem airport with a smooth decrease in traffic levels, when transitioning into the actual interval of constrained capacity. The implementation time refers to the time when the program is put into effect, usually in advance of the start time. Figure 2.19 illustrates the critical time events for a GDP on a timeline.

![Ground Delay Program Timeline](image)

**Figure 2.19: Ground Delay Program Timeline**

The GDP leadtime is the difference between the implementation time and start time. Leadtimes are critical. In an extreme case, an airport may experience severe capacity reductions with no warning. A program may be implemented with a start time that matches the implementation time, thus having no leadtime. As a result, many inbound flights will have previously departed. If this occurs, the effectiveness of the program for the initial hours of the constrained time interval will be limited by the few available flights, bound for the problem airport, remaining on the ground. Only flights departing airports that neighbor the problem airport will be available for ground delays at the onset of the constrained time interval. In this situation, airborne holding techniques will most likely be required to control the arrival rate of flights that had previously departed.

Leadtime is related to the certainty of the constraining problem. Assuming there is a highly predictable, but severe weather pattern anticipated in the vicinity of an airport, traffic management specialists will have a long leadtime (*long* being in excess of four hours for most CONUS airports). A program initiated four hours in advance of the start time will have access to a high percentage of candidate flights bound for the problem airport. Candidate flights will have yet to depart their departure airports, and will be therefore available, or candidates, for ground delays. Figure 2.20 displays approximate leadtimes required to assign ground delays to flights bound for St. Louis. The map in Figure 2.20 indicates that in order to assign a ground delay to a flight leaving Boston, the program must be run with approximately 3 hours of leadtime.
If the certainty of a constraining problem is very high, ATCSCC will most likely implement the program as soon as possible, maximizing the leadtime and availability of candidate flights.

On the other hand, if a decrease in capacity is anticipated, but with uncertainty, there is a risk associated with implementing a program too soon. If a program is implemented and the capacity decrease does not materialize, ATCSCC will have unnecessarily held flights on the ground to the displeasure of the airlines and passengers. From a traffic management viewpoint, this constitutes an inefficient use of airport capacity and may result in derivative problems later in the day. The decision may be made to delay the implementation of a program until traffic management specialists gain more confidence in the forecast of the constraining problem. ATCSCC weighs the tradeoffs between numbers of candidate flights versus more accurate forecasts when determining a program leadtime.

Step 3. The remaining program inputs are:

- *flight information and status:* Flight information (OAG schedules and current flight plan data) and flight status (i.e. airborne, preflight, previously delayed, canceled etc.) are provided by ETMS.

- *maximum ground delays:* This parameter sets the upper limit on the length of ground delays, restricting the program from assigning excessive ground delays to
individual flights. Often a program is initially run with unlimited maximum ground delay. In this manner, this parameter exerts no influence on the GDP. If some flights receive disproportionately high ground delays, the program will be re-run with an upper limit.

- **Center regions:** Specialists may choose to limit the distribution of ground delays to flights departing from within specific Center boundaries. Those Centers, adjacent to the Center area with the problem airport, are referred to as first tier Centers. If the specialist selects only first tier Centers, the program is referred to as a first tier program. Similarly, a second tier program includes first tier Centers and all Centers adjacent to first tier Centers. A full program includes all Centers.

- **General Aviation (GA) factor:** The GA factor is an estimate of the anticipated hourly GA traffic during the constrained time interval. This input reserves a portion of the landing slots for unscheduled GA.

After defining all program parameters, a controller runs the GDP. The objective of the GDP algorithm is to distribute demand at the problem airport in order to maintain a flow rate at or below capacity during the entire constrained time interval. The mechanics of the GDP algorithm resemble the local GDP process. All flights are sorted by arrival time within discretized time segments over the entire constrained interval. Using a first-scheduled first-serve priority basis, the GDP algorithm selects flights for ground delays that cause projected demand to exceed projected capacity for a given segment. The process is repeated for every time segment, where for any given segment, flights with previously assigned delays are given priority over flights with no delays.

Flights may be delayed beyond the stop time of a program. If so, the number of flights delayed beyond the stop time constitutes the *stack value*, which serves as an input for a Follow-on Program. Follow-on Programs will be discussed in the next section.

**Step 4.** The output of a program is ground delays. The delays are analyzed statistically prior to implementation. The process by which a traffic management specialist reviews the output of a program is significantly enhanced by the data processing, sorting, and filtering features built into the GDP application.

The following results pertain to a general program that was run on the NAS computers at ATCSCC in Washington D.C., on October 13, 1992. The scenario was hypothetical with Denver Stapleton Airport as the problem airport. The inputs for the program were:
Start Time: 1500 (3:00 PM) GMT
Stop Time: 1800 (6:00 PM) GMT
Arrival Rate:
   1500 - 1559:  50 flights
   1600 - 1659:  45 flights
   1700 - 1759:  60 flights
Maximum Delay: 999 minutes
Stack Value: 0
GA Factor: 2
General Program, Include All Centers

Setting the maximum ground delay parameter to 999 minutes is the equivalent of no upper limit. The program was run with a leadtime of less than one hour.

The aggregate results for the Denver program were:

  156 flights predicted to arrive
  73 flights require ground delays
  Total minutes of delay: 3940
  Average delay: 25 minutes
  Maximum delay: 40 minutes

By analyzing the aggregate results, it is possible for an ATCSCC controller to ascertain the necessity of the program. If the average delay figure were low (below 15 minutes) and the number of flights delayed were insignificant, the program may not be unnecessary. If this occurs, ATCSCC may contact the Center and recommend that local efforts be continued.

In the Denver example, it appears the program is valid. Approximately half of the expected arrivals receive delays and the average delay figure is substantial. Once the validity of a program is established, ATCSCC analyzes the proportion of delayed flights by departure Center and by airline. Departure Center refers to the Center region that contains the departure airport for a flight. Tables 2.2 presents the distribution of delayed flights for the Denver example by departure Center.
Table 2.2: Ground Delay Program Results By Departure Center

<table>
<thead>
<tr>
<th>Center Name</th>
<th>Predicted Arrivals</th>
<th>% of Total Arrivals</th>
<th>Flights Delayed</th>
<th>% of Total Flights Delayed</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZAB</td>
<td>11</td>
<td>7.1%</td>
<td>6</td>
<td>8.2%</td>
</tr>
<tr>
<td>ZAU</td>
<td>6</td>
<td>3.8%</td>
<td>3</td>
<td>4.1%</td>
</tr>
<tr>
<td>ZBW</td>
<td>2</td>
<td>1.3%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>ZDC</td>
<td>3</td>
<td>1.9%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>ZDV</td>
<td>37</td>
<td>23.7%</td>
<td>33</td>
<td>45.2%</td>
</tr>
<tr>
<td>ZFW</td>
<td>7</td>
<td>4.5%</td>
<td>6</td>
<td>8.2%</td>
</tr>
<tr>
<td>ZHU</td>
<td>5</td>
<td>3.2%</td>
<td>1</td>
<td>1.4%</td>
</tr>
<tr>
<td>ZID</td>
<td>2</td>
<td>1.3%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>ZJK</td>
<td>2</td>
<td>1.3%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>ZKC</td>
<td>5</td>
<td>3.2%</td>
<td>4</td>
<td>5.5%</td>
</tr>
<tr>
<td>ZLA</td>
<td>21</td>
<td>13.5%</td>
<td>8</td>
<td>11.0%</td>
</tr>
<tr>
<td>ZLC</td>
<td>9</td>
<td>5.8%</td>
<td>2</td>
<td>2.7%</td>
</tr>
<tr>
<td>ZMA</td>
<td>1</td>
<td>0.6%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>ZME</td>
<td>2</td>
<td>1.3%</td>
<td>2</td>
<td>2.7%</td>
</tr>
<tr>
<td>ZMP</td>
<td>8</td>
<td>5.1%</td>
<td>6</td>
<td>8.2%</td>
</tr>
<tr>
<td>ZMY</td>
<td>8</td>
<td>5.1%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>ZOA</td>
<td>12</td>
<td>7.7%</td>
<td>1</td>
<td>1.4%</td>
</tr>
<tr>
<td>ZOB</td>
<td>5</td>
<td>3.2%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>ZSE</td>
<td>7</td>
<td>4.5%</td>
<td>1</td>
<td>1.4%</td>
</tr>
<tr>
<td>ZTL</td>
<td>3</td>
<td>1.9%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>156</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>73</strong></td>
<td><strong>100.0%</strong></td>
</tr>
</tbody>
</table>

The length of ground delays is not reflected in Table 2.2. If a flight receives a ground delay, it is counted in the “Flights Delayed” column, regardless of the length of the delay.

It is not unexpected to discover that the departure Centers most affected by the program are the Denver Center (ZDV) and the neighboring Los Angeles Center (ZLA). This is partly influenced by the fact that the program had minimal leadtime and therefore, only flights with short flight times were candidates for ground delays. Additionally, a high proportion of most traffic into an airport originates from same-Center or neighboring Center airports.

A review of delays by Center is conducted to find flights from departure Centers that minimally impact the results of the program and may be exempt from delays. In the example, there is only one flight delayed from the Seattle Center (ZSE). It is probable that the ZSE area can be excluded from the program without significantly impacting the overall program. By excluding low-impact departure Centers, ATCSCC attempts to reduce unnecessary delays to individual flights. Furthermore, Centers with low numbers of flights receiving delays are often far from the Center of the problem airport. Because of the uncertainty in actual flight times, a ground delay may be an unnecessary adjustment to a flight with a long flight time. If the Denver scenario were real, the program would most likely be re-run, excluding ZSE and other low-impact departure Centers.
Table 3.3 contains the distribution of delayed flights by airline.

<table>
<thead>
<tr>
<th>Airline Name</th>
<th>Total Aircraft</th>
<th>Percent of Total</th>
<th>Total Delayed</th>
<th>Percent of Total Delayed</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAL</td>
<td>4</td>
<td>2.6%</td>
<td>3</td>
<td>4.1%</td>
</tr>
<tr>
<td>AWE</td>
<td>1</td>
<td>0.6%</td>
<td>1</td>
<td>1.4%</td>
</tr>
<tr>
<td>BTA</td>
<td>17</td>
<td>10.9%</td>
<td>13</td>
<td>17.8%</td>
</tr>
<tr>
<td>COA</td>
<td>45</td>
<td>28.8%</td>
<td>17</td>
<td>23.3%</td>
</tr>
<tr>
<td>DAL</td>
<td>3</td>
<td>1.9%</td>
<td>1</td>
<td>1.4%</td>
</tr>
<tr>
<td>GLA</td>
<td>2</td>
<td>1.3%</td>
<td>1</td>
<td>1.4%</td>
</tr>
<tr>
<td>LN</td>
<td>1</td>
<td>0.6%</td>
<td>1</td>
<td>1.4%</td>
</tr>
<tr>
<td>MEP</td>
<td>1</td>
<td>0.6%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>SME</td>
<td>16</td>
<td>10.3%</td>
<td>15</td>
<td>20.5%</td>
</tr>
<tr>
<td>NWA</td>
<td>3</td>
<td>1.9%</td>
<td>1</td>
<td>1.4%</td>
</tr>
<tr>
<td>TWA</td>
<td>1</td>
<td>0.6%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>UAL</td>
<td>55</td>
<td>35.3%</td>
<td>18</td>
<td>24.7%</td>
</tr>
<tr>
<td>USA</td>
<td>2</td>
<td>1.3%</td>
<td>0</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Airline results are reviewed to determine if there are airlines which contribute minimally to the results of the program. If low-impact airlines exist, they may be excluded, and the program re-run. It is not surprising to discover the high percentages of Continental (COA) and United Airlines (UAL) flights affected by the program, because Denver is a major operating facility, i.e., a hub, for these airlines.

The GDP Implementation Cycle (Figure 2.18) demonstrates the different options available to a traffic management specialist after a review of GDP results. A specialist may redefine the departure Centers to be included in the assignment of delays (returning to Step 3). Or a specialist may modify the time interval (returning to Step 2). Also, a program may be re-run, if updated airport capacity forecasts are received (returning to Step 1). Thus, the GDP process is potentially dynamic and iterative.

**Step 5.** Assuming the program results are accepted, the implementation process transitions to an optional step. The ATCSCC planner may coordinate the program results with the airline or airlines most affected by the ground delays prior to implementation. The rationale for involving the airlines, is to help reduce the negative impact of delays on airline operations.

Typically, airlines schedule many connecting flights through hub airports, in order to create multiple connection opportunities for passengers. A ground delay on a single flight may result in derivative delays on subsequent flights of the same aircraft. Since airlines attempt to achieve successful hub operations, ATCSCC frequently transmits the proposed ground delays to the airlines most affected by a program, prior to implementation. Upon receipt of ground delays,
airline schedulers are permitted to adjust their schedules with flight substitutions and cancellations, without modifying the aggregate flow objectives of the GDP. An amended schedule is submitted to ATCSCC.

Airlines benefit from involvement in the assignment of ground delays. From an airline perspective, some connecting flights are more critical than others. A certain flight's aircraft may be scheduled for multiple flight legs later in the day, or a flight may have many passengers transferring to different connecting flights. Such a flight is more critical than a terminating flight.

A flight substitution involves transferring delays assigned to critical flights, to less critical flights, in order to minimize the overall expected delays or connection problems that will result from ground delays. Cancellations occur when an airline determines that a ground delay or another problem warrants eliminating a flight. Scheduled passengers will be transferred to other flights or perhaps other carriers.

The process of ground delay coordination with the airlines is presently under review by the FAA. Under the current guidelines, airlines may perform flight substitutions and cancellations based on the following criteria:

- Prior to any flight substitutions, an airline must perform a flight cancellation.

- Airlines may only substitute a flight, for example Flight F, for another flight, Flight G, if F's scheduled arrival time is within 10 minutes before G's arrival time or within 20 minutes after G's arrival time.

The proposed changes to the guidelines are:

- No flight cancellations are required.

- Flight F may only be substituted for Flight G if F's arrival time is within 20 minutes after G's arrival time.

- Airlines may rely on estimated arrival times rather than scheduled arrival times.

By using estimated arrival times, airlines will be able to modify their schedules based on more accurate, updated information. This is predicted to improve the efficiency of GDAs.

Because of the high return on effective flight substitutions, airlines have invested significant research in developing automated programs and algorithms to execute the substitution process in a timely and cost-effective manner. However, coordinating with the airlines is conditional upon the availability of adequate GDP leadtime.
Step 6. In Step 6, the program is implemented. Ground delays are transmitted simultaneously to Centers, towers, and airlines. Center TMUs are directed to monitor all flights with ATCSCC assigned ground delays and confirm that departure clearances correspond to the adjusted departure times. At the same time, Centers are to assign general aviation flights bound for the problem airport with delays that correspond to the delay factors of a program.

Step 7. Finally, as the problem airport enters the constrained time interval, ATCSCC monitors the progress of traffic at the problem airport. Every two hours, ATCSCC statistically reviews the efficacy of the distributed ground delays. The primary reason for poor performance of a program is error in the forecast capacity. As stated earlier, if airport capacity is underestimated, an airport will be “starved” due to unnecessary ground delays. Conversely, if airport capacity is overestimated, the airport will experience higher than expected congestion. ATCSCC may adjust subsequent departure times to correct dynamically for an ineffective program. The Ground Delay Program is described in detail in ATC Order 7210.3J (FAA 1992).

- Follow-On Programs

Due to uncertainties in capacity forecasts, traffic management specialists are reluctant to implement a program that exceeds 4 hours in length. However, at times constraining problems exceed 4 hours, e.g., adverse weather, which require follow-on programs.

The criteria for the implementation of follow-on programs resembles the framework for an initial program. In order to transition smoothly between programs, follow-on programs are implemented prior to the stop time of the initial program. Follow-on programs often benefit from more accurate capacity forecasts, since the constraining problem at an airport, as in a weather condition, will have materialized during the constrained interval of the initial program. A follow-on program may follow an initial program or another follow-on program. Figure 2.21 displays the time sequence of an initial program (GDP1) and a follow-on program (GDP2).

![Figure 2.21: Follow-on Program Timeline](image)

An input into a follow-on program is the stack value from the initial program. The stack value represents the number of flights delayed by the initial program past the stop time, and therefore, into the constrained time interval of the follow-on program. The GDP offsets the initial
hour of the follow-on program according to the stack value, in order to reserve arrival slots for flights that are carried over.

### 2.2.2.4. Airlines

Before concluding the discussion on air traffic flow management, it is necessary to comment on the subject from the perspective of the airlines. Clearly, flight delays concern service-oriented and profit-motivated airlines. Although they are aware that the ATCSCC, TRACONs and towers are attempting to manage the aggregate flow of air traffic in the network safely, airlines are nonetheless focused on their individual priorities. Three key airline priorities that are not considered in ATC flow management decisions are the effects of a flight delay on:

- flight interdependencies,
- banks of flights, and
- flight crew compliance with ATM regulations.

• **Flight Interdependencies**

A single aircraft typically executes several flight legs throughout the day, creating flight interdependencies. As a result, the delay of a particular flight may cause negative downstream effects in the form of delays of the same aircraft on subsequent flight legs. While the *turn time* of an aircraft (time required to refuel, execute passenger transfers and recenter the aircraft) varies, the average is approximately 30 minutes. Airlines accommodate this constraint by scheduling connecting flights with sufficient layovers, or *ground times*. However, if an aircraft's scheduled ground time is equal to or only slightly greater than the aircraft's turn time, even a small delay on an arriving flight may result in a late departure for the connecting flight, because of inadequate *real* ground time. On the other hand, a flight with a long layover (ground time greatly exceeds turn time) will be able to absorb a delay assigned to a previous flight leg of the same aircraft. For this reason, airlines look downstream and evaluate the ability of an aircraft to absorb a delay in the future when reviewing and accepting the Controlled Departure Times transmitted from ATCSCC.

• **Banks of flights**

The hub system adopted by the major carriers in the 1980s, further complicates the effect of flight delays. Under the hub system, an airline will schedule many flights to arrive into a hub airport within a specific interval of time. Such a group of flights is referred to as an *arrival bank*. Following a short time period to allow for passenger connections and aircraft preparation, the
airline schedules a *departure bank* of flights. This scheduling approach is a cost-effective method of creating multiple connection opportunities for passengers traveling in the related arrival and departure banks. However, a major weakness in the hub system is its sensitivity to flight delays. FM delays, e.g., ground delays or airborne holds, imposed on one or two flights in an arrival bank, may result in airline-imposed delays on multiple flights in the related departure bank. Decisions to delay departing flights are based on the percentages of connecting passengers on the late arriving flights and, as mentioned above, potential downstream effects. A high percentage of connecting passengers will make an airline more apt to hold departing aircraft.

- **Flight Crew Compliance**

  Finally, delays may also affect flight crews. Crews must abide by FAA regulations which state that the maximum allowable duty time is 12 hours before requiring crew rest. Unplanned lengthy delays may cause a crew to reach its maximum allowable duty time prior to completing all scheduled flights. When this occurs, airlines adjust crew schedules where possible; however, they may also react by canceling flights. In addition, crews do not always continue on the same aircraft when passing through an airport. A late arriving aircraft may hold up a different aircraft waiting for the crew.
Chapter 3

3. Airport Demand Schedule Modeling

This chapter describes the development of an algorithm for generating hypothetical, OAG\(^1\)-like airport demand schedules. The objective of the algorithm is to produce daily airport schedules that adhere to flight connectivity and arrival rate parameters for a set of user-specified airports. The flight schedules produced by the algorithm are statistically valid for air traffic flow management (FM) research. The algorithm is implemented in a computer application known as the Pseudo-OAG Generator or POAGG. The POAGG algorithm extends flight schedule generation research initiated by Savari [14].

The schedule modeling research required analysis of OAG schedules and interviews with airline scheduling personnel. These interviews were conducted in order to gain a general understanding of airline scheduling procedures. To a casual observer, the flight itineraries produced by the algorithm resemble actual airline flights; however, airline information in the schedules is approximate. While valid for FM research, the generated schedules are not designed to be implemented by the airlines.

This chapter is organized into three sections. The first section describes the motivation for the research and some characteristics of commercial air traffic scheduling. The second section presents the problem formulation and solution approach. The third section describes the implementation and test results. The concluding section discusses extensions to the model.

3.1. Motivation and Objectives

The development of a schedule generator is motivated by the need for airport demand schedules to support air traffic management research. The decision to model commercial demand is based on observations of the dominance of commercial traffic at major airports over other forms of air traffic. In addition, the current practices of forecasting demand within the FAA primarily involve commercial traffic.

\(^1\)OAG or Official Airline Guide is a monthly publication that lists all commercial flights.
3.1.1. Flow Management Analysis

The primary motivation for developing a multiple-airport schedule generator is to support research in air traffic FM. FM is the process of adjusting the itineraries of flights within the air traffic system due to actual or predicted congestion. The goal of FM is to relieve overcrowding at airports and airspace sectors. One FM technique is ground holding, which involves delaying the departure of selected flights bound for airports with insufficient landing capacity.

POAGG is designed to generate schedules that can be used as input schedule data in the analysis of FM techniques. With airport demand schedules, traffic management specialists can compare and evaluate the effectiveness of various FM strategies within a computer simulation of the air traffic environment. Through simulation and the schedule modeling algorithm, analysts may compare multiple techniques, experimenting with alternative scenarios of demand. Any change to the current processes of FM, such as ground holding, or the integration of a new strategy, may be rigorously analyzed prior to implementation in the actual national FAA control system.

The current source of demand schedules for FM analysis is OAG flight schedules. There are limitations in how much one can use OAG schedules, since they represent existing or historic air traffic demand. Given the predicted, ten-year 4% annual growth rate in air traffic demand levels, OAG schedules are insufficient for a simulation of the future of the air traffic environment. Furthermore, future changes in airline scheduling practices are not reflected in current or past OAG schedules. By offering the capability to generate hypothetical schedule scenarios that capture the growth in air traffic demand and the dynamic nature of airline scheduling, POAGG schedules could supplement OAG schedules as demand input for FM analysis.

As a secondary motivation, hypothetical schedules may be useful in the analysis of proposed air traffic capacity infrastructure enhancements. Given the long term and costly effort required to substantially increase an airport’s capacity, e.g., building a new runway, it is appropriate to measure the cost effectiveness of an enhancement against future demand levels.

3.1.2. Commercial versus Total Air Traffic Demand

Commercial air traffic is one of three types of air traffic demand. Commercial air traffic includes air carriers providing service between metropolitan areas, and commuter and air taxi

---

2FAA annual forecast for domestic traffic released April 1993.
services, consisting of shorter trips by mostly propeller driven aircraft. In addition, there is general aviation air traffic, which includes all private or corporate travel, and military air traffic. Figure 3.1 displays the 1992 airport demand profile based on “total itinerant\(^3\) airport operations” for the ten busiest airports in the United States [3].

![Pie chart showing airport demand by category]

**Figure 3.1: Airport Demand by Aviation Category**

Figure 3.1 demonstrates the dominance of commercial air traffic at busy airports. The preponderance of commercial traffic justifies initiating demand schedule modeling based on commercial demand. The rationale for presenting data for busy airports, is that busy airports routinely experience congestion. As a result, busy airports are the focus of FM and capacity improvements, the principal motivations for developing POAGG.

### 3.1.3. Enhanced Traffic Management System

Another motivation for modeling commercial demand (rather than general aviation or military air traffic) is derived from observations of current traffic management practices. Commercial air traffic demand, represented by OAG schedules, is the primary source of demand data available to the FAA. Traffic management specialists predict future demand on the system by integrating OAG schedule data, in electronic format, into FAA computers. The Enhanced Traffic Management System (ETMS), a network of FAA computers and displays dedicated to Air Traffic Management, enables traffic management specialists to selectively process flight data and project demand, based on 15 minute intervals, for each airport and airspace sector. A specialist can access flight data from ETMS for times up to 12 hours in the past and 30 days in the future.

---

\(^3\)An itinerant operation refers to a flight that exceeds a 20-mile radius around a single airport.
OAG flights are augmented in realtime as flight plans or flight cancellations are received by the FAA from the airlines. General aviation and military flights are rarely included in forecasts for flow management purposes. POAGG, therefore, attempts to generate demand data that replicate the current demand forecasting method.

3.1.4. Modeling Objectives

The general objective is to develop an approximate model of commercial air traffic scheduling. The basic characteristics of commercial air traffic schedules modeled in the POAGG algorithm are flight connections, hubbing operations, and shuttle services. These characteristics are the principal factors that influence, and are significantly affected by, FM actions and system delays. The modeling effort attempts to produce schedules that preserve these characteristics, in such a manner that hypothetical schedules generated by the algorithm are statistically indistinguishable from real OAG schedules, for the same airports and similar demand levels. Airline information for the hypothetical flight itineraries is modeled to enhance schedule realism, but does not reflect actual airline scheduling practices at a high fidelity.

The following listing is an example flight from the OAG.

<table>
<thead>
<tr>
<th>Departure</th>
<th>Arrival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Airport</td>
</tr>
<tr>
<td>6:24 AM</td>
<td>ATL</td>
</tr>
</tbody>
</table>

Analyzing OAG schedules provides some statistical insight into airline scheduling practices. For example, using OAG schedules, it is possible to ascertain the overall flight arrival and departure frequency distributions at airports of interest. Flight connection information, however, is not provided in the OAG schedules.

Connecting flights, hubbing operations and shuttle services are planned by scheduling groups within each airline. In order to protect a particularly efficient or effective scheduling technique, airlines do not publish detailed aircraft information in the schedules. Without aircraft information, OAG schedules are insufficient for tracking individual aircraft through the system. In certain situations, it is possible to infer which flights are scheduled with the same aircraft by pairing flights with the same flight number. Often however, aircraft change flight numbers throughout the day.

- Connecting Flights

Airlines routinely schedule an aircraft to perform multiple flight legs throughout the day, creating an interdependency between an incoming flight and the departing flight of the same
aircraft. The departing flight is referred to as the connecting flight. Layover time refers to the scheduled time interval between the arrival time of an incoming flight and departure time of its connecting flight. Due to flight interdependencies caused by connections, congestion delays often propagate to connecting flights throughout the day. In the same manner, flow management actions by the FAA, may also influence connecting flights. A ground hold on one flight may result in derivative delays on subsequent flights of the same aircraft.

* Hubbing Systems

The ripple effect of delays due to flight interdependencies is exacerbated by the airline hubbing system. A hub airport is an airline's central operating location within a geographical region. Major airlines have 3 to 4 hubs. At a hub airport, an airline will typically schedule 30 to 40 incoming flights within a specific interval of time. This grouping of flights constitutes an arrival bank. Following a sufficient period of time to allow for passenger connections and aircraft preparation, the airline schedules a departure bank. This scheduling practice is an efficient use of available aircraft and creates multiple connection opportunities for passengers traveling in the coupled arrival and departure banks.

However, banks create many flight interdependencies. Departure bank flights mostly consist of connecting flights from an associated arrival bank. As a result, airline hubbing systems are sensitive to air traffic delays. A delay of a single flight in an arrival bank, may cause an airline to delay the departures of several flights in the departure bank to accommodate incoming passengers connecting to multiple outgoing flights.

* Shuttle Services

A shuttle service is established by an airline to provide air travel between an airport pair on a regular basis throughout the day. Typically, shuttle flights depart on an hourly basis. Most of the aircraft within a shuttle service are dedicated to travel back and forth between the airport pair. A shuttle flight is a st Case 2 - Results: Airport 4ght, and therefore, also potentially affected by FM actions and congestion delays.

3.1.4.4. Further Schedule Observations

Three observed characteristics of airline schedules which are not in the POAGG model, per se, are aircraft specific flight and layover times, and airline departure time biases.

Actual flight times between two given airports differ for jet engine and propeller-driven aircraft. Accordingly, airlines schedule a flight's arrival time to reflect aircraft type. In POAGG,
the model assumes a uniform fleet on each route, i.e., scheduled flight times are assumed to be the same for all flights traveling between two given airports.

This assumption does not affect the validity of the model, because in practice, most routes are scheduled with similar aircraft types. The uniform fleet assumption is largely correct except for some short-range routes, in which propeller-driven aircraft and jet engine aircraft may be scheduled between the same airport pair. POAGG does model the difference in a flight's direction for a given airport pair, e.g., the flight time from Airport 1 to Airport 2 is modeled separately from the flight time from Airport 2 to Airport 1. This reflects the difference in East-West and West-East flight times between the same airports due to jet stream winds.

In practice, airlines schedule layover times according to aircraft type. The type and respective size of an aircraft influences the time required to prepare for connecting flights and transfer passengers, i.e., the turnaround time. Furthermore, certain airlines have faster turnaround times than others for the same aircraft type. Although POAGG models layover times as random variables, the variation does not reflect aircraft types or airlines. Correct layover times would enhance the realism of generated schedules; however for FM analysis, precise layover times are not critical.

Departure time biases refers to the practice of favoring departures on the hour and half-hour. To a lesser degree, airline schedules contain departures on the five minute interval, e.g., 8:05, 8:10, 8:15 etc., over other minutes within the hour. Bias toward departures on the hour and half-hour is indirectly simulated in the shuttle service model.

3.2. Problem and Approach

The approach to solving the schedule-generation problem is to formulate the modeling objectives into constraints of a mathematical program. The solution approach is to define an objective function, by relaxing a constraint, and implement a heuristic solution technique. The mathematical formulation and constraints will be referred to as the baseline formulation and baseline constraints.

The baseline formulation is presented in the next section in order to introduce the variables and baseline constraints of the problem. The baseline formulation does not fully describe the schedule-generation problem, e.g., shuttle service and hub modeling are omitted. The full schedule-generation problem, including shuttle service and hub modeling, will be provided in the description of the heuristic solution technique. In this manner, the emphasis is on the heuristic approach.
3.2.1. Baseline Formulation

The formulation is designed to generate a fictitious flight schedule for a set of N airports. The problem requires that each specified airport, i.e., one of the N airports, is described by a frequency distribution of arrival flights over a day. The number of intervals, or periods, in the arrival distributions is the variable T. If the frequency distributions are based intervals of one hour, T=24, for half-hour intervals, T=48, etc. Further inputs for the formulation include flight connectivity parameters and flight times for each airport pair. Flight connectivity refers to the number of connecting flights between an airport pair.

An individual flight is represented by \( f \). Flight \( f' \) is a connecting flight of \( f \) if both flights are scheduled to use the same aircraft and \( f' \) immediately follows \( f \). In a connecting flight pair, e.g., \( f' \) and \( f \), the layover time between the flights is assigned to \( f \).

Given a fixed number of flights \( F \), the specific objective of the problem is to calculate arrival and departure times for each flight that satisfy the given arrival distributions. Simultaneously, the problem solution technique must assign connecting flights between specified airports that satisfy the flight connectivity parameters. The model assumes that flights landing after period T, i.e., the next day, are not constrained by arrival distributions. The baseline formulation follows.

Given

\[
\begin{align*}
F & \quad \text{total number of flights} \\
N & \quad \text{total number of airports} \\
T & \quad \text{total time periods} \\
a_{i,t} & \quad \text{number of scheduled arrival flights at airport } i \text{ in period } t \\
d_{i,j} & \quad \text{number of connecting flights scheduled from airport } i \text{ to airport } j, \text{ during the } T \text{ time periods} \\
g_{i,j} & \quad \text{flight time from airport } i \text{ to airport } j \\
l_f & \quad \text{scheduled layover time for flight } f \text{ upon arrival, if } f \text{ is assigned a connecting flight } f' \\
\end{align*}
\]

solve for

\[
\begin{align*}
s_f & \quad \text{scheduled departure time for flight } f \\
o_f & \quad \text{scheduled arrival time for flight } f \\
x_{i,j}^f & \quad 0-1 \text{ variable describing connection status of flight } f \\
y_{i,t}^f & \quad 0-1 \text{ variable describing arrival period and arrival airport of flight } f \\
\end{align*}
\]
where \( x_{i,j}^f \) and \( y_{i,t}^f \) are defined as

\[
\begin{align*}
  x_{i,j}^f &= \begin{cases} 
  1 & \text{if flight } f \text{ is a connecting flight from airport } i \text{ to airport } j \\
  0 & \text{otherwise} 
  
end{cases} \\
y_{i,t}^f &= \begin{cases} 
  1 & \text{if flight } f \text{ is scheduled to arrive at airport } i \text{ in period } t \\
  0 & \text{otherwise} 
  
end{cases}
\]

The following baseline constraints must hold:

\[
\begin{align*}
  \sum_{f=1}^{F} y_{i,t}^f &= a_{i,t} \quad \forall i, t & (1) \\
  \sum_{f=1}^{F} x_{i,j}^f &= d_{i,j} \quad \forall i, j & (2) \\
  o_f - s_f &= g_{i_f, i_f} \quad \forall f & (3) \\
  s_{f'} - o_f &= l_f \quad \forall f' & (4)
\end{align*}
\]

In order to understand variable \( x_{i,j}^f \), it is necessary to note the difference between a connecting flight versus a flight that originates from an airport, i.e., an originating flight. A connecting flight is scheduled to use the same aircraft of a previous flight that lands within the same day. Conversely, an originating flight is the first departure of an aircraft in the day. Although an originating flight may travel from airport \( i \) to airport \( j \), it does not constitute a connecting flight with regard to the variable \( x_{i,j}^f \).

Concerning the baseline constraints, constraint (1) ensures that a flight schedule solution satisfies the airport arrival distributions. Constraint (2) ensures that a solution contains the desired flight connectivity between each airport pair. Constraint (3) specifies that the difference between each flight’s arrival and departure time must be the scheduled flight time between the arrival and departure airports. Finally, constraint (4) ensures that connecting flights in a solution are scheduled to depart after their preceding flights are scheduled to land and receive the assigned layover time. The number of decision variables in the baseline formulation is

\[
F + F + F \cdot N \cdot (N - 1) + F \cdot N \cdot T
\]

The four terms in equation (5) corresponds to the variables \( s_f, o_f, x_{i,j}^f, \) and \( y_{i,t}^f \) respectively.

An objective function can be formulated by relaxing one of the baseline constraints. If constraint (1) is relaxed the problem becomes:
Chapter 3: Airport Demand Schedule Modeling

\[
\min \sum_{i=1}^{N} \sum_{t=1}^{T} \left( a_{i,t} - \sum_{f=1}^{F} y_{i,t}^{f} \right) \quad (6)
\]

s.t. \[
\sum_{f=1}^{F} x_{i,j}^{f} = d_{i,j} \quad \forall i, j \quad (2)
\]

\[
o_f - s_f = g_{i,j}^{f} \quad \forall f \quad (3)
\]

\[
s_f - o_f = l_f \quad \forall f' \quad (4)
\]

Equation (6) minimizes the difference between the desired (or input) and realized arrival distributions.

If constraint (2) is relaxed, the problem becomes:

\[
\min \sum_{i=1}^{N} \sum_{j=1}^{N} \left( d_{i,j} - \sum_{f=1}^{F} x_{i,j}^{f} \right) \quad (7)
\]

s.t. \[
\sum_{f=1}^{F} y_{i,t}^{f} = a_{i,t} \quad \forall i, t \quad (1)
\]

\[
o_f - s_f = g_{i,j}^{f} \quad \forall f \quad (3)
\]

\[
s_f - o_f = l_f \quad \forall f' \quad (4)
\]

Equation (7) minimizes the difference between the number of desired and realized connecting flights.

A third option is to allow flight layover time, i.e., \( l_f \), to vary as a decision variable. In practice, competitive airlines attempt to minimize the amount of time operational aircraft spend on the ground during the day. The ground time is incorporated in airline schedules as layover times. However, there is a minimum ground time required by the airlines in order to deplane arriving passengers, board departing passengers, and prepare aircraft. An additional input variable is defined

\[\alpha \quad \text{- minimum required layover time}\]

The model assumes \( \alpha \) is equivalent for all flights. The problem may be formulated to minimize total layover times as follows:
Chapter 3: Airport Demand Schedule Modeling

\[
\min \sum_{f=1}^{F} l_f \\
\text{s.t. } \sum_{f=1}^{F} y_{i,t}^f = a_{i,t} \quad \forall i,t
\]

(1)

\[
\sum_{f=1}^{F} x_{i,j}^f = d_{i,j} \quad \forall i,j
\]

(2)

\[
o_f - s_f = g_{f,i}, \quad \forall f
\]

(3)

\[
s_f - o_f - l_f = 0 \quad \forall f'
\]

(4)

\[
l_f \geq \alpha \quad \forall f
\]

(9)

The heuristic solution technique is a variation of the third formulation that minimizes layover times.

3.2.2. Heuristic Solution Technique

The objective function of the heuristic approach of generating hypothetical airport demand schedules is to minimize the total number of flights with layovers that exceed a specified threshold. Given the objective function, the solution technique is a two stage process. The first stage involves assigning initial values to intermediate variables that will govern the generation of flight itineraries, which is the second stage. During the second stage, the initial values of the intermediate variables determine the assignment of arrival and departure times and flight connections.

3.2.2.1. Objective Function

The objective function for the heuristic solution technique is a variation of the objective function that minimizes total layover times. The technique requires an additional variable defined as

\[
\lambda \quad - \text{maximum allowable layover time}
\]

Constraint (9) is modified to

\[
\alpha \leq l_f \leq \lambda \quad \forall f
\]

(10)
Constraint (10) is subsequently relaxed, allowing layovers to exceed \( \lambda \) in order to form an objective function using the 0-1 variable \( z_j \) defined as

\[
    z_j = \begin{cases} 
        1 & \text{if } l_j > \lambda \\ 
        0 & \text{otherwise} 
    \end{cases}
\]

The objective function is

\[
    \min \sum_{j=1}^{F} z_j
\]

which attempts to minimize the total number of flights with layover times that exceed \( \lambda \). The constraints governing the heuristic solution are baseline constraints (1) through (4).

3.2.2.2. Stage One: Assigning Values to Intermediate Variables

The first stage of the solution technique involves assigning initial values to intermediate variables. The intermediate variables are

- \( a_{i,t} \) - the number of scheduled arrival flights at airport \( i \) in period \( t \)
- \( d_{i,j} \) - the total number of connecting flights from airport \( i \) to airport \( j \)
- \( c_{i,j,t} \) - number of flights arriving at airport \( i \) in period \( t \), with connecting flights to airport \( j \)

Although \( a_{i,t} \) and \( d_{i,j} \) are defined as givens in the baseline formulation, they are intermediate variables, derived from the following input variables,

- \( e_i \) - the total number of arrivals at airport \( i \)
- \( k_{i,t} \) - percentage of total arrivals at airport \( i \), that arrive in period \( t \)
- \( p_{i,j} \) - percentage of total arrivals at airport \( i \) with connecting flights to airport \( j \)

\( k_{i,t} \) is the relative frequency distribution of arrival flights for airport \( i \). \( p_{i,j} \) is referred to as the connection percentage between airports \( i \) and \( j \).

Calculation of \( a_{i,t} \)

The calculation the flight arrival distributions for each specified airport, i.e., \( a_{i,t} \) must ensure that
\[
\sum_{t=1}^{T} a_{i,t} = e_i \quad \forall i
\]

(12)

The formula for calculating \( a_{i,t} \) is

\[
a_{i,t} = k_{i,t} \cdot e_i
\]

(13)

Because the formula does not necessarily produce integer values for \( a_{i,t} \), it is necessary to implement a numerical rounding process to correct for rounding errors. In the equations that describe the rounding process below, \( r_{i,t} \) is the cumulative total rounding error from all periods preceding \( t \) for airport \( i \). The rounding process is:

\[
a_{i,t}=\begin{cases} 
\text{round}(k_{i,t} \cdot e_i) & \text{if } -1.0 < r_{i,t} < 1.0 \\
\lfloor k_{i,t} \cdot e_i \rfloor & \text{if } r_{i,t} < -1.0 \\
\lceil k_{i,t} \cdot e_i \rceil & \text{if } r_{i,t} > 1.0
\end{cases}
\]

(14)

\[
r_{i,t} = r_{i,t-1} + a_{i,t-1} - (k_{i,t-1} \cdot e_i)
\]

(15)

where \( r_{i,1} = 0.0 \) for all \( i \). The function \( \text{round}(\cdot) \), rounds the argument to the closest integer value. \( \lfloor \cdot \rfloor \) rounds the argument to the nearest lower integer value, while \( \lceil \cdot \rceil \) rounds the argument to the nearest higher integer value.

The mechanics of the rounding process work in such a manner that excessive cumulative rounding error is detected and corrected in the subsequent calculation of \( a_{i,t} \) for a given airport. Table 3.1 demonstrates the functioning of the rounding process for a five-period example at airport \( \hat{i} \). In the example, \( e_i = 110 \) arrival flights and \( k_{i,t} \) is described by the graph in Figure 3.2.

![Figure 3.2: Relative Frequency Distribution of Arrivals at Airport \( \hat{i} \)](image-url)

73
Table 3.1: Rounding Process for \( a_{i,t} \)

<table>
<thead>
<tr>
<th>( t )</th>
<th>( k_{i,t} \cdot e_i )</th>
<th>( r_{i,t} )</th>
<th>( a_{i,t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1</td>
<td>0.0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3.3</td>
<td>-0.1</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4.4</td>
<td>-0.4</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>2.2</td>
<td>-0.8</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>4.4</td>
<td>-1.0</td>
<td>5</td>
</tr>
</tbody>
</table>

In period 5, the process rounds \( k_{i,t} \cdot e_i \) to the nearest higher integer, in order to adjust for successive rounding to lower integers in previous periods. The process ensures that rounding \( k_{i,t} \cdot e_i \) for the final period at each airport, i.e., \( T \), does not violate (12).

3.2.2.2.2. Calculation of \( d_{i,j} \)

The total number of connecting flights between any two specified airports is calculated using the equation

\[
d_{i,j} = \left\lfloor p_{i,j} \cdot e_i \right\rfloor \quad \forall i,j
\]

(16)

By rounding to the nearest lower integer for all values of \( d_{i,j} \), the equation ensures that there is no attempt to assign more connecting flights from an airport than total arrival flights into the airport, i.e.

\[
\sum_{j=1}^{N} d_{i,j} \leq e_i \quad \forall i
\]

(17)

3.2.2.2.3. Calculation of \( c_{i,j,t} \)

The values of \( c_{i,j,t} \) describe the distributions, by period, of connecting flights between all airport pairs. The relationship between \( a_{i,t} \) and \( c_{i,j,t} \) for three specified airports is demonstrated in Figure 3.3.
Figure 3.3: Relationship Between $a_{i,t}$ and $c_{i,j,t}$

The arrows in Figure 3.3 represent groups of flights. Of the $a_{i,t}$ arrivals at Airport 1 in period $i$, $c_{1,2,1}$ and $c_{1,3,1}$ are designated as having connecting flights bound for Airports 2 and 3, respectively. The layover times are not necessarily equivalent for all flights.

The initial values of $c_{i,j,t}$ are assigned using a numerical process similar to the rounding process used for calculating $a_{i,t}$. The requirements for the connection assignment process are

$$\sum_{t=1}^{T} c_{i,j,t} = d_{i,j} \quad \forall i, j$$

(18)

and

$$\sum_{j=1}^{N} c_{i,j,t} \leq a_{i,t} \quad \forall i, t$$

(19)

The equality in (18) ensures that the desired numbers of total connecting flights are distributed to the individual values of $c_{i,j,t}$. The inequality in (19) ensures that the process is restricted from assigning more connecting flights than incoming flights for each period and airport.

An additional requirement is that the values of $c_{i,j,t}$ are spread out over the time periods proportional to the arrival frequency distributions. This ensures the connection process is not biased in assigning connections for different intervals of the day.

The basic connection assignment formula for calculating $c_{i,j,t}$ is

$$c_{i,j,t} = p_{i,j} \cdot a_{i,t}$$

(20)

The connection formula does not necessarily produce integer values for $c_{i,j,t}$. In order to correct for rounding errors, the basic formula is augmented with the following numerical process:
\[ c_{i,j,t} = \begin{cases} \text{round}(p_{i,j} \cdot a_{i,t}) & \text{if } -1.0 < r_{i,j,t} < 1.0 \\ p_{i,j} \cdot a_{i,t} & \text{if } r_{i,j,t} < -1.0 \\ p_{i,j} \cdot a_{i,t} & \text{if } r_{i,j,t} > 1.0 \end{cases} \]  

(21)

where

\[ r_{i,j,t} = r_{i,j,t-1} + c_{i,j,t-1} - (p_{i,j} \cdot a_{i,t-1}) \]  

(22)

and

\[ r_{i,j,1} = 0.0 \quad \forall i,j \]  

(23)

In equation (21, 22, and 23), \( r_{i,j,t} \) is the cumulative total rounding error for periods preceding period \( t \), for each airport pair \( i,j \). The mechanics of the connection assignment process are identical to the mechanics for the rounding process used in the calculation of \( a_{i,t} \).

The connection assignment process is applied to each specified airport pair. There is potential for the inequality in (19) to be violated, if the process selects the function \( [\cdot] \) to calculate the values of \( c_{i,j,t} \) during the same period for multiple airport pairs. This can be illustrated through an example.

The parameters for the example are

\[ N = 5, \]

\[ p_{i,j} = 25.0\% \text{ for } j = 2, 3, 4, \text{ and } 5, \]

\[ t = \tilde{t}, \]

\[ a_{i,t} = 2, \text{ and} \]

\[ r_{1,j,t} < -1.0 \text{ for } j = 2, 3, 4, \text{ and } 5. \]

In this example, the process will simultaneously select the function \( [\cdot] \) to calculate the values of \( c_{1,i,t} \). As a result, \( c_{1,j,i} = 1 \text{ for } j = 2, 3, 4, \text{ and } 5 \), which violates (19).

If this situation occurs, the process is designed to lower selected values of \( c_{1,j,i} \) on a priority basis until (19) is no longer violated. In the example, it is necessary to reduce the collective values of \( c_{1,i,t} \) by 2. The process selects airport pairs that are closest to achieving their total connections for the day. If \( d_{t,i} \) is closest to being completely distributed for Airports 2 and 3, the values of \( c_{1,2,i} \) and \( c_{1,3,i} \) are adjusted to 0, and thus, (19) is no longer violated.

This correction, however, may result in assigning insufficient total connecting flights for some airport pairs, i.e.,
\[
\sum_{t=1}^{T} c_{i,j,t} < d_{i,j}
\]

If, after completing the connection assignment process for all airports, the above inequality occurs for one or more airport pairs, the process will seek to increment the values of \(c_{i,j,t}\), starting at \(t=1\), until all connecting flights are assigned between all airport pairs. This is always achievable, because the formula for calculating \(d_{i,j}\) ensures there is no attempt to assign more connecting flights than arrival flights at each airport.

The following example illustrates the assignment of values to \(c_{i,j,t}\) over a day between two airports (Airports 1 and 2). The example demonstrates the effectiveness of the connection assignment process in distributing the values of \(c_{i,j,t}\) proportional to the arrival distribution of Airport 1. The example also demonstrates that it is insufficient to use the functions \(\text{round}()\), \([\cdot]\), and \(\lceil\cdot\rceil\) by themselves, as alternative methods of calculating \(c_{i,j,t}\). The input parameters are:

\[
\begin{align*}
e_1 &= 400 \text{ arrival flights} \\
p_{1,2} &= 5.0\% \\
T &= 48 \text{ (30 minute periods)} \\
k_{1,2} &= \begin{cases} 
0.0\% & 1 \leq t \leq 12 \\
2.0\% & 13 \leq t \leq 25 \text{ and } 39 \leq t \leq 48 \\
4.0\% & 26 \leq t \leq 32 \text{ and } 35 \leq t \leq 38 \\
5.0\% & t = 33 \text{ and } t = 34 
\end{cases}
\end{align*}
\]

From the input parameters, \(d_{1,2} = (0.05 \cdot 400) = 20\) connecting flights. The values of \(a_{i,t}\) for the example are depicted graphically in Figure 3.4.

\[\text{Figure 3.4: Arrival Distribution for Connection Assignment Example}\]

The arrival distribution in Figure 3.4 is an example of arrivals at a major airport that experiences low rates of arrival in the morning and late evening, a high rate of arrivals in the afternoon with a peak between the hours of 4:00 and 5:00 PM.
Table 3.2 presents a comparison of the results from the connection assignment process in the “POAGG” column, to the functions \( \text{round}(\cdot) \), \( \lfloor \cdot \rfloor \), and \( \lceil \cdot \rceil \), for the given example.

**Table 3.2: Assignment of Initial Values for \( c_{1,2,t} \)**

<table>
<thead>
<tr>
<th>Period</th>
<th>( a_{1,t} )</th>
<th>( p_{1,2} \cdot a_{1,t} )</th>
<th>( \text{round}(\cdot) )</th>
<th>( \lfloor p_{1,2} \cdot a_{1,t} \rfloor )</th>
<th>( \lceil p_{1,2} \cdot a_{1,t} \rceil )</th>
<th>POAGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 12</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>1.0</td>
<td>2.0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>34</td>
<td>1.0</td>
<td>2.0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>35</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>36</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>37</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>38</td>
<td>0.8</td>
<td>1.6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>39</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>41</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>42</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>43</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>46</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>47</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>48</td>
<td>0.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>400</td>
<td>100.0</td>
<td>13</td>
<td>20</td>
<td>2</td>
<td>20</td>
</tr>
</tbody>
</table>

As seen in Table 3.2, the numerical connection process spreads the values for \( c_{1,2,t} \) over the 48 periods. The other functions assign insufficient total connections, or values for \( c_{1,2,t} \) that are biased towards either earlier or later periods.
3.2.2.3. Stage Two: Generating Flight Itineraries

The second stage of the heuristic solution technique is the generation of individual flight itineraries. Flight itineraries are incrementally generated, using a sequential, period-by-period algorithm.

All flight itineraries will reflect either an originating or a connecting flight. Originating flights are the first scheduled flights in the day for their respective aircraft, i.e., originating flights have no preceding flights. All other flights are connecting flights. Connecting flights may be preceded by originating flights or other connecting flights. Figure 3.5 graphically depicts an originating flight $f_O$, and a connecting flight $f_C$, scheduled from Airport 1 to Airport 2.

![Figure 3.5: Originating and Connecting Flights](image)

The vertical line segments in each aircraft's trajectory represent layover times. Both flights in Figure 3.5 are scheduled with follow-on connecting flights. Although not depicted in Figure 3.5, if a flight is the last scheduled flight for an aircraft in a day, the flight is referred to as a terminating flight.

The objective of the second phase is to generate individual flight itineraries that collectively constitute a flight schedule that satisfies all input parameters. The data fields for each flight itinerary are:

- arrival airport
- arrival time
- departure airport
- departure time
- flight number
- carrier name
- aircraft type
- "next flight"
- "previous flight"

The "next flight" and "previous flight" data fields contain the flight numbers of connecting flights and preceding flights, respectively.

This section will present the algorithm for generating individual flight itineraries and the process of completing all flight data fields. First, we introduce the concept of Airport X, which models all airports outside the network of specified airports.

3.2.2.3.1. Airport X

In the model, all airports which are not specified in the set of N airports are grouped together and modeled as Airport X. A representation of an airport network with three specified airports and Airport X is presented in Figure 3.6. The arrows in Figure 3.6 represent flights.

![Airport Network Diagram](image)

Figure 3.6: Airport Network

Airport X is the scheduled departure airport for originating flights which are not scheduled from specified airports. Airport X is also the scheduled arrival airport for connecting flights which are not bound for specified airports. Flight times to and from Airport X are modeled as discrete random variables requiring probability mass functions (PMFs) as input values. Flights connecting to Airport X or arriving from Airport X are assigned flight times using
a Monte Carlo method. The probability mass functions (PMFs) for the Airport X flight times are input parameters.

The schedule-generation algorithm does not attempt to generate past or future connecting flight information for flights scheduled to and from Airport X. Therefore, all flights scheduled from Airport X appear to be originating flights, and all flights scheduled to Airport X appear to be terminating flights. This does not degrade the utility of the hypothetical schedules for FM purposes, since generated schedules preserve connection data for the specified airports of interest.

3.2.2.3.2. Assigning an Arrival Airport

Originating flights are generated in order to satisfy the arrival distributions. Therefore, assigning an arrival airport to an originating flight is determined by the arrival distribution being satisfied by the algorithm. The algorithm assigns an arrival airport to an originating flight before assigning it a departure airport. Assigning an arrival airport to a connecting flight is determined by the connection parameter being satisfied by the algorithm.

Concerning connecting flights, Airport X is the arrival airport for flights connecting outside of the network of specified airports. The purpose of scheduling flights into Airport X is to enhance the realism of the generated schedules. Without Airport X, all aircraft would appear to terminate at the specified airports. In order to terminate some aircraft at the specified airports, there is an input parameter that defines the cutoff time after which flights are no longer connected to Airport X.

3.2.2.3.3. Assigning a Departure Airport

After receiving an arrival airport, an originating flight is assigned a departure airport using a Monte Carlo method. The method is illustrated in the following example. Assume the parameters for total scheduled arrivals for each specified airport in Figure 3.6 above are

\[ e_1 = e_2 = e_3 = 100 \text{ arrival flights} \]

and the numbers of scheduled connecting flights into Airport 1, from Airports 2 and 3 are

\[ d_{2,1} = 10 \text{ connecting flights} \]
\[ d_{3,1} = 20 \text{ connecting flights} \]
In this example, it is necessary to schedule an additional 70 arrival flights into Airport 1. The remaining 70 flights are originating flights that arrive from either the other specified airports (Airports 2 and 3) or Airport X.

The departure airport is modeled as a discrete random variable. The PMF for departure airports is defined by converting the connection percentages into probabilities. Each PMF is completed with a probability that the departure airport will be Airport X.

In the example above, the connection percentages into Airport 1 are

\[ p_{2,1} = 10.0\% \]
\[ p_{3,1} = 20.0\% \]

The resulting PMF for the departure airport is

\[ P(\text{departure airport} = \text{Airport 2}) = 0.1 \]
\[ P(\text{departure airport} = \text{Airport 3}) = 0.2 \]
\[ P(\text{departure airport} = \text{Airport X}) = 0.7 \]

In order assign an originating flight a departure airport, a random number \( w \) is generated from \( W, \) a uniformly distributed random variable over the interval \((0, 1)\), i.e., \( U(0,1) \). A departure airport will be assigned using the criteria

\[ \text{departure airport} = \begin{cases} 
\text{Airport 2} & \text{if } 0.0 \leq w < 0.1 \\
\text{Airport 3} & \text{if } 0.1 \leq w < 0.3 \\
\text{Airport X} & \text{otherwise}
\end{cases} \]

The PMF may be defined entirely by the connecting percentages into Airport 1, such that

\[ P(\text{departure airport} = \text{Airport X}) = 0 \]

In which case, no flights will be scheduled into Airport 1 from Airport X.

The departure airport for a connecting flight is simply the arrival airport of its preceding flight. Due to the sequential nature of the algorithm, connecting flight itineraries will always be generated after their preceding flight itineraries.

3.2.2.3.4. Assigning Airline Information and Flight Numbers

Each flight itinerary requires airline information, in the form of a carrier name and aircraft type. Originating flights are assigned airline information using a Monte Carlo process.
The process references input PMFs that define carrier names and aircraft types by arrival airport. An example PMF describing three different carriers operating at a specified airport may be

\[
P(\text{carrier} = \text{Carrier}_1) = .3 \\
P(\text{carrier} = \text{Carrier}_2) = .3 \\
P(\text{carrier} = \text{Carrier}_3) = .4
\]

The assignment of a flight's aircraft type is a similar process. Every carrier at a specified airport is described by an input PMF that defines the airline's fleet. After assigning a flight a carrier name, the Monte Carlo process is applied to determine a flight's aircraft type. A connecting flight is assigned airline information that matches the airline information of its preceding flight.

By using a Monte Carlo method to assign airline information, flight itineraries created by the model resemble flights within an OAG schedule. However, this not intended to model the scheduling practices of any specific airline.

Flights are assigned flight numbers by means of a counting variable that tracks the total number of generated flight itineraries. The flight for the tenth generated flight itinerary will be assigned flight number 10.

3.2.2.3.5. Process of Flight Itinerary Generation

The algorithm for generating flight itineraries is similar to a time-advance algorithm for a discrete-time simulation. Time (t) in periods, is initialized to the first period of the day, i.e., \( t = 1 \), and advances period-by-period. In each period, the algorithm generates itineraries for originating and connecting flights for each specified airport.

A connecting flight between two specified airports will satisfy a connection requirement. At the same time, if the connecting flight is scheduled to land on or before period T, the connecting flight will satisfy an arrival requirement. An additional intermediate variable is introduced in the second stage to monitor the scheduled arrivals of connecting flights into specified airports. The variable is

\[
b_{i,t} \quad \text{- number of connecting flights arriving at airport } i \text{ in period } t
\]

While generating flight itineraries, the algorithm ensures that the arrival distributions are not exceeded by the scheduled arrivals of connecting flights. In variable terms, the algorithm ensures that

\[
b_{i,t} \leq a_{i,t} \quad \forall i, t
\]  \hspace{1cm} (24)
Figure 3.7 illustrates the relationship between arrival flights and connecting flights.

![Diagram showing the relationship between arrival flights and connecting flights](image)

**Figure 3.7: Connecting Flights versus Arrival Flights**

In Figure 3.7, flights \( f_1 \) and \( f_2 \) both arrive in period \( t_1 \) at Airport 1. Flights \( f_1' \) and \( f_2' \) are two connecting flights from Airport 1 to Airport 2. Flights \( f_1' \) and \( f_2' \) are also two arrivals at Airport 2. Due to the different arrival times and layover times for \( f_1 \) and \( f_2 \), \( f_1' \) and \( f_2' \) are scheduled to arrive in different periods at Airport 2. This demonstrates that the arrival period of a connecting flight is unknown until its preceding flight is assigned an arrival time and a layover time. For this reason, \( b_{i,t} \) is needed in the second phase to track the arrivals of connecting flights.

* Generating Originating Flight Itineraries *

Assume the algorithm has advanced to period \( \hat{t} \), and airport \( \hat{i} \). The algorithm first generates, if necessary, itineraries for originating flights into \( \hat{i} \) from Airport X or the other specified airports, in order to satisfy the arrival distribution for \( \hat{i} \) during \( \hat{t} \). The required number of originating flight itineraries to be generated is \( a_{i,t} - b_{i,t} \).

The arrival time, \( o_f \), is the first calculation in generating an originating flight itinerary for \( f \). Originating flights are assigned arrival times using a Monte Carlo method. \( o_f \) is assigned a random number generated from a uniform distribution over the interval of the period in the arrival distribution being satisfied, i.e., \( \hat{t} \) from above. \( o_f \) is originally calculated in absolute minutes, where 7:40 AM is equivalent to \((7 \cdot 60) + 40 = 460\) absolute minutes. \( o_f \) is subsequently converted into an arrival hour and minute using the equations.
Chapter 3: Airport Demand Schedule Modeling

\[
\text{arrival hour}_f = \left\lfloor \frac{o_f}{60} \right\rfloor
\]  
(25)

\[
\text{arrival minute}_f = o_f - 60 \cdot \text{arrival hour}_f
\]  
(26)

After the departure airport, i.e., \(i_f\), for an originating flight \(f\) is determined using the process described in Section 3.2.2.3.3, \(f\) is assigned a departure time, i.e., \(s_f\), using baseline constraint (3), where \(g_{i_f,i_f}\) is the input flight time.

\[
s_f = o_f - g_{i_f,i_f} \quad \forall f
\]  
(3)

If the flight is assigned Airport X as the departure airport, the flight time is assigned randomly according to a PMF for flight times to and from Airport X. The flight times between specified airports, i.e., \(g_{i,j}\), are input parameters for all specified airport pairs.

By generating the scheduled arrival times from a uniform distribution, the interarrival and interdeparture times among flights within each period are exponentially distributed. As a result, the scheduled departure times may exhibit a “bunching” effect within each period. This effect models an observation of actual OAG flight schedules.

All originating flights have no previous flights. The “previous flight” data field is set to N/A for not applicable. The “next flight” data field is also set to N/A. If a flight is assigned a connecting flight at a later time, the “next flight” data field will be appropriately updated.

- **Generating Connecting Flight Itineraries**

After the algorithm satisfies the arrival distribution for period \(\hat{i}\) at airport \(\hat{i}\), the algorithm attempts to generate flight itineraries for connecting flights from \(\hat{i}\) based on the arrivals in \(\hat{i}\). The algorithm randomly selects \(c_{i,j,\hat{i}}\) of the \(a_{i,i}\) arrivals to have connecting flights for each specified airport \(j\).

The first step in generating a flight itinerary for connecting flight \(f'\) is to generate a layover time for the preceding arrival flight, \(f\). The initial layover time, \(l_f\), is generated from the continuous random variable, described by the probability density function (PDF) in Figure 3.8.
\( \alpha \) is the minimum required layover time. \( m \) is an input parameter with the restriction that \( m \leq \lambda \), where \( \lambda \) is the maximum allowable layover time.

The difference between the parameters \( \lambda \) and \( m \) is that \( m \) defines the shape of the PDF for the initial calculation for a flight’s layover time, while \( \lambda \) is a threshold value for maximum layover time. The initial value for \( l_f \) may be adjusted by the algorithm, as will be explained below, in which case, \( \lambda \), becomes a factor. If no values for \( l_f \) are adjusted, the PDF guarantees the constraint \( \alpha \leq l_f \leq \lambda \) is satisfied. The values for \( \lambda \) and \( m \) are separate input parameters in order to model schedules with initial layover time distributions that are independent of the maximum allowable layover time.

After \( l_f \) is generated, the departure time of \( f' \) is determined by baseline constraint (4).

\[
s_{f'} = l_f + o_f \quad \forall f'
\]

(4)

and the arrival time of flight \( f' \) is calculated using baseline constraint (3).

At this point, the algorithm verifies that the arrival distribution for the arrival airport of \( f' \) will not be violated by the scheduled arrival time of \( f' \). The arrival period of \( f' \), represented by \( t' \), is calculated using the equation

\[
t' = \left[ \frac{o_{f'}}{\theta} \right]
\]

(27)

If \( \hat{j} \) is the arrival airport of \( f' \), the algorithm examines the value of \( b_{j,t'} \). If \( b_{\hat{j},t'} < a_{\hat{j},t'} \), the arrival distribution for airport \( \hat{j} \) has not been satisfied by previously generated connecting flight itineraries. In this situation, the algorithm proceeds to generate an itinerary for flight \( f' \) with the initial value of \( o_{f'} \) calculated from baseline constraint (3). The value of \( b_{\hat{j},t'} \) is incremented by one.

However, if \( b_{\hat{j},t'} = a_{\hat{j},t'} \), this implies that the number of scheduled arrivals into airport \( \hat{j} \) in period \( t' \) has been satisfied by previously generated itineraries for connecting flights.
Scheduling flight \( f' \) with the initial value of \( o_f \) will result in excess arrivals in period \( t' \) at airport \( j \).

The algorithm invokes a connection sub-algorithm in order to resolve the conflict. The overall objective of the sub-algorithm is to minimize the number of flights with layover times that exceed \( \lambda \). The sub-algorithm contains two alternative correction techniques.

The first technique of the sub-algorithm is to search for future periods at \( j \) that have not been satisfied by connecting flights. The value of \( l_f \) is incrementally increased by the length of one period, \( \theta \), as demonstrated in Figure 3.9.

![Figure 3.9: First Correction Technique](image)

If \( t' = t4 \) is satisfied at Airport 2, \( t' \) is incremented to \( t5 \) and the arrival distribution at Airport 2 is examined for unsatisfied arrivals in \( t5 \).

If \( b_{j,t'} = a_{j,t''} \), the technique continues the search, again incrementing \( l_f \) by \( \theta \). The technique stops if a period is discovered for which \( b_{j,t'} < a_{j,t''} \), or if the layover time exceeds the maximum threshold, i.e., \( l_f > \lambda \). If the technique discovers a period with unsatisfied arrivals prior to exceeding the maximum layover time, the connecting flight itinerary is generated with arrival and departure times based on the adjusted value of \( l_f \).

If \( l_f > \lambda \), the sub-algorithm switches to a second correction technique, which attempts to assign a connecting flight to an earlier arrival flight at airport \( i \). The technique searches for two flights itineraries:
Flight 1 ($f_1$) - an originating flight, arriving at airport $\hat{i}$, with either no connecting flight or a connecting flight to Airport $X$, and

Flight 2 ($f_2$) - an originating flight, arriving at airport $\hat{j}$, with either no connecting flight or a connecting flight to Airport $X$.

Flights $f_1$ and $f_2$ will replace flights $f$ and $f'$ as a connection pair. The second technique requires that the following inequality be true:

$$a_{f_2} \geq a_{f_1} + \alpha + g_{i,j}$$

(28) ensures that a connection can be accomplished using $f_1$ and $f_2$ without adjusting either flight's arrival times, and hence, the arrival distributions for airports $\hat{i}$ and $\hat{j}$ will be unaffected. If Flights 1 and 2 are discovered, the departure airport of $f_2$ is changed to $\hat{i}$, and the departure time is modified to reflect the flight time from $\hat{i}$ to $\hat{j}$. If multiple flight pairs meet the necessary criteria, the technique selects the pair with the minimum layover time. The original flight $f$ is assigned a connecting flight to Airport $X$.

The second technique increases the value of the intermediate variable $c_{i,j,j',1}$ by one, reflecting the arrival period of Flight 1. The value of $c_{i,j,j',1}$ is decreased by one to reflect the unsuccessful attempt of the first technique to assign a connecting flight for $f$ within the specified airport network.

If there are no flight pairs that meet the necessary criteria in the second technique, the sub-algorithm returns to the first correction technique. The search for available arrivals continues, however, with layover times that exceed $\lambda$. If all arrivals are accounted for in future periods, i.e.,

$$b_{j,t} = a_{j,t} \quad \forall t, \quad t' < t \leq T$$

the sub-algorithm schedules the connecting flight, $f'$, with an arrival time in the next day. The model assumes that flights scheduled to land after midnight, are not constrained by arrival distributions.

Completing the remaining data fields for a connecting flight is accomplished as described in previous sections. The data field "previous flight" is completed to reflect the flight number of the preceding arrival and, the data field "next flight" of flight $f$ is completed to reflect the flight number of $f'$.
3.2.2.4. Modeling Shuttle Services

In the POAGG shuttle service model, an airline schedules a group of aircraft to fly back and forth between an airport pair, with departures on a regularly scheduled basis. The shuttle service model is a variation of the model for generating itineraries for regular, i.e., non-shuttle, originating and connecting flights. The shuttle model requires as input: the airport pair, start and stop times, airline information for flights in the shuttle, and the interval between departures of shuttle flights.

*Calculating the Number of Shuttle Aircraft*

The shuttle model first determines the minimum number of aircraft required to schedule a particular shuttle service. The model assumes that the first shuttle flights of a day depart simultaneously from both airports. Therefore, there is a minimum requirement of one aircraft at the start time at each airport. The example in Figure 3.10 displays a scenario, in which two aircraft are sufficient for all shuttle flights. \(\alpha\) is the minimum layover time between flights.

![Graphical View of Shuttle Service](image)

*Figure 3.10: Graphical View of Shuttle Service*

In Figure 3.10, \(l(1)\) and \(l(2)\) represent the layover times at Airports 1 and 2 respectively. Additional aircraft will be required if either \(\alpha + g_{1,2} > h\) or \(\alpha + g_{2,1} > h\) is true. For example, if \(\alpha + g_{1,2} > h\) occurs, it will be infeasible to schedule Aircraft 1 for the second departure flight of
the shuttle service at Airport 1. As a result, an additional aircraft will be required in order to achieve the desired frequency of departures. This scenario is illustrated in Figure 3.11.

![Three-Aircraft Shuttle Service Diagram]

**Figure 3.11: Three-Aircraft Shuttle Service**

In Figure 3.11, Aircraft 3 is scheduled for the second shuttle departure at Airport 1. Aircraft 1 receives an extended layover time and is scheduled as the third shuttle departure at Airport 1.

The variable \( n_i \) is defined as

\[
n_i = \text{number of aircraft required in a shuttle service, originating from airport } i
\]

The formula for calculating \( n_i \) for shuttle airport pair \( i,j \) is

\[
n_i = \left\lfloor \frac{g_{i,j} + \alpha}{h} \right\rfloor
\]

Equation (29) has been verified through empirical analysis.
• **Generating Shuttle Flight Itineraries**

Shuttle flights are either originating or connecting flights. Generating itineraries for flights in a shuttle service differs from generating regular flight itineraries. The first flight for any aircraft in a shuttle service is an originating flight. In a shuttle service, an originating flight’s arrival and departure airports are always the airports in the shuttle airport pair, and assigned based on the direction of the flight. Airline information is set according to the shuttle input parameters for carrier name and aircraft type. The arrival and departure times for shuttle flights are based upon the departure times, which are determined by the start time of the shuttle service and the frequency of departures.

The first flights of Aircraft 1, 2 and 3, in Figure 3.11, are originating flights. All remaining flights in the shuttle service are connecting flights. Whereas layover times are generated randomly for regular flight itineraries, shuttle flight layover times are determined by the departure time of the aircraft’s next flight. The “previous flight”, “next flight”, and airline information data fields are completed as for regular connecting flights.

Generating flight itineraries for a shuttle service aircraft terminates after a shuttle flight arrival time exceeds the stop time for both airports. In this manner, the last flights in a shuttle service are scheduled to depart on or before the stop time.

The shuttle model may be utilized to bias the departure times within a generated schedule by designing multiple shuttle services between sets of airport pairs. A generated schedule with multiple shuttle services will have departure times which are biased according to values of the start times and frequencies of departures. For example, multiple hourly shuttle services, starting on the hour, involving a specific airport will bias departure times on the hour throughout the day.

### 3.2.2.5. Modeling Airline Hubs

Airline hubs are modeled by constructing arrival and departure banks within the hypothetical schedules. Whereas shuttle service modeling is a separate process, arrival and departure bank modeling is integrated within the second stage of flight schedule generation.

The relationship between an arrival bank and a departure bank is depicted in Figure 3.12. The arrows in Figure 3.12 represent individual flights.
Figure 3.12: Arrival and Departure Banks

The parameter $\delta$ is the minimum objective ground time. $\delta$ represents the time interval established by an airline, that separates the last arrival flight in an arrival bank from the first departing flight in a departure bank. The model assumes all flights in a departure bank are connecting flights of the flights in the arrival bank, which is a reasonable approximation.

Generating itineraries for flights in an arrival and departure bank pair requires the following input parameters:

- $i_H$ - airport for hub modeling
- $\delta$ - minimum objective ground time
- $n_b$ - number of flights in the arrival bank
- $\beta$ - start time of arrival bank
- $\chi$ - end time of arrival bank

The process also requires the hub airline carrier name.

• Constructing an Arrival Bank

An arrival bank of flights is constructed by interrupting the flight schedule generation process and modifying previously generated flight itineraries. The process is interrupted immediately after period $t_\chi$, defined as
At \( t_x \), all itineraries for originating and connecting flights for \( t \leq t_x \) have been generated for airport \( i_H \). The hub model identifies all scheduled arrivals into airport \( i_H \) during the interval \((\beta, \chi)\). Of the total arrivals, \( n_b \) are selected as candidate arrival bank flights.

In the model, a candidate arrival bank flight cannot be a shuttle flight, nor can the flight be already assigned to an arrival or departure bank. If the total number of available candidate flights exceeds the required number of flights, \( n_b \) of the total flights are selected at random. The carrier name for each selected flight is modified to the carrier name of the hub airline. The aircraft type of each selected flight is modified to reflect an aircraft type in the hub airline's fleet. If a selected flight is a connecting flight, the model also converts all previous flight itineraries of the same aircraft to reflect the change in carrier name and aircraft type. If, while converting previous flight itineraries, the model discovers that a previous flight was assigned to an arrival or departure bank of a different airline, the candidate flight is disqualified from the arrival bank and an alternate flight is selected as a replacement, if one is available.

After selecting and converting all candidate flights, the model identifies the last arriving flight in the arrival bank, \( f_A \). The arrival and departure times of \( f_A \) are modified such that \( o_{f_A} = \chi \). Figure 3.13 demonstrates this process. The gray arrow represents the original scheduled arrival position of \( f_A \).

![Figure 3.13: Arrival Times in Arrival Bank](image)

The construction of the arrival bank is complete.
• Constructing a Departure Bank

After period $t_x$, all connecting flight itineraries for arrivals at $i_H$ on or before $t_x$ have been generated. The departure bank is constructed by modifying the departure times of the connecting flights of the arrival bank flights. The departure times are adjusted to conform to the parameters of the departure bank. The model assumes the departure bank has the same number of flights as the arrival bank, and the time interval of the departure bank is equivalent in length to the time interval of the arrival bank, i.e., $\chi - \beta$. The departure bank parameters are

\[
\begin{align*}
\tau & \quad \text{- start time of departure bank} \\
u & \quad \text{- end time of departure bank}
\end{align*}
\]

where

\[
\begin{align*}
\tau &= \chi + \delta \\
u &= \chi + \delta + (\chi - \beta) = 2\chi - 2\beta
\end{align*}
\]

The hub algorithm generates $n_h - 1$ departure times randomly from a uniform distribution over the interval $(\tau, u)$. The departure time for the first departure bank flight will be $\tau$. The existing connecting flight itineraries are modified randomly to reflect $n_h - 1$ generated departure times. As a result, the departure sequence of aircraft is shuffled with regard to the arrival sequence.

Due to the fact that departure times of connecting flights are modified, the arrival times into other airports are also affected. It is necessary to verify that the arrival time adjustments of departure bank flights bound for specified airports do not violate any arrival distributions.

Constructing an arrival and departure bank pair may be infeasible. For example, adjusting the arrival times of connecting flights may violate an arrival distribution. Modifying layover times in search of available arrivals is limited by the interval of the departure bank. The problem is also infeasible if the arrival bank interval $(\beta, \chi)$ does not contain at least $n_h$ qualified candidate arrival flights.
3.3. Implementation of the Model

The algorithm for generating hypothetical flight schedules is implemented in a
computer application, known as the Pseudo-OAG Generator or POAGG. POAGG has been
developed at the Charles Stark Draper Laboratory (CSDL) on a Macintosh platform. The
algorithms and sub-algorithms of POAGG are coded in C. POAGG operates with a user
interface for the Macintosh written in C++.

POAGG can function as a standalone application or as a support module for the CSDL
ATC Flow Management Simulation Testbed. The Testbed is a simulation environment for
analyzing alternative flow management strategies. The Testbed will accept either actual OAG
flight schedules or POAGG generated schedules as air traffic demand input. The structure and
operation of the POAGG application is illustrated in Figure 3.14.

![Diagram of POAGG Flow Chart]

*Figure 3.14: POAGG Flow Chart*
3.3.1. POAGG Functionality

POAGG is a user-driven application. The user constructs a network of specified airports by defining parameters that describe each airport. The components of the flow chart represent different phases of the schedule-generation process.

The following list is a summary of all input parameters for POAGG. The respective variables from the model are provided in parentheses. The general parameters are:

- number of specified airports (N)
- period length (θ)
- cutoff time for connecting flights to Airport X
- probability mass functions (PMFs) for flight times to Airport X
- minimum required layover time (α)
- maximum allowable layover time (λ)
- upper parameter for layover probability density function (PDF) (m)

For every specified airport, the parameters are:

- three-letter airport name, e.g., JFK
- total scheduled arrivals (e_i)
- relative frequency distribution of arrivals (k_{i,t})
- PMF and carrier names for airlines operating at the airport
- PMFs for airline fleets

Input parameters describing the relationships between airport pairs are:

- directed scheduled flight times (g_{i,j})
- connection percentages (p_{i,j})

Shuttle input parameters are:

- airport pair
- carrier name
- aircraft type
- start time
- stop time
- interval between shuttle flight departures (h)

Arrival/Departure banks input parameters are:

- hub airport name
- arrival bank start time (β)
Chapter 3: Airport Demand Schedule Modeling

- arrival bank stop time (τ)
- minimum objective time (δ)
- number of flights in arrival bank (n_b)

The feasibility subroutine depicted in the flow chart in Figure 3.14 will be described separately in Section 3.3.2 below. After the input parameters have been checked or modified for feasibility, the process transitions into Stage 1. The connection algorithm in Stage 1, assigns values to the intermediate variables that will govern the generation of flight itineraries during Stage 2. Stage 1 distributes the desired number of connecting flights between all specified airports over the day by time period.

If the user requests one or more shuttle services in the schedule, POAGG enters the shuttle services modeling phase. All shuttle flight itineraries are generated prior to generating regular flight itineraries. The essence of a shuttle service is that shuttle flights will be scheduled with a fixed group of aircraft, dedicated to traveling back and forth between the shuttle airport pair. A shuttle flight will never connect outside of the network of specified airports.

After all shuttle flight itineraries are generated, or if no shuttles services are requested, POAGG generates the remaining flight itineraries, creating a schedule that satisfies the parameters defined by the user. The algorithm advances period-by-period, generating itineraries and assigning connecting flights according to the values of the input parameters and intermediate variables. Arrival and departure banks are constructed by temporarily interrupting the schedule-generation process. Previously generated flight itineraries are modified in order to conform to the parameters of the arrival and departure banks. After an arrival-departure bank pair are successfully constructed the schedule-generation process is resumed.

The output of POAGG is an OAG-like flight schedule. Each flight contains the following data fields: arrival airport, arrival time, departure airport, departure time, flight number, carrier name, aircraft type, “previous flight”, and “next flight”. The user may save the generated flight schedule in a text file. Since POAGG is fully integrated in the Draper ATC Testbed, the user may also transmit the generated schedule electronically into the Testbed for use as airport demand data, in the analysis of flow management strategies.
3.3.2. Feasibility Checks

Prior to generating a flight schedule, POAGG executes a series of feasibility checks on the input parameters. If infeasibility is detected, the user is offered three alternatives. The user may: 1) abort the schedule generation attempt, 2) manually modify the values of the input parameters, or 3) request that the application modify the parameters using built-in correction algorithms. For the third alternative, the application invokes a subroutine designed to seek new values for the input parameters. The objective of the subroutine is to achieve feasibility with minimal modifications to the input parameters.

3.3.2.1. Total Scheduled Arrivals versus Connecting Flights

Generating a flight schedule is infeasible if the user attempts to schedule more connecting flights into an airport, than requested total arrivals into the airport. Mathematically,

\[ \sum_{i=1}^{N} d_{i,j} \leq e_{j} \quad \forall j \]  

(32)

The feasibility subroutine offers the user two options if (32) is violated: 1) lowering the connection percentages, \( p_{i,j} \), or 2) raising the total scheduled arrivals, \( e_{i} \). The parameters are minimally modified by formulating and solving linear programs. In the formulations, \( e_{i}' \) and \( p_{i,j}' \) are the modified values of \( e_{i} \) and \( p_{i,j} \), respectively. The formulation that minimally raises total arrivals is

\[
\min \sum_{i=1}^{N} e_{i}'
\]

(33)

s.t. \[ \sum_{i=1}^{N} p_{i,j} \cdot e_{i}' \leq e_{j}' \quad \forall j \]  

(34)

\[
e_{i}' \geq e_{i} \quad \forall i
\]

(35)

Feasibility for the above formulation is trivial. Raising all values of \( e_{i}' \) to the greatest value of \( e_{i} \) will always produce a feasible solution.

The linear program that minimally lowers the connection percentages is
\[
\max \sum_{i=1}^{N} \sum_{j=1}^{N} p_{i,j} \quad (36)
\]

subject to \[
\sum_{i=1}^{N} p_{i,j} \cdot e_i \leq e_j \quad \forall j \quad (37)
\]

\[
p_{i,j} \leq p_{i,j} \quad \forall i,j \quad (38)
\]

\[
p_{i,j} \geq 0 \quad \forall i,j \quad (39)
\]

Feasibility is again trivial. The formulation may be solved by lowering all connection percentages to 0.

3.3.2.2. Shuttle Service versus Connecting Flights

The connecting flights in a shuttle service between an airport pair will achieve a connection percentage between the airport pair, referred to as \(p_{i,j}^s\). If the user requests shuttles services between an airport pair, for which \(p_{i,j}^s > p_{i,j}\), there is a conflict between the shuttle input parameters and the connection percentages. The solution offered by POAGG is to raise \(p_{i,j}\) to the value of \(p_{i,j}^s\). If \(p_{i,j}\) is modified, the feasibility subroutine returns to the feasibility check in Section 3.3.2.1, permitting only modifications to total arrivals.

3.3.2.3. Shuttle Service versus Arrival Distribution

Shuttle service flights are scheduled as a function of the parameter describing the interval between departures. The exact arrival times of all shuttle flights can be calculated prior to generating the shuttle flight itineraries. If \(q\) is the shuttle service start time, the \(n\)th shuttle departure from airport \(\hat{i}\) to airport \(\hat{j}\) will arrive at

\[
o_{f_n} = q + (n-1) \cdot h + g_{i,j}\quad (40)
\]

where \(h\) is the departure interval, and \(g_{i,j}\) is the flight time from \(\hat{i}\) to \(\hat{j}\).

The feasibility subroutine verifies that the scheduled arrivals of shuttle flights does not exceed the parameters of the arrival distributions. For example, assume \(a_{i,j}^s\); represents the scheduled arrivals at airport \(\hat{i}\) in period \(\hat{i}\) achieved by all shuttle services. If \(a_{i,j}^s > a_{i,j}\), a conflict exists between the arrivals of shuttle flights and the values of the arrival distribution parameters.
If requested by the user, POAGG corrects the conflict by raising the value of \( a_{i,j} \), until \( a_{i,j}^S = a_{i,j} \). At the same time, in order to avoid violating (12), i.e.,

\[
\sum_{i=1}^{T} a_{i,t} = e_i \quad \forall i
\]  

the subroutine decreases the values of \( a_{i,t} \) in periods other than period \( \hat{t} \). Values of \( a_{i,t} \) are only decreased for periods in which \( a_{i,t}^S < a_{i,t} \). The subroutine seeks to adjust \( a_{i,t} \) in periods closest to \( \hat{t} \).

3.3.2.4. Arrival Banks versus Arrival Distribution

Constructing an arrival bank is infeasible if there are insufficient candidate flights arriving at the hub airport during the interval of the user designated arrival bank. If \( n_a \) represents the arrivals during the interval of the arrival bank, then the bank is infeasible if \( n_a < n_b \). If requested by the user, POAGG will decrease the number of flights in the arrival bank until \( n_a = n_b \).

3.4. Test Results

Multiple flight schedules have been generated by POAGG and verified as satisfying the requested input parameters. The following test cases will be presented in this section:

Test Case 1 - Generating a flight schedule using input parameters derived from an actual OAG schedule.

Test Case 2 - Generating flight schedules for 5, 10, 15, and 20 airports with varying total connection percentages, flight times and arrival distributions.

Test Case 3 - Generating flight schedules for 5 airports with multiple shuttle services.

Test Case 4 - Generating a flight schedule for 10 airports with hubbing operations at multiple airports.

The test cases were designed to test the robustness of POAGG. The input parameters for the test cases were established to reflect scenarios characteristic of major airports in the United States. Schedules were analyzed in order to assess if POAGG preserves the desired airport demand characteristics, i.e., flight connections, shuttle services and hubbing operations. In addition, the performance of POAGG was measured by tracking the total number of flights.
with layovers that exceeded the maximum allowable layover time for each generating schedule. This metric is derived from the objective function of the heuristic solution technique, described in Section 3.2.2.1. A POAGG “run” refers to the generation of an individual schedule. All test cases were executed on a Macintosh Quadra 950.

The following descriptions of input parameters apply to Test Cases 2, 3, and 4. The connection percentages were derived by introducing a new parameter which represents the total connection percentage parameter: \( p'' \). Individual connection percentages between all airport pairs for a POAGG run were calculated using the equation

\[
p_{i,j} = \frac{p''}{N-1} \quad \forall i, j
\]  

(41)

where

\[
p_{i,i} = 0 \quad \forall i \quad \text{and} \quad 0.0 \leq p'' \leq 100.0
\]

Total arrivals for each airport were generated from a uniform distribution over the interval (400 flights, 600 flights), which simulates total arrivals at major airports. Flight times were generated based upon airport relationships depicted in Figure 3.15.

![Diagram of airport groups with flight times](image)

**Figure 3.15: Test Cases - Flight Times**

Each clustered airport group contains 5 airports, and represents a geographical region. Within any group, flight times among airports were generated from a uniform distribution over the interval \( U(1 \text{ hour}, 2 \text{ hours}) \). Flight times between airports in adjacent groups were generated over the interval \( U(2 \text{ hours}, 3 \text{ hours}) \). For airports that were two groups apart, the interval was \( U(3 \text{ hours}, 4 \text{ hours}) \), and for airports in Groups I and IV, the interval was \( U(4 \text{ hours}, 5 \text{ hours}) \). For Test Cases 2, 3, and 4, the number of airports, \( N \), was always a multiple of 5.
The flight times between each airport pair \( g_{i,j} \) and \( g_{j,i} \) were assigned correlated values. That is, after \( g_{i,j} \), for any airport pair, was generated according to the intervals in Figure 3.15, then \( g_{j,i} \) was subsequently calculated using the equation

\[
    g_{j,i} = g_{i,j} + v
\]

\( v \) is a random number realized over the interval \( U(-10 \cdot x, 10 \cdot x) \) where

\[
    x = \begin{cases} 
        1 & \text{if airports } i \text{ and } j \text{ are in the same group} \\
        2 & \text{if airports } i \text{ and } j \text{ are in adjacent groups} \\
        3 & \text{if airports } i \text{ and } j \text{ are separated by one group} \\
        4 & \text{if airports } i \text{ and } j \text{ are separated by two groups}
    \end{cases}
\]

In this manner, the expected differences of the directed flight times between two airports increases as the airports are further apart.

The relative frequency distributions for each airport for Test Cases 2, 3 and 4 were selected at random from a set of three actual OAG distributions. The distributions were derived from a January 1992, OAG flight schedule by examining the arrivals at Boston Logan (BOS), LaGuardia (LGA), and Washington National (DCA) airports. The distributions are depicted in Figures 3.16, 3.17, and 3.18.

![Figure 3.16: Relative Frequency Distribution of Arrivals for Logan, January 1992](image_url)
Chapter 3: Airport Demand Schedule Modeling

Figure 3.17: Relative Frequency Distribution of Arrivals for LaGuardia, January 1992

Figure 3.18: Relative Frequency Distribution of Arrivals for National, January 1992

The remaining input parameters used in all test cases were:

- period length: 30 minutes
- cutoff time for connecting flights to Airport X: 10:00 PM
- PMF for flight time to and from Airport X:
  \[ P(g_X = 112 \text{ minutes}) = 0.4 \]
  \[ P(g_X = 156 \text{ minutes}) = 0.4 \]
  \[ P(g_X = 136 \text{ minutes}) = 0.15 \]
  \[ P(g_X = 235 \text{ minutes}) = 0.5 \]
- minimum required layover time: 30 minutes
- maximum allowable layover time: 120 minutes
upper parameter for layover PDF: 60 minutes

The difference between the maximum layover time and the upper parameter for the PDF is described in Section 3.2.2.3.5. The PDF for the generation of layover times is displayed in Figure 3.19.

![PDF Diagram](image)

**Figure 3.19: Test Cases - Layover Time PDF**

### 3.4.1. Test Case 1: OAG Flight Parameters

Test Case 1 consisted of deriving arrival and connection parameters from an actual flight schedule (OAG, January 1992), and using the parameters as input for a POAGG run. The specified airports in Test Case 1 were BOS, LGA, and DCA. Two shuttle services were requested for BOS-LGA and LGA-DCA. Flight times among the airports were selected from a sample of flights in the OAG schedule. The input flight times were:

- BOS to LGA: 65 minutes
- BOS to DCA: 80 minutes
- LGA to DCA: 65 minutes
- LGA to BOS: 60 minutes
- DCA to BOS: 85 minutes
- DCA to LGA: 60 minutes

The fact that the above scheduled flight times are greater than actual flight times between the airport pairs indicate that airlines pad flight itineraries to account for ground taxi time and possibly in anticipation of unexpected delays.

The Test Case 1 generated schedule satisfied all parameter inputs derived from the OAG schedule. The desired and realized connection percentages are depicted in Figure 3.20. The numbers over the columns represent actual connecting flights.
Chapter 3: Airport Demand Schedule Modeling

Figure 3.20: Test Case 1 - Connection Results

The slight difference in percentages are due to quantization error. The desired and realized total arrivals were: BOS - 595 flights, LGA - 463 flights, and DCA - 362 flights.

Table 3.3 contains the desired and realized airline information for Test Case 1. In the actual OAG schedule, each airport had more than six operating airlines each. The input parameters in Table 3.3 contain a sampling of the major airlines at each airport. Similarly, the input aircraft types are a sampling of the aircraft types in each airline's fleet.

The column headings under "Aircraft Type:..." are the input percentages for the PMFs describing the airline fleets. The values in parentheses are the realized percentages for each aircraft type for each airline. For example, at Boston, the input PMF for AA's airline fleet was specified as

\[
P(\text{aircraft type} = \text{M80}) = 0.4 \\
P(\text{aircraft type} = \text{72S}) = 0.4 \\
P(\text{aircraft type} = \text{D10}) = 0.2
\]

The realized aircraft type percentages for AA at BOS were for M80, 37.05%, for 72S, 40.24% and for D10, 22.71%.

The "other" category reflects arrival flights from carriers outside of the assigned carriers at an airport. The presence of these flights indicates that connecting flights were scheduled into an airport from airports with different desired airlines. For example, assume a US arrival flight at LGA is assigned a connecting flight into BOS. The connecting flight will also be a US flight, however, since US is not an assigned carrier at BOS, the connecting flight will classify as "other" at BOS.
### Table 3.3: Test Case 1 - Airline Results

<table>
<thead>
<tr>
<th>Airport</th>
<th>Airline</th>
<th>Desired %</th>
<th>Realized %</th>
<th>Aircraft Type: Desired and (Realized %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOS</td>
<td>NW</td>
<td>20.00</td>
<td>17.82</td>
<td>D9S (30.19) 320 (50.00) D08 (19.81)</td>
</tr>
<tr>
<td></td>
<td>DL</td>
<td>40.00</td>
<td>36.81</td>
<td>72S (35.62) SH6 (44.29) SF3 (20.09)</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>40.00</td>
<td>42.18</td>
<td>M80 (37.05) 72S (40.24) D10 (22.71)</td>
</tr>
<tr>
<td></td>
<td>(other)</td>
<td>40.00</td>
<td>3.19</td>
<td></td>
</tr>
<tr>
<td>LGA</td>
<td>AA</td>
<td>35.00</td>
<td>28.51</td>
<td>M80 (37.88) 72S (40.91) D10 (21.21)</td>
</tr>
<tr>
<td></td>
<td>US</td>
<td>15.00</td>
<td>18.36</td>
<td>100 (30.59) 73S (47.06) DH7 (22.35)</td>
</tr>
<tr>
<td></td>
<td>DL</td>
<td>50.00</td>
<td>51.19</td>
<td>72S (43.88) SH6 (38.82) SF3 (17.30)</td>
</tr>
<tr>
<td></td>
<td>(other)</td>
<td>50.00</td>
<td>1.94</td>
<td></td>
</tr>
<tr>
<td>DCA</td>
<td>DL</td>
<td>30.00</td>
<td>31.49</td>
<td>72S (42.98) SH6 (40.35) SF3 (16.67)</td>
</tr>
<tr>
<td></td>
<td>US</td>
<td>40.00</td>
<td>39.78</td>
<td>100 (36.11) 73S (49.31) DH7 (14.58)</td>
</tr>
<tr>
<td></td>
<td>UA</td>
<td>30.00</td>
<td>22.93</td>
<td>733 (48.19) 727 (25.30) 72S (26.51)</td>
</tr>
<tr>
<td></td>
<td>(other)</td>
<td>30.00</td>
<td>5.80</td>
<td></td>
</tr>
</tbody>
</table>

Two shuttles services were requested in order to simulate actual shuttle services between the represented airports. The shuttle inputs were

**Shuttle Service 1 - BOS, LGA**
- **Airline:** DL  
- **Aircraft Type:** 72S  
- **Start Time:** 6:30 AM  
- **Stop Time:** 10:00 PM  
- **Frequency:** hourly

**Shuttle Service 2 - LGA, DCA**
- **Airline:** US  
- **Aircraft Type:** 73S  
- **Start Time:** 7:00 AM  
- **Stop Time:** 9:00 PM  
- **Frequency:** every 90 minutes

Both shuttle services were successfully modeled. Each service required more than 2 aircraft to complete the scheduling. Figure 3.21 contains the flight itineraries of a single aircraft for Shuttle Service 1. The spreadsheet in Figure 3.21 is an image of the flight display “window” in POAGG.
Due to the fact that none of the three airports in the test case are hub airports, there was no attempt to construct airline and departure banks.

Figure 3.22 contains a small sample of regular flight (non-shuttle) itineraries from the generated schedule in Test Case 1, sorted by arrival airport. AX2 and AX3 represent Airport X, with differing flight times. The total schedule contained 2536 total flight itineraries.

The generated schedule for Test Case 1 contained one flight with a layover time exceeding 2 hours. This flight was identified in the schedule as the first arrival at Logan, with a 1:44 AM arrival time (period 4). The flight was scheduled to have a connecting flight to LaGuardia, however, due to the arrival distribution at LaGuardia, the first arrival was not permitted until 6:24 AM (period 13).

### 3.4.2. Test Case 2: Connections

In Test Case 2, flights were generated for \( N = 5, 10, 15 \) and 20 airports. Initially, four schedules were generated for each value of \( N \), with total connection percentages of \( p'' = 25.0\% \), \( 50.0\% \), \( 75.0\% \), and \( 100.0\% \). POAGG runs with \( p'' = 100.0\% \) were extreme scenarios in which all flights connected within the network of specified airports, precluding flights to or from Airport
X. Airline and aircraft fleet input parameters were simplified to a single airline with a single
aircraft type at all airports.

All generated schedules for Test Case 2 conformed to the values of the arrival and
collection parameters. Table 3.4 contains the verification statistics for the N=5 and $p''=50.0\%$
run.

Table 3.4: Test Case 2 - Results for N=5, $p''=50.0\%$

<table>
<thead>
<tr>
<th>Airport</th>
<th>Total Arrivals: Desired: 467 Realized: 467</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Next day arrivals: 0</td>
</tr>
<tr>
<td></td>
<td>Deviations from RFD of Arrivals: 0</td>
</tr>
<tr>
<td>Connections:</td>
<td>To</td>
</tr>
<tr>
<td>Airport 2</td>
<td>12.5</td>
</tr>
<tr>
<td>Airport 3</td>
<td>12.5</td>
</tr>
<tr>
<td>Airport 4</td>
<td>12.5</td>
</tr>
<tr>
<td>Airport 5</td>
<td>12.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Airport</th>
<th>Total Arrivals: Desired: 504 Realized: 504</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Next day arrivals: 1</td>
</tr>
<tr>
<td></td>
<td>Deviations from RFD of Arrivals: 0</td>
</tr>
<tr>
<td>Connections:</td>
<td>To</td>
</tr>
<tr>
<td>Airport 2</td>
<td>12.5</td>
</tr>
<tr>
<td>Airport 3</td>
<td>12.5</td>
</tr>
<tr>
<td>Airport 4</td>
<td>12.5</td>
</tr>
<tr>
<td>Airport 5</td>
<td>12.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Airport</th>
<th>Total Arrivals: Desired: 548 Realized: 548</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Next day arrivals: 0</td>
</tr>
<tr>
<td></td>
<td>Deviations from RFD of Arrivals: 0</td>
</tr>
<tr>
<td>Connections:</td>
<td>To</td>
</tr>
<tr>
<td>Airport 2</td>
<td>12.5</td>
</tr>
<tr>
<td>Airport 3</td>
<td>12.5</td>
</tr>
<tr>
<td>Airport 4</td>
<td>12.5</td>
</tr>
<tr>
<td>Airport 5</td>
<td>12.5</td>
</tr>
</tbody>
</table>
### Airport 4

Total Arrivals: Desired: 427  Realized: 427  
Next day arrivals: 0  
Deviations from RFD of Arrivals: 0  

<table>
<thead>
<tr>
<th>Connections</th>
<th>To</th>
<th>% Desired</th>
<th>% Realized</th>
<th>Flights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Airport 1</td>
<td>12.5</td>
<td>12.41</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Airport 2</td>
<td>12.5</td>
<td>12.41</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Airport 3</td>
<td>12.5</td>
<td>12.41</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Airport 5</td>
<td>12.5</td>
<td>12.41</td>
<td>53</td>
</tr>
</tbody>
</table>

### Airport 5

Total Arrivals: Desired: 460  Realized: 460  
Next day arrivals: 0  
Deviations from RFD of Arrivals: 0  

<table>
<thead>
<tr>
<th>Connections</th>
<th>To</th>
<th>% Desired</th>
<th>% Realized</th>
<th>Flights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Airport 1</td>
<td>12.5</td>
<td>12.39</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Airport 2</td>
<td>12.5</td>
<td>12.39</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Airport 3</td>
<td>12.5</td>
<td>12.39</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Airport 4</td>
<td>12.5</td>
<td>12.39</td>
<td>57</td>
</tr>
</tbody>
</table>

Minimum Layover Time: 30 minutes  
Maximum Layover Time: 115 minutes  
Average Layover Time: 44.86 minutes  
Total Flight Layovers Exceeding 2 hours: 0

The next day arrival at Airport 2 refers to a connecting flight from a specified airport that was scheduled to land after midnight. The performance of POAGG for Test Case 2 is shown in Figure 3.23.
Figure 3.23: Test Case 2 - Flights with Layovers Exceeding 2 Hours

Figure 3.23 demonstrates that as the connection percentages and number of airports increased, more flights were assigned layover times which exceeded 2 hours. These results are predictable. As $N$ increases the total number of flight itineraries generated for each schedule increases, and as a result a higher number of flights will have excessive layovers. The increase appears proportional to the increase in number of airports. The decrease from $N=15$ to $N=20$, for 75% total connections, is attributed to the randomness of the generated arrival data and flight times.

Flight times with layover times over 2 hours increased significantly as total connection percentages reached 100%. This is expected, due to the nature of the second correction technique in the connection algorithm. The second technique searches for connection opportunities earlier in the day, by identifying flights with connections to Airport X. For $p''=100\%$, every flight is assigned a connecting flight to a specified airport, and thus, the technique will never find an earlier flight connecting to Airport X. The connection algorithm returns to the first correction technique which involves extending layover times, in search of available arrivals in later periods.

In order to assess the relative effectiveness of POAGG, Figure 3.24 graphs the proportion of flights with layover times exceeding 2 hours, with respect to the total number of connecting flights between specified airports in the generated schedules.
In the worst case scenarios, i.e., $p''=100\%$, the percentage of flights with layovers over 2 hours does not exceed 9%. POAGG reduces the proportion to under 3% for all runs with $p''=75\%$.

From Figures 3.23 and 3.24, it is evident that the POAGG connection algorithm is very effective for values of $p''$ at or below 50%. In order to gain a better assessment of POAGG performance above 50%, flight schedules were generated for $N=10$ and $p''=60\%, 70\%, 80\%$ and $90\%$. The results are plotted in Figure 3.25.
As before, the vertical axis represents the percentage of total connecting flights with layover times exceeding 2 hours. Figure 3.25 reveals that the connection algorithm contains flights with excessive layovers below 3% for all test runs except for $p'' = 100\%$.

Figure 3.26 displays the total number of flight itineraries generated for each run in Test Case 2.

![Bar chart showing total flights for different connection percentages and airport counts](image)

**Total Connection Percentage ($p''$)**

*Figure 3.26: Test Case 2 - Total Flights*

Generally, the number of flight itineraries is higher for lower total connection percentages, due to the fact that few arrivals at specified airports are achieved by connecting flights. As the number of connecting flights increases, the number of total flight itineraries decreases. The increase in itineraries for $p'' = 100\%$ is due to the fact that many connecting flights are scheduled to arrive in the next day. The largest schedule (N=20, $p'' = 25\%$) contained 16916 flights.

Concerning the speed of the POAGG application, all schedules for N=5 and 10 were generated in under 2 minutes. Schedules for N=10 and 20 with $p'' = 25\%$ and $50\%$, were also generated in under 2 minutes. The remaining four schedules (N=10 and 20, $p'' = 75\%$ and $100\%$) required between 2 to 5 minutes.

As part of Test Case 1, an additional flight schedule was generated for N=5 and $p'' = 75\%$. The individual airport connection percentages, i.e., $p_{i,j}$, were manually modified to test POAGG with non-homogeneous airport connection percentages. The statistics of the resulting flight schedule confirmed that the schedule satisfied the input parameters. Table 3.5 contains the input parameter values and results for Airport 4.
Table 3.5: Test Case 2 - Results for Airport 4, N=5, \( p'' = 75\% \), Non-homogeneous \( p_{i,j} \)

<table>
<thead>
<tr>
<th>Connections to Airport</th>
<th>% Desired</th>
<th>% Result</th>
<th>Flights</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15.0</td>
<td>14.85</td>
<td>71</td>
</tr>
<tr>
<td>2</td>
<td>25.0</td>
<td>24.90</td>
<td>119</td>
</tr>
<tr>
<td>3</td>
<td>5.0</td>
<td>4.81</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>30.0</td>
<td>29.92</td>
<td>143</td>
</tr>
</tbody>
</table>

Table 3.5 compares the POAGG performance for \( N=5, \ p'' = 75\% \) for non-homogenous values of \( p_{i,j} \) and homogeneous values, where \( p_{i,j} = 18.75\% \) for all \( p_{i,j} \).

<table>
<thead>
<tr>
<th>( p_{i,j} ) varied</th>
<th>( p_{i,j} = 18.75 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>average layover</td>
<td>52.50</td>
</tr>
<tr>
<td>layovers &gt; 2 hours</td>
<td>43</td>
</tr>
<tr>
<td>% of connecting flights</td>
<td>2.39%</td>
</tr>
</tbody>
</table>

The third row contains the percentage of total connecting flights with layovers exceeding 2 hours for both schedules. It is assumed the differences can be attributed to the randomness of the input data.

3.4.3. Test Case 3: Shuttles

Two flight schedules were generated to verify the POAGG shuttle service model. The basic input parameters were \( N=5 \) and \( p'' = 50\% \). The first flight schedule was designed with a shuttle service by a single airline between each airport pair in the network, resulting in 10 shuttle services. The shuttle service parameters for each airport pair were:

- start time: 7:00 AM
- stop time: 9:00 PM
- frequency: hourly

The second flight schedule for Test Case 3, contained two shuttle services, by two independent airlines, between each airport pair in the 5 airport network, resulting in 20 shuttle
Chapter 3: Airport Demand Schedule Modeling

services. Shuttle service parameters for the first airline were the same as for the first schedule. Shuttle service parameters for the second airline for each airport pair were:

start time: 8:00 AM  
stop time: 10:00 PM  
frequency: every 90 minutes

Both schedules contained sufficient shuttle flights to meet the shuttle input parameters, and satisfied all additional input parameters. Neither run generated any flight itineraries with layovers exceeding 2 hours.

3.4.4. Test Case 4: Airline Hubbing

Test Case 4 involved generating a flight schedule with arrival and departure banks, in order to model hubbing operations. The schedule was generated for a network of 10 airports with $p^{*}$=50%. 5 airports were modeled as hub airports for two independent airlines. Arrival and departure banks were designed in the morning for one airline, and in the afternoon for the other airline. The parameters for the banks were:

Arrival Bank 1
  Interval: 8:00 AM to 9:30 AM  
  Minimum Objective Ground Time: 35 minutes  
  Number of Flights: 45

Arrival Bank 2
  Interval: 2:00 PM to 4:00 PM  
  Minimum Objective Ground Time: 35 minutes  
  Number of Flights 45

The generated schedule satisfied all input parameters. The schedule contained 281 flights with layovers over 2 hours. This high statistic is attributed to the technique of shuffling the departure times of flights in the departure banks. Some flights arriving early in arrival banks were randomly assigned connecting flights with late departure times in the departure bank, resulting in long layover times. Furthermore, the interval of Arrival Bank 2 was 2 hours. Therefore, all flights arriving in the first 35 minutes of Arrival Bank 2 received in excess of 2 hours of layover time, independent of the sequence of departures.
3.5. Conclusion and Extensions

The airport demand model presented in this chapter offers the capability to generate OAG-like schedules for multiple airports with variable connection percentages. The current implementation of the model in the POAGG application is fast and robust. The stand-alone application has been delivered to multiple air traffic flow management researchers as a support tool. The application has also been used extensively as a support module within the Draper ATC Testbed.

The numerical method implemented in POAGG effectively models commercial demand for a network of specified airports. The method can potentially be extended to model general aviation (GA) and military air traffic. One alternative is to insert GA and military flights into the commercial schedule using a Monte Carlo method. GA and military flights could randomly be assigned arrival times into specified airports based upon historical or hypothetical distributions of GA and military airport demand.

Additionally, the POAGG numerical method can be extended to include a higher fidelity model of individual airline scheduling. Enhancements could include modeling flight times as airline and aircraft specific, and biasing schedule departure times based on airline preferences. The process of assigning connecting flights could be modified to model connections at the airline level. This would require input parameters and constraints that describe the scheduling practices of the individual airlines.
Chapter 4

4. Airport Weather and Capacity Modeling

In this chapter, we present research on the application of an existing weather model to support airport capacity modeling. The weather model was created for the United States Air Force (USAF) by Boehm [18] using multidimensional sawtooth waves. The model will be referred to as the sawtooth wave model (SWM). The objective of the research is to determine the suitability of the SWM for air traffic flow management (FM) research and simulation efforts.

This chapter is organized in four sections. The first section presents background information on FM simulation and the basic requirements of airport weather modeling. The second section describes the mathematics of the SWM. The third section contains an analysis of simulation test runs for airports at Boston, Washington and New York. The concluding section discusses the suitability of the model for FM research.

4.1. Background and Requirements

Through simulation and analysis, air traffic researchers are attempting to improve the efficiency of the air traffic system and reduce delays. Many FM automation aids are aimed specifically at improving operations at groups of airports. In order to analyze any particular tool in an air traffic simulation environment, the simulation must contain a model of system capacity, particularly the capacities of major airports where congestion and delays are common. The FAA maintains that 65% of all delays in the air traffic system are weather related [15]. Due to the dominant influence of weather, airport capacities must be modeled as a function of airport weather.

4.1.1. Airport Weather

Airports are the bottlenecks in the system, in that most congestion and related delays occur at airports or in airspace near airport terminal areas. At an airport, adverse weather will impact operations by increasing the required spacing intervals between aircraft or by reducing the number of usable runways, thus decreasing an airport’s capacity. The principal components of weather which influence airport capacity are cloud ceiling, visibility, wind
speed, wind direction, and precipitation. The term airport weather, in this chapter, will refer exclusively to these weather components.

4.1.1.1. Cloud Ceiling and Visibility

Cloud ceiling, or simply ceiling, refers to the base altitude of the lowest cloud in the airport area. Visibility refers to horizontal visibility at an airport, at ground level. These two aviation weather components are used to classify an airport as under Instrument or Visual Flight Rules, i.e., IFR or VFR. During periods of low visibility and/or low cloud ceiling, an airport operates in IFR. An IFR airport classification implies that the airport may only conduct landings on runways with special landing equipment. In addition, the aircraft and crews must be rated for IFR landings. The landing intervals between aircraft are greater under IFR, than VFR, and as a result, capacity is lower. Figure 4.1 specifies the visibility and ceiling thresholds for airport classifications. LIFR refers to Low IFR. MVFR is Marginal VFR. Different rules apply to each category.

![Figure 4.1: IFR and VFR Thresholds](image)

4.1.1.2. Wind Speed and Direction

Wind affects airport capacity by restricting the availability of runways. The prevailing wind vector at an airport is decomposed by tower personnel into a headwind component and a crosswind component for each airport runway. (A wind vector perpendicular to a runway will have zero headwind and full crosswind.) The tower supervisor selects active
runways from runways that are oriented for aircraft landings and takeoffs with headwind. Headwind reduces the speed of aircraft relative to the ground, thus requiring less runway length, and less time, for landings and departures. Since, the FAA requires that an aircraft safely exits a runway before a subsequent aircraft enters the same runway (either for a landing or a takeoff), as aircraft move off runways more quickly the overall capacity of an airport improves. Landing or taking off with headwind is also safer. Reduced takeoff and landing distances provide pilots with additional runway length that may be necessary if a pilot error or equipment failure occurs. The active configuration is selected from the available runways in order to maximize the capacity of an airport, accounting for other influencing factors, e.g. noise levels and maintenance issues.

The crosswind component affects the selection of the active runways. Generally, the maximum allowable crosswind for an active runway is 15 knots. While controllers may defer to a pilot's discretion concerning crosswinds, the tower supervisor determines the operational configuration considering runways with crosswind components of less than 15 knots. It is conceivable, therefore, that wind speed and direction may dictate an active configuration with less than optimal capacity. Potentially, due to wind, capacity may be insufficient to meet airport demand.

4.1.1.3. Precipitation

Precipitation impacts airport operations by increasing the required runway length for operations. Aircraft maneuver on and off runways at slower speeds if precipitation is present, reducing arrival and departure rates.

4.1.2. Definitions

Prior to discussing the requirements of an airport weather model, some terms will be defined in order to clarify their use in this chapter. The following definitions are verbatim from the USAF Technical Note 82/004, Basic Techniques in Environmental Simulation [17].

* Model

A model is a representation or imitation of a system or process (e.g., weather) through a different medium (e.g., mathematics). A model is a generalization of a more complex reality usually involving simplifying assumptions. A good model accurately and realistically represents the behavior being imitated while preserving the essential properties of the system.
• **Simulation**

A simulation is a realistic synthesis of a behavior consistent in space and time achieved using various techniques, e.g. computers. Most weather simulations are constructed as support modules within a primary simulation to analyze the environmental affects on a system or policy, e.g. FM. A simulation produces *synthetic* outputs (e.g., weather observations and forecasts). A model is implemented within a simulation.

• **Correlation**

Correlation is a measure of the relatedness or association among variables. Statistically, correlation can be quantified by the theoretical coefficient of correlation $\rho$ between two random variables. Given $X$ and $Y$ are two random variables with respective means of $\mu_X$ and $\mu_Y$ and standard deviations of $\sigma_X$ and $\sigma_Y$, then

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} \quad (1)$$

where

$$\sigma_{XY} = Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \quad (2)$$

Equations (1) and (2) can reflect both positive and negative correlation. If related values of $X$ and $Y$ vary from their respective means in the same direction, the resulting covariance will be positive. Conversely, if $X$ and $Y$ vary in opposite directions from their means, the covariance will be negative resulting in negative correlation.

For problems involving the sampling of actual data, e.g. wind speed measurements or cloud ceiling observations, correlation can be measured by the sample correlation coefficient $r$, which can be calculated by the Pearson Product Moment (PPM) formula

$$r_{XY} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \quad (3)$$

where $X_i$ and $Y_i$ are sample observations, and $\bar{X}$ and $\bar{Y}$ are sample means. Both $\rho$ and $r$ can vary between -1.0, implying perfect negative correlation, and 1.0 implying perfect positive correlation. The PPM assumes that the two random variables are related through a linear function. All references to correlation in this chapter refer to linear correlation.
4.1.3. Requirements for an Airport Weather Model

The requirements for airport weather modeling are driven by the needs of air traffic FM simulations. For FM purposes, both projected and real-time weather is important. Much of the planning performed by traffic management specialists is based on projected airport capacity levels, which in turn is influenced by weather forecasts. The accuracy of projected capacity is contingent upon the skill level of the forecasters, leadtime and the type of weather being predicted [25]. A synthetic forecasting method that accurately models an operational forecasting technique should generate forecasts that are correlated with synthetic observations and other forecasts of the same variable at spatially correlated sites. In other words, forecasts for ceiling at Boston and New York should be correlated to the future observations of ceiling at each city, as well as to each other. Naturally, real-time weather observations are necessary in order to simulate real-time airport capacity.

Also for FM purposes, the model should be numerical and multi-variate. Numerical, rather than analytical, in order to be implemented into a Monte Carlo simulation environment, and multi-variate in order to model the different variables that constitute airport weather.

The characteristics of airport weather that are important for flow management analysis are temporal, spatial, and cross-variable correlation. Additionally, for a weather simulation to be realistic, the model should preserve the independent, historical frequency distributions of the weather variables at the individual sites of interest. Characteristics are preserved, if a model produces synthetic data that are statistically indistinguishable from historical data for the same region and time interval [17]. The following descriptions of temporal, spatial and cross-variable correlation will explain the importance of each correlation from an FM viewpoint, calculation techniques for each correlation, and sample correlation values.

4.1.3.1. Temporal Correlation

Temporal correlation, also referred to as serial correlation, refers to the autocorrelation of weather observations in the time domain. A model must capture the relationship between time-series values of a variable in order to accurately represent the transient and dynamic aspects of weather.

Temporal autocorrelation for a weather variable can be measured by calculating the lag $s$ autocorrelation of sample observations, where $s$ represents the time interval between observations. The Pearson Product Moment formula in equation (3) can be used to calculate autocorrelation, as is illustrated below.
Assume there are $N$ discrete observations of a weather variable $V_t$, say visibility, where $t$ corresponds to the time of an observation, such that $t = 1$ to $N$. $V_{t-s}$ corresponds to an observation $s$ time units before $V_t$. In equation (3), $V_t$ corresponds to $X_t$, and $V_{t-s}$ corresponds to $Y_t$. The autocorrelation between $V_t$ and $V_{t-s}$ can be calculated by solving (3) using the observations of $V_t$ in the range $(s,N)$ for values of $X_t$, and observations of $V_t$ in the range $(1,N-s)$ for values of $Y_t$.

Figure 4.2 displays the temporal autocorrelation for ceiling at Washington National Airport, for $s=0$ to 40. The sample data consisted of actual, hourly observations for the months of January and February over a five year period [24].

![Temporal Autocorrelation for Ceiling at DCA](image)

**Figure 4.2: Temporal Autocorrelation for Ceiling at DCA**

The graph in Figure 4.2 suggests an exponential decay in ceiling autocorrelation as observations are further apart. It can be inferred that observations taken 30 hours apart or greater are approximately independent.

The accuracy of autocorrelation calculations are subject to quantization error caused by the quantification of observations. According to U.S. meteorological standards, ceiling measurements are taken to the nearest hundred feet up to 5,000 feet, and to the nearest 500 feet from 5,000 feet to 10,000 feet and to the nearest 1000 feet above 10,000 feet [23]. Although errors may occur due to quantization, the standards reduce errors that would be introduced into data due to different measurement techniques among various sites.

4.1.3.2. Spatial Correlation

Spatial correlation between two sites is a function of the separation distance between the sites. The correlation among different geographic regions will vary according to the
stability of the weather patterns and seasonal aspects. Figure 4.3 depicts a 30-year, average spatial correlation curve for cloud ceiling for sample sites in midwestern states for the month of July, compiled by the Air Force [21]. The quantification of ceiling measurements is described in the previous section.

![Figure 4.3: Ceiling Spatial Correlation - Midwestern States](image)

Within 700 miles, there is an attenuation in correlation as sites are further apart. The curve suggests that at certain distances correlation between sites is negative. This is attributable to inclement weather conditions that migrate over a region, which are followed by favorable weather conditions, and vice versa [32]. For sites over 1000 miles apart, there was some indication of positive spatial correlation.

Spatial correlation is important for FM research because it reflects the relatedness amongst neighboring airports. An FM strategy that improves the overall efficiency of the air traffic system, must account for simultaneous capacity problems at neighboring airports. A common example is the congestion and capacity problems that occur at the New York, Boston, and Washington airports. These airports often operate at demand levels close to maximum capacity. A problem at one airport usually affects the other two, simply due to flight interdependencies. Typically, however, if one airport experiences a capacity decrease due to weather, the other airports are likely to be influenced by the same weather pattern. FM strategies must be evaluated in simulation environments that model this correlation. Additionally, the model should be sufficiently flexible to be tailored to model spatial and temporal autocorrelation for different regions and seasons.
4.1.3.3. Cross-Variable Correlation

Cross-variable correlation, or simply cross-correlation, refers to the correlation between two weather variables at the same site. Cross-correlation is high amongst airport weather variables. Cross-correlation between ceiling and visibility observations at Boston Logan Airport was calculated for the months of January and February using five years of historical data [22]. Ceiling quantifications are described in Section 4.1.3.1. Visibility observations are recorded in sixteenths of a nautical mile (nm) up to 10 nm [21]. Above 10 nm, the quantifications are unspecified. Using equation (3), the sample cross-correlation coefficient based on raw ceiling and visibility data for Boston was 0.64.

Due to the fact that both ceiling and visibility influence an airport's classification, modeling cross-correlation would provide more realistic scenarios in an air traffic simulation. It would be unreasonable for a simulation to produce a synthetic ceiling observation of zero feet, while simultaneously generating a favorable observation for visibility.

4.1.4. Alternative Weather Modeling Techniques

4.1.4.1. Discrete Markov Process

In connection with air traffic delay research, Peterson [16] proposed modeling airport capacity with regard to weather as a discrete, first-order Markov chain. The states of the Markov chain are defined as different weather conditions and runway configurations at an airport. The state transition probabilities can be derived using historical records of weather observations and airport configurations. The Peterson model was designed primarily to support delay studies for two weather-independent hub airports. The primary limitations of the Markov chain model for FM purposes, are that it does not easily extend to produce spatially and cross-correlated weather data for multiple airports. In addition, the model does not provide weather forecasts. Furthermore, by generalizing weather conditions into states, the fidelity of the temporal aspects of weather may be insufficient.

4.1.4.2. Historical Data

There are reasons for developing a numerical model for aviation weather rather than using direct historical data [17]. With historical data, there is a risk of choosing a sample set of weather with conditions that are either too mild or too severe, adversely affecting FM analysis. Furthermore, a sample set selected as typical, may in fact contain biases, or be
uncharacteristic of the future. It is conceivable that the typical data set will become the only weather data used in an FM simulation. This may result in adjusting an FM algorithm to manage congestion for the sample data only. Furthermore, a numerical model may reveal future realizations that have yet to occur. Finally, synthetic weather is easier to manage than historic weather data files, which often require manipulation, filtering and large memory storage.

However, historical data sets play a key role in the development of a numerical model. Real data provide the basis to develop frequency distributions and correlations needed in the mathematical model. Furthermore, validating a model requires extensive use of past data to test the accuracy and realism of a model once implemented.

4.2. Sawtooth Wave Model

The SWM was developed by the USAF to support wargaming efforts in different regions of the world. The SWM was developed as a submodel for wargaming simulations to generate cloud cover ceiling and visibility observations and forecasts with spatial, temporal, and cross-correlation.

The basic function of the SWM consists of superimposing randomly generated, multidimensional sawtooth waves within a time and space coordinate system [17]. At sites of interest within the coordinate system, the heights of the random waves are summed together. The sums will be normally distributed if sufficient waves are used in the model, i.e., more than 12, due to the Central Limit Theorem [18]. The sums are transformed into standard normally distributed values, i.e., with a mean of zero and variance of unity, i.e., N(0,1).

The N(0,1) values are referred to as Equivalent Normal Deviates, or ENDs [18]. The ENDs are converted into observations and forecasts of weather variables, through a process known as inverse transnormalization. Time-series data are generated by adjusting the sawtooth waves in the time dimension and repeating the process.

The SWM will be described in four parts. The first part will describe the geometry of multidimensional sawtooth waves. The second part presents the correlation curve that relates correlation in the space and time domains to the wavelengths of the sawtooth waves. The third and fourth sections cover forecast and cross-correlation modeling.

Following a description of the model, is a discussion on the process of converting ENDs into raw weather values through inverse transnormalization. Additionally, the assumptions of the SWM will be presented.
4.2.1. Sawtooth Wave Geometry

A sawtooth wave can be described in terms of its direction, wavelength, phase shift and amplitude. A one-dimensional sawtooth wave is depicted in Figure 4.4.

![Figure 4.4: 1-D Sawtooth Wave](image)

where \( u \) is the phase shift at the origin and \( L \) is the wavelength. The wave in Figure 4.4 has an amplitude or height of 1, as will all waves in the SWM. The dimensionality of the wave is determined by the number of dimensions required to describe the direction of the wave. Thus, amplitude is not an additional dimension. The direction of the sawtooth wave in Figure 4.4 is positive. A 2-D sawtooth wave exists on a plane. A physical representation of a 2-D wave is the surface of a woodfile. A 3-D wave becomes more difficult to visualize. One cycle of a 3-D wave can be represented by a sheet of plywood, dark on one side and gradually fading in color through the plywood to white on the other side. The intensity of color corresponds to amplitude. A 3-D wave would consist of multiple plywood sheets stacked together [17].

Figure 4.5 is a standing 2-D wave. Assume the phase shift and direction of the wave are selected at random from uniform distributions over the intervals

\[
0 \leq u < L
\]  (4)

and

\[
0 \leq \alpha < 2\pi \]

\[
\left( \beta = \alpha - \frac{\pi}{2} \right)
\]  (5)

Under these conditions, the height of the wave at any point in the 2-D coordinate system is a uniformly distributed random variable, with a mean of 1/2 and a variance of 1/12. Figure 4.5 displays the parameters required in order to determine the height of the wave at a random point \( p \) [32].
Figure 4.5: 2-D Wave with Point p

If the amplitude of the sawtooth wave is unity, the height of the wave at point $p$, $h_p$, is equivalent to the fractional position of $p$ within the cycle of the sawtooth wave containing $p$. To calculate the fractional position of $p$ within its respective cycle, it is necessary to determine the normal distance between $p$ and a line passing through the origin, parallel to the wave, represented by $d_p$ and calculated

$$d_p = x_p \cos \alpha + y_p \cos \beta$$ (6)

where $(x_p, y_p)$ are the coordinates of $p$, and $\cos \alpha$ and $\cos \beta$ are referred to as the direction cosines for $p$.

Accounting for the phase shift $u$, it is possible to calculate the position of $p$ within its respective wave cycle, and hence the height of the wave at $p$, by the equation

$$h_p = \frac{d_p + u}{l} - \left\lfloor \frac{d_p + u}{l} \right\rfloor$$ (7)
where the notation $\lfloor \rfloor$ implies rounding a non-integer value to the nearest lower integer. $d_p$ will be negative, if the direction of the sawtooth wave opposes the direction of a vector from the origin to point $p$.

Assume, in Figure 4.5, that $N$ sawtooth waves are generated, of the same wavelength, with random directions and random phase shifts, all with an amplitude of unity. At $p$, there will be $N$ uniformly distributed waveheights. If $N$ is sufficiently large, i.e. $N \geq 12$, the sum of the waveheights at $p$, $w_p$, will approximate a normally distributed random variable because of the Central Limit Theorem [18]. $w_p$ will have a mean of $\bar{w}_p$ and standard deviation of $s_{w_p}$, where,

$$\bar{w}_p = \frac{N}{2}$$

$$s_{w_p} = \sqrt{\frac{N}{12}}$$

It is possible to transform $w_p$ into an N(0,1) END by the equation

$$z = \frac{w_p - \bar{w}_p}{s_{w_p}}$$

The SWM simulates spatially correlated, time-series data for a single N(0,1) variable on the surface of the earth by extending the concept of multiple superimposed waves into four dimensions, for space and time. In essence, the model creates a grid of spatially correlated random variables in 3-D, for any moment in time, $t$. At $t+1$, the grid changes in such a way that successive sums of waveheights at the same location will possess temporal autocorrelation.

The center of the earth is the origin for the space dimensions as shown in Figure 4.5 [20].
The $y$ axis passes through the equator and the prime meridian (in Africa), the $x$ axis passes through the equator and $90^\circ$E (in the Indian Ocean), while the $z$ axis passes through the North Pole. It is possible to determine the coordinates of a random point $p$, on the surface of the earth, with a latitude of $lat_p$ and a longitude of $lon_p$, using the equations

$$
x_p = E \cdot \cos(lat_p) \cdot \sin(lon_p)
$$
$$
y_p = E \cdot \cos(lat_p) \cdot \cos(lon_p)
$$
$$
z_p = E \cdot \sin(lat_p)
$$

where $E$ is the radius of the earth. One can reasonably assume a spherical earth with $E = 3438.26$ nm, without affecting the results of the model.

To illustrate the geometry of a single sawtooth wave in the spatial dimensions, assume a standing 3-D wave is created with a randomly generated direction and phase shift. Again, all waves in the model have a height of unity. The height of a wave at any point on the surface of the earth is a uniformly distributed random variable with a mean of $1/2$ and variance $1/12$, as in the 2-D model. The distance $d_p$, in the 3-D model, represents the length of a line connecting $p$ to a two dimensional plane, normal to the line and passing through the center of the earth [32]. Using the coordinates of $p$, one can calculate $d_p$ by extending equation (6) to incorporate the direction cosine of $p$ in the $z$ direction, i.e.,

$$
d_p = x_p \cdot \cos(\alpha) + y_p \cdot \cos(\beta) + z_p \cdot \cos(\gamma)
$$

The equation to calculate the waveheight at $p$, for the 2-D example, i.e., equation (7), is true for all dimensions.
The fourth dimension is time, or \( t \). The principles of the model are the same for any dimension, however, it is necessary to standardize the units of space and time into units of wavelength, in order to calculate \( h_p \) for each wave at each point of interest [32]. \( \Lambda_s \) and \( \Lambda_t \) represent the wavelengths for the sawtooth waves in space and time units respectively. Determining appropriate values for \( \Lambda_s \) and \( \Lambda_t \) according to desired levels of spatial and temporal correlation will be discussed in the next section.

Assume \( \Lambda_s \) and \( \Lambda_t \) are given. \( d_p^\Lambda \) is the normal distance, in wavelengths, from point \( p \) with coordinates \((x_p, y_p, z_p, t_p)\), to a 3-D hyperplane which passes through the origin. The equation for \( d_p^\Lambda \) is

\[
d_p^\Lambda = \left( \frac{x_p}{\Lambda_s} \right) \cdot \cos(\alpha) + \left( \frac{y_p}{\Lambda_s} \right) \cdot \cos(\beta) + \left( \frac{z_p}{\Lambda_s} \right) \cdot \cos(\gamma) + \left( \frac{t_p}{\Lambda_t} \right) \cdot \cos(\omega) \tag{12}
\]

\( \cos(\omega) \) is the direction cosine for \( p \) in the time direction. The waveheight at \( p \) for a single wave is

\[
h_p = \left( d_p^\Lambda + u^\Lambda \right) - \left[ d_p^\Lambda + u^\Lambda \right] \tag{13}
\]

where \( u^\Lambda \) is the phase shift in units of wavelength such that,

\[
0 \leq u^\Lambda < 1 \tag{14}
\]

\( \text{N(0,1)} \) variables are created at sites of interest on the surface of the earth by generating multiple 4-D waves with random phase shifts and directions, summing the waveheights at the sites of interest and transforming the sums into \( \text{N(0,1)} \) random variables. A time-series of realizations at each site is achieved by incrementally increasing the value of \( t_p \) in equation (12) and performing successive calculations of \( h_p \) for each wave at each site. The time-series ENDS correspond to observations of a weather variable and can be transformed into raw data using the process of inverse transnormalization.

4.2.2. Correlation Curves

Boehm [19] has derived formulas for correlation curves that describe the theoretical relationship between the separation distance between two points in a SWM and correlation in the generated grid of random variables. Boehm calculates correlation as an expected value by
integrating over all possible directions of a wave in the SWM given the probability density of wave directions.

The correlation formula is

$$ r_n = 1 - \frac{12 \cdot \delta}{(n-1) \cdot \beta\left(\frac{1}{2}, \frac{n-1}{2}\right)} + \frac{6 \cdot \delta^2}{n} $$

(15)

where $r_n$ is correlation in $n$ dimensions and $\delta$ is the separation distance between two points in units of wavelength [19]. The beta function $\beta$ is defined as

$$ \beta\left(\frac{1}{2}, \frac{n}{2}\right) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \cdot \Gamma\left(\frac{n-1}{2}\right)} $$

(16)

It is assumed that all waves in the SWM have the same wavelength. Figure 4.7 plots equation (15) for $n=4$ and $n=5$.

![Graph](image)

**Figure 4.7:** 4-D and 5-D Correlation Curves

The application of equation (15) and the curves in Figure 4.8 is as follows. The objective is to calibrate the curve in Figure 4.7 along the horizontal axis, to fit a desired curve for either spatial or temporal correlation, e.g., the curves in Figures 4.2 and 4.3. In order to model both space and time, it is necessary to calibrate the curve in both units of measurement.

Generically, in order to calculate a wavelength $\Lambda$ in any unit of measurement that corresponds to some desired correlation in $n$ dimensions, it is necessary to have a reference
distance that corresponds to a reference correlation. (Distance refers to either distance in space units or time units.)

Assume \( m_n \) represents the distance in space units of measurements (e.g., nm) in \( n \) dimensions, between two points for which there is a desired correlation of \( r_n \). Using the inverse of equation (15), it is possible to calculate the separation distance in units of wavelengths, i.e., \( \delta \), that corresponds to \( m_n \). Subsequently, the wavelength for the SWM in space units of measurement, can then be determined using the equation

\[
\Lambda = \frac{m_n}{\delta}
\]  

(17)

In the 4-D model for a single variable in space and time, it is necessary to calibrate the curve in the space and time dimensions. Calibrating in the time dimension will produce a value for \( \Lambda_t \) to be applied in equation (12). Calibrating in the space dimension will produce a value for \( \Lambda_s \) for the same equation.

4.2.3. Forecast Modeling

Correlation between forecasts and observations are achieved by exploiting the omnidirectional correlation grid created by the SWM. The correlation grid created by an \( n \) dimensional SWM is omnidirectional, in that any two points in the \( n \) dimensional coordinate system exhibit correlation based upon the sawtooth wavelength. Therefore, points above and below the surface of the earth exhibit the same correlation properties as points on the surface of the earth.

Forecasts can be generated by adjusting the value of the earth’s radius, \( E \), by an amount that corresponds to a desired forecast skill level, i.e. a forecast-observation correlation [32]. The adjustment of \( E \) is determined by calculating the distance in units of wavelength that corresponds to a desired correlation. This requires using the inverse of equation (15). The distance in wavelengths is transformed into space units of measurement using equation (17) and subtracted from, or added to, the earth’s radius, creating fictitious sites for generating forecasts. Similar to the process of generating ENDS for observations, waveheights are summed at the fictitious forecast sites, above or below the earth, using adjusted values of \( E \), to generate “forecast” ENDS.

Figure 4.9 depicts observation and forecast sites for two example sites (A and B), using points below the surface of the earth for fictitious forecast sites. In Figure 4.8, \( O(t) \) corresponds to observations and \( F(t+t') \) corresponds to forecasts \( t' \) time units in the future of the \( O(t) \)
observations. The arrows indicate the difference in distance required to obtain a specific correlation between observations at time $t$, and forecasts for $t+t'$.

![Diagram showing observation and forecast sites](image)

**Figure 4.8: Observation and Forecast Sites**

This forecasting technique simultaneously achieves correlation amongst the forecasts, as well as forecast-observation correlation. As seen in Figure 4.9, forecasts will be more closely correlated than the actual observations, since the fictitious forecast sites are closer together, which, according to the USAF, is representative of actual forecasting over a geographical region.

At a single site, research suggests that the accuracy of forecasts decreases as a function of leadtime [25]. In the SWM forecast model, longer leadtimes would correspond to smaller values for $E$. This forecast model also allows modeling different forecasting skills at different sites by adjusting values of $E$ appropriately. The model is referred to as the “breathing earth” model [32].

### 4.2.4. Cross-Variable Correlation

There are two techniques for extending the SWM to model cross-correlation between two variables. The first involves the breathing earth model. Assume $X_1$ is a variable being modeled on the surface of the earth and $X_2$ is a second variable to be modeled with cross-correlation to $X_1$. It is possible to select points sufficiently above the surface of the earth, to generate ENDS for $X_2$ that are cross-correlated to the ENDS for $X_1$ on the surface of the earth, and spatially correlated amongst each other. A limitation of this technique involves generating forecast ENDS for $X_2$. Using smaller values of $E$ to forecast $X_2$ may result in forecasts that are perfectly correlated to the observations of $X_1$. Using larger values of $E$ to generate forecasts for $X_2$ will result in forecasts for the two variables that are less cross-correlated than the cross-correlation for the observations.

The second technique for achieving cross-correlation involves adding a new dimension to the SWM [32]. The wavelengths are recalculated in the higher dimension using the inverse
of equation (15). Equation (12) for calculating the heights of the individual waves at sites of interests, is modified to include the additional dimension as shown in equation (18).

\[
d_p^A = \left( \frac{x_p}{\Lambda_s} \right) \cdot \cos(\alpha) + \left( \frac{y_p}{\Lambda_s} \right) \cdot \cos(\beta) \\
+ \left( \frac{z_p}{\Lambda_s} \right) \cdot \cos(\gamma) + \left( \frac{f_p}{\Lambda_l} \right) \cdot \cos(\delta) + k \cdot \cos(\theta)
\] (18)

\(\theta\) is the direction angle for a wave in the new dimension. \(k\) is a factor in units of wavelength that corresponds to a desired cross-correlation between two variables. Returning to the \(X_1 - X_2\) example, the ENDs for \(X_1\) are calculated by setting \(k = 0\) for all waveheight calculations. \((X_1\) is perfectly cross-correlated to itself). In calculating the waveheights for \(X_2\), \(k\) is set to the separation distance in units of wavelength that corresponds to the desired cross-correlation. Conceptually, the model creates a separate earth for simulating \(X_2\) exactly \(k\) wavelengths away from the earth where \(X_1\) is being simulated.

Forecasting in the 5-D model for both variables is an extension of the 4-D model. The value for \(E\) is decreased when calculating the coordinates of fictitious forecasting sites in the space dimensions. For forecasts of \(X_1\), \(k = 0\). For forecasts of \(X_2\), \(k\) is set to the separation distance in units of wavelength. Cross-correlation for more than two variables can be simulated by further increasing the number of dimensions in the model.

4.2.5. Inverse Transnormalization

Most meteorological variables are not normally distributed. In order to apply the SWM to weather variables, it is necessary to transform the N(0,1) variables into simulated weather observations, a process referred to by the USAF as in inverse transnormalization [17].

Inverse transnormalization can be illustrated through an example. ENDs are transformed into weather values according to the cumulative distribution functions (CDFs) of the weather variable being modeled. Assume \(\tilde{C}_i\) is the END for ceiling at a particular site, \(C_i\) represents the raw synthetic ceiling value to be generated, and \(\bar{c}_i\) and \(c_i\) are realizations of \(\tilde{C}_i\) and \(C_i\), respectively. Values of \(\tilde{C}_i\) are transformed into raw values of ceiling, in such a manner that

\[
P(C_i < c_i) = P(\tilde{C}_i < \bar{c}_i)
\] (19)

For example, assume the SWM generates a realization for \(\tilde{C}_i\) of \(\bar{c}_i = -0.253\). From the standard normal lookup table, \(P(\tilde{C}_i < -0.253) = 0.4\). The inverse transnormalization scheme...
converts the END $\bar{c}_i$ into a realization for ceiling, based on the CDF for the site being modeled. Figure 4.9 contains a hypothetical CDF for a fictitious Site A.

![Figure 4.9: Ceiling CDF for Site A](image)

From Figure 4.9, the ceiling realization that matches an END of -0.253 is 7,500 feet.

### 4.2.6. Basic Model Assumptions

The fundamental assumption of the SWM is that spatial and temporal correlation can be modeled using the quadratic decay function in equation (15). Additional assumptions include, isotropic and homogeneous spatial correlation [18]. Isotropic correlation implies that correlation between two sites is independent of the orientation of the sites to each other. Homogeneous correlation implies that correlation is the same everywhere. These assumptions do not present a problem if the model is applied to a limited area. If the model is applied to a large area where correlation is not homogeneous the sawtooth wavelength should be chosen that represents the correlation of a subarea of greatest interest [18]. Moreover, there are more sophisticated methods of calibrating the correlation formula using multiple reference points, that modifies the form of the correlation curves.

The model also assumes that all variables display the same decay in temporal autocorrelation. USAF tests suggest that for ceiling and visibility, this assumption is largely valid [23]. This conclusion was confirmed in correlation statistics calculated for airports at Boston, LaGuardia and Washington, the results of which will be presented in the next section.
The SWM presented in this section is the basic model. Air Force weather specialists have verified and validated the model in multiple simulation environments [17] [18]. Furthermore, extensions have been developed to model non-homogeneous spatial correlation and variations in the correlation curves presented in Figure 4.7 [32].

4.3. Implementation

In order to evaluate the SWM’s characteristics, the basic SWM was implemented in a computer simulation. Synthetic hourly weather observations for ceiling and visibility were generated for three sites: Boston (BOS), New York (LGA) and Washington D.C. (DCA). Additionally, 3- and 6-hour ceiling forecasts were generated every hour at Boston. The objective was to test the ability of the model to simulate spatial correlation, temporal autocorrelation, and cross-correlation for observations and forecasts. A secondary motivation was to test the validity of the SWM assumptions.

4.3.1. Simulation Procedure

Four simulation runs were performed using 5-D sawtooth waves. (It was necessary to use 5-D waves in order to model two variables (ceiling and visibility) at each site with cross-correlation.) Each run generated 1000 data points for each site and variable of interest. Figure 4.10 displays the process that was followed for each run and components of each step.
Figure 4.10: Simulation Run Process

The simulation runs were conducted using an implementation of the SWM in a computer routine coded in the C language. The input cumulative distribution functions (CDFs) for ceiling and visibility at all sites were also calculated using a C program. All correlation calculations, transformations and inverse transnormalizations were performed using the Microsoft Excel spreadsheet application on a Macintosh platform. The units of time are hours, and the units of distance are nautical miles (nm) for all calculations.

The simulation run process will be described below.

4.3.1.1. Step 1: Setting the Input Parameters

The input parameters defined the weather characteristics and sites that were modeled in the simulation runs. The sites and their geographical coordinates were:

- **BOS** 42° 20' (N) 71° 05' (E)
- **LGA** 40° 43' (N) 74° 01' (E)
- **DCA** 38° 55' (N) 77° 00' (E)
These sites were selected due to the fact that they are major airports within the same geographical region (the East Coast). In addition, it would be possible to compare synthetic observations generated by the simulation runs to actual observations for the same sites using five years of hourly weather data for the months of January and February, procured from the National Oceanic and Atmospheric Administration (NOAA) [24]. The NOAA data consisted of 1416 consecutive observations at each site for each variable by year; except for 1988, a leap year, for which there were 1440 consecutive observations. Correlation statistics involving the NOAA data were calculated by averaging the correlation values for each year.

The NOAA quantifications of ceiling and visibility observations are as follows. Ceiling measurements are taken to the nearest hundred feet up to 5,000 feet, and to the nearest 500 feet from 5,000 feet to 10,000 feet and to the nearest 1000 feet above 10,000 feet. Visibility observations are recorded in sixteens of a nautical mile (nm) up to 10 nm [23]. Above 10 nm, the quantifications are unspecified.

* Reference Spatial Distance and Correlation

The SWM in the simulation runs required a reference separation distance and spatial correlation, as well as a reference separation time and temporal correlation. These values calibrated the 5-D correlation curve in Figure 4.7 in the space and time dimensions. In order to simulate real weather characteristics at the sites of interest, all reference values were calculated using the NOAA weather data.

Lag $s$ spatial correlations were calculated for each site pair, where $s$ represents the difference in hours between observations. For example, lag 2 spatial correlation between BOS and LGA for a specific weather variable, refers to the correlation between observations at Boston at time $t$ and observations at LaGuardia at time $t-s$. Lag 0 represents the correlation between sites for same-hour observations. The assumption was that because winter weather patterns generally migrate from west to east, low ceiling or visibility values would usually be observed first at DCA, then at LGA, and finally at BOS. Lag $s$ spatial correlations between the sites should reflect the average weather migration trends.

The assumption was confirmed as valid, as displayed by the dashed lines in Figures 4.11 through 4.16 (in Section 4.3.2.1). The curves represent the observed lag $s$ spatial correlations for both ceiling and visibility at the sites of interest, for $s=-1$ to 10. The peak lag $s$ spatial correlation values are provided in Table 4.1 below, along with the separation distance between the site pairs. The site pairs are ordered in such a manner that all lag values are positive. (Note: BOS-LGA lag $s$ spatial correlation is equivalent to LGA-BOS lag -$s$ spatial correlation.)
Table 4.1: Observed Lag $s$ Spatial Correlation and Site Separation Distances

<table>
<thead>
<tr>
<th></th>
<th>Ceiling</th>
<th>Visibility</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOS-LGA</td>
<td>0.67 (lag 2)</td>
<td>0.61 (lag 4)</td>
<td>200 nm</td>
</tr>
<tr>
<td>LGA-DCA</td>
<td>0.56 (lag 4)</td>
<td>0.64 (lag 4)</td>
<td>210 nm</td>
</tr>
<tr>
<td>BOS-DCA</td>
<td>0.42 (lag 7)</td>
<td>0.44 (lag 8)</td>
<td>405 nm</td>
</tr>
</tbody>
</table>

Lag $s$ correlations were calculated (rather than same-hour correlations, exclusively) in order to enhance the accuracy of the input correlation parameters.

The BOS-LGA and LGA-DCA results in Table 4.1 suggest that spatial correlation for the sites of interest is approximately homogeneous, since the site pairs are separated by similar distances and exhibit similar correlations. Table 4.1 also indicates that spatial correlation for ceiling and visibility is roughly equivalent for all site pairs.

In order to approximate the BOS-LGA and LGA-DCA statistics in Table 4.1, the reference separation distance for the simulation runs was selected as 200 nm with a reference correlation of 0.60. Incorporating the lag component of spatial correlation will be discussed in the final step of the simulation.

- **Reference Separation Time and Temporal Autocorrelation**

Using the NOAA data and lag $s$ autocorrelation as described in Section 4.1.3.1, temporal autocorrelation curves were derived for ceiling and visibility at each site. The variability between the curves for different sites was relatively small. The dashed lines in Figures 4.17 and 4.18 depict the average temporal autocorrelation decay functions for $s=0$ to 40, for both ceiling and visibility.

Using the ceiling curve in Figure 4.17, a reference separation time was selected as 24 hours with a reference autocorrelation of 0.14. 24 hours was selected in order to attempt to calibrate the correlation curve in Figure 4.7 for temporal autocorrelation accuracy within a single day. The ceiling curve was arbitrarily selected over the visibility curve, although the differences between the curves are negligible.

- **Cross-correlation**

Ceiling and visibility cross-correlation values were calculated at the sites of interest using the NOAA data and the Pearson Product Moment formula. The results were: 0.64 at BOS, 0.61 at LGA and 0.48 at DCA, which uncovered an inaccuracy in the SWM assumption that
cross-correlation is equivalent at all sites. The input parameter for cross-correlation was selected as 0.60 in order to approximate cross-correlation at BOS and LGA.

- **Forecast Skill Levels**

The final input parameters to the SWM were the forecast skill levels for the 3- and 6-hour ceiling forecasts at Boston. According to Shumsky [25], forecasting skill varies significantly depending upon the type of weather being forecast and leadtimes. Rather than attempting to simulate actual forecast skills, two arbitrary values were selected as 0.6 and 0.4. These values reflect the decrease in forecast accuracy with longer leadtime. The SWM assumes that forecast skill and forecast-observation correlation is equivalent.

- **Cumulative Distribution Functions (CDFs)**

The inverse transnormalization process requires as input, CDFs for the variables being simulated. Using January and February observations in the NOAA data [24], historical CDFs were derived for visibility and ceiling at the three sites of interest. The columns in Figures 4.19 through 4.24 represent the observed CDFs. In the figures, the absence of a column for a quantile, indicate that no observations were made within the range of that quantile. The observed CDFs for visibility at all sites were very similar. Concerning ceiling however, the CDFs revealed more low cloud observations for DCA than for LGA or BOS. This regional variation does not affect the model.

#### 4.3.1.2. Step 2: Calculating the SWM Parameters

The six parameters for the SWM in the simulation runs were:

- \( \lambda_s \) - length of sawtooth waves in space units of measurement: (nm)
- \( \lambda_t \) - length of sawtooth waves in time units of measurement (hours)
- \( k \) - ceiling and visibility cross-correlation factor
- \( E(3) \) - Earth radius for 3-hour forecasts
- \( E(6) \) - Earth radius for 6-hour forecasts
- \( x_i, y_i, z_i \) - Earth-centered origin coordinates for each site \( i \), i.e., BOS, LGA and DCA

The wavelengths were a function of the input parameter values of spatial and temporal correlation as described by the reference values. The correlation formula (15) for a 5-D model is
$r_5 = 1 - 2.25 \cdot \delta + 1.2 \cdot \delta^2 \quad \forall \delta < 1 \quad (20)$

where $r_5$ is correlation in 5-D with regard to $\delta$, the separation distance between two points in time and space, in units of wavelength. The inverse of (20), i.e., $\delta$ as a function of correlation, is

$$\delta = 0.9375 - \sqrt{0.833 \cdot r_5 + 0.0459} \quad (21)$$

With (21) it is possible to determine the lengths of the sawtooth waves in time and space.

- Calculating $\Lambda_s$

$\Lambda_s$, the sawtooth wavelength in the space dimensions, was calculated based on the reference distance and correlation, i.e., 200 nm and 0.60. The sawtooth wavelength was calculated by finding the units of wavelength, $\delta$, that corresponded to a separation distance of 200 nm. Using (21) with $r_5=0.6$, a value for $\delta$ was determined to be 0.198 wavelengths. It was then possible to calculate the value for the wavelength in the space dimension as

$$\Lambda_s = \frac{200 \, \text{nm}}{0.198 \, \text{wavelengths}} = 1006.54 \, \text{nm}$$

- Calculating $\Lambda_t$

$\Lambda_t$, the sawtooth wavelength in the time dimension, was calculated based on the reference separation time and temporal autocorrelation, i.e., 24 hours and 0.14. Using equation (21) with $r_5=0.14$, $\delta=0.534$. The wavelength in the time dimension was therefore,

$$\Lambda_t = \frac{24 \, \text{hours}}{0.534 \, \text{wavelengths}} = 44.91 \, \text{hours}$$

- Generating 14 Random Sawtooth Waves

14 waves were generated for each run. The number of waves was chosen for reasons of computational simplicity. Additionally, the number is sufficiently large in order for the sum of the waveheights at sites of interest to approximate normally distributed random variables.

A random wave is defined by a random phase shift and a random direction. All waves had an amplitude of unity. Since the equation to be used for calculating individual waveheights at sites of interest was (13), the phase was determined for each wave using a Monte Carlo method, where a random number was generated from a uniformly distributed random variable over the interval (0,1).
Generating a random direction for a wave in multidimensions is non-trivial. The direction of a wave is defined by the angles between a wave’s direction vector and each axis. The simulation runs utilized a Monte Carlo technique involving a unit hyper-cube, which generates the direction cosines for each wave [22].

* Calculating $k$

The cross-correlation factor, $k$ as explained in Section 4.2.4, represents the separation distance in units of wavelength that corresponds to a desired cross-correlation between two variables. For the simulation, the desired cross-correlation was 0.6. Using equation (21) with $r = 0.6$, a value for $k$ was calculated as 0.198.

* Calculating $E(3)$ and $E(6)$ for Forecast Sites

In order to simulate ceiling forecasts at Boston, the simulation runs were designed to use the breathing earth technique explained in Section 4.2.3. The breathing earth technique establishes forecast sites below the surface of the earth, that correspond to desired forecast-observation correlation levels. The sites for the forecasts were calculated according to desired forecast-observation correlation levels for 3- and 6-hour forecasts of 0.6 and 0.4.

The lat/long coordinates for the forecast sites are equivalent to the lat/long coordinates of the actual site of interest, i.e., BOS. The forecast-observation correlation levels were used to determine values for $E(3)$ and $E(6)$. First, we calculated the units of wavelength that corresponded to each skill level, i.e., $\delta_3$ and $\delta_6$, using equation (21). The results were

\[
\delta_3 = 0.198 \\
\delta_6 = 0.322
\]

Subsequently, using the parameter $A_x$, it was possible to find values for $E(3)$ and $E(6)$ by subtracting the distances in nautical miles that correspond to $\delta_3$ and $\delta_6$ from the earth’s radius, i.e., 3438.26 nm. The results were

\[
E(3) = 3438.26 - (0.198 \cdot 1006.54) = 3238.96 \text{ nm} \\
E(6) = 3438.26 - (0.322 \cdot 1006.54) = 3114.15 \text{ nm}
\]

As expected the forecast site for the 6-hour forecasts was further below the surface of the earth corresponding to the lower forecast skill.
• **Earth-Centered Coordinates**

The final parameters for the SWM were the earth-centered coordinates for each site, including the forecast sites. The lat/long coordinates were converted to x,y,z coordinates using the equations in (25). For conversion purposes, eastern longitudinal and southern latitudinal values are negative.

### 4.3.1.3. Step 3: Superimposing Sawtooth Waves - Generating ENDS

Superimposing the sawtooth waves refers to calculating and summing the waveheights at the sites of interest. It was possible to simplify equations (13) and (18), due to the fact that the location coordinates were fixed and only the time variable was modified during the simulation runs. All time-independent terms in (13) and (18) were calculated once for each run and stored as constants. In recalculating the waveheights for each observation or forecast site, it was only required to determine the change for the term in the time dimension.

This step was performed by a simple computer program coded in C. Waveheights were calculated, using (13) and (18) at each site with \( k=0 \), to represent ceiling observations, and \( k=0.198 \) for visibility observations. The waveheights were also calculated using the BOS coordinates with the values for \( E(3) \) and \( E(6) \), and \( k=0 \) to represent the 3- and 6-hour forecasts for ceiling.

1008 time-series data points for each variable at each site were generated corresponding to hourly observations and forecasts. The goal was to simulate 1000 data points at each site. The additional 8 values were generated in order model the lag spatial correlation among the sites. This will be explained further in the final step of the simulation process.

Since the number of waves, i.e., \( N \), was 14, the theoretical mean \( \bar{w} \) and standard deviation \( s_w \) for the wave heights at each site were

\[
\bar{w} = \frac{14}{2} = 7.0
\]

\[
s_w = \sqrt{\frac{14}{12}} = 1.08
\]

Based on these values, all waveheights were transformed into Equivalent Normal Deviates (ENDs) using the standard normal transformation method in equation (9), i.e.,

\[
z_p = \frac{w_p - 7.0}{1.08}
\]

142
where $z_p$ is the END that corresponds to waveheight $w_p$, at site $p$.

### 4.3.1.4. Step 4: Inverse Transnormalization of ENDS into Raw Data

The final step of each simulation run was transforming the ENDS into raw synthetic observations and forecasts using the process of inverse transnormalization described in Section 4.2.5. The first step in inverse transnormalizing an END, is determining the normal cumulative probability from $-\infty$ to the END value according to the standard normal distribution, i.e., $\text{N}(0,1)$. All standard normal, cumulative probabilities were calculated using the function `NORMSDIST()` embedded in the Microsoft Excel software application.

Subsequently, referencing the input CDFs for each variable, the cumulative probabilities were converted in raw weather observations. Because the CDFs were computed based on discrete ceiling thresholds, it was necessary to use linear approximations in between the thresholds, in order to complete the continuous functions. Table 4.2 contains an example of a single realization of a synthetic ceiling observation for BOS. Each entry in Table 4.1 corresponds to an intermediate calculation in the simulation process. $w_{\text{BOS}}$ is the sum of the waveheights, $z_{\text{BOS}}$ is the END value, $P(Z < z_{\text{BOS}})$ is the cumulative normal probability, and $c_{\text{BOS}}$ is the raw ceiling observation.

<table>
<thead>
<tr>
<th>$w_{\text{BOS}}$</th>
<th>$z_{\text{BOS}}$</th>
<th>$P(Z &lt; z_{\text{BOS}})$</th>
<th>$c_{\text{BOS}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.84</td>
<td>-1.074</td>
<td>0.141</td>
<td>1916 ft</td>
</tr>
</tbody>
</table>

Lag $s$ spatial correlation was modeled in the simulations by offsetting the raw data values at the LGA and DCA sites with regard to BOS, according to the visibility lag values of BOS-LGA: 4, and BOS-DCA: 8. In other words, the $n$th realizations for ceiling and visibility became the $(n-4)$th observations at LGA and the $(n-8)$th observation at DCAs.

In order to offset the data, it was necessary to discard the first 8 observations at DCA, the first 4 observations at LGA, the last 8 observations at BOS, and the last 4 observations at DCA. As a result, each simulation contained 1000 data points. The lag $s$ values for visibility were selected over ceiling lag $s$ values due to the fact that the visibility lag for BOS-DCA was exactly twice the lag for BOS-LGA, which facilitated offsetting the data.
4.3.2. Simulation Results

The synthetic raw data were analyzed in order to evaluate the effectiveness of the model and the validity of the model assumptions. The correlation statistics and CDFs for the synthetic data were calculated with the same methods that were used for the actual NOAA weather data statistics. The statistics presented in this section represent the average values from the four simulation runs.

4.3.2.1. Spatial Correlation

Figures 4.11 through 4.16 display the lag s spatial correlation for the raw values of synthetic ceiling and visibility (dashed lines) along the with the observed spatial correlation from the NOAA data (solid lines). Concerning the spatial correlation input parameters, the model was calibrated for 0.6 spatial correlation at 200 nautical miles, offset by the appropriate lag values. The performance of the model, with reference to 200 nm, can be ascertained from the ceiling and visibility curves for BOS-LGA and LGA-DCA, due to the fact that both airport pairs are separated by approximately 200 nm. At lag 4, the average correlation at 200 nm is 0.596, which is reasonably close to the desired value of 0.6. The assumption that spatial correlation is equivalent for all variables at all sites is demonstrated as somewhat inaccurate, in the difference between the observed and synthetic curves for BOS-LGA in Figure 4.11.

The BOS-DCA curves suggest that the correlation formula (21) does not exactly model the decay in spatial correlation for the three sites of interest, since both the synthetic ceiling and visibility spatial correlations are below the observed correlations. It is possible to calculate the theoretical spatial correlation for BOS and DCA by calculating the separation distance in units of wavelength and then applying equation (20).

\[ \delta_{BOS-DCA} = \frac{405}{1006.54} = 0.40 \]

\[ r_{BOS-DCA} = 0.29 \]

The theoretical value is confirmed in the simulation runs which produced on average between synthetic ceiling (0.25) and visibility (0.34), an equivalent correlation. However, observed correlation values derived from the NOAA data are 0.42 and 0.44 for ceiling and visibility. This demonstrates that spatial correlation decay produced by the basic model inaccurately models spatial correlation for the sites of interest.
Chapter 4: Airport Weather and Capacity Modeling

Figure 4.11, 4.12, and 4.13: Lag's Spatial Correlation for Ceiling
Chapter 4: Airport Weather and Capacity Modeling

Figure 4.14, 4.15, and 4.16: Lag Spatial Correlation for Visibility
4.3.2.2. Temporal Autocorrelation

Figures 4.17 and 4.18 depict the observed and synthetic temporal autocorrelation curves for ceiling and visibility respectively. The synthetic curves represent the average temporal autocorrelations over the four simulation runs for the raw synthetic data at the three sites. Similarly, the observed curves represent the average temporal autocorrelations over the five years of NOAA data for January and February at the three sites. For both the observed and synthetic data, the curves for the individual sites did not deviate significantly from the average curves.

Figure 4.17 and 4.18 suggest that the SWM correlation curve is a reasonable approximation for temporal autocorrelation within 24 hours. Actual temporal autocorrelation exhibits an exponential decay, which is only partially modeled by the quadratic formula for SWM correlation. From a flow management viewpoint, 24 hours should be a sufficient temporal autocorrelation model, since most FM strategies are initiated and completed within a single day of air traffic activity.
Figures 4.17 and 4.18: Temporal Correlation for Ceiling and Visibility
4.3.2.3. Cross-correlation and Forecasts

Table 4.3 presents desired and synthetic results for cross-correlation.

Table 4.3: Cross-Variable Correlation Results

<table>
<thead>
<tr>
<th></th>
<th>Desired</th>
<th>Synthetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOS</td>
<td>0.60</td>
<td>0.56</td>
</tr>
<tr>
<td>LGA</td>
<td>0.60</td>
<td>0.53</td>
</tr>
<tr>
<td>DCA</td>
<td>0.60</td>
<td>0.59</td>
</tr>
</tbody>
</table>

By using the same value of $k$ in calculating the waveheights for visibility observations at every site, the SWM modeled cross-correlation as equivalent for each site. This assumption was approximately valid for BOS and LGA with cross-correlation values of 0.64 and 0.61 respectively. However, this assumption was invalid at DCA where cross-correlation was 0.48.

Concerning forecasts, the SWM model was designed to provide 3- and 6- hour forecasts of Boston ceiling observations with forecast skills of 0.6 and 0.4 respectively. The resulting forecast-observation correlations were 0.62 and 0.46. The fact that the forecast-observation correlations were somewhat above the desired levels, and the cross-correlation values were slightly low are possibly attributable to insufficient data points.

4.3.2.4. Cumulative Distribution Functions

The observed and synthetic CDFs for ceiling and visibility at each site are presented in Figures 4.19 through 4.24.
Figures 4.19, 4.20, and 4.21: CDFs for Ceiling
Figures 4.22, 4.23, and 4.24: CDFs for Visibility
The visibility data are only displayed up to 5 nm, due to the fact that 5 nm is the first threshold between two different airport weather categories (VFR to MVFR). Also, visibility observations within the weather data were aggregate over 5 nm at all three airports, i.e., most observations over 5 nm were recorded as 10 nm, 15 nm or 20 nm.

The results in Figures 4.19 through 4.24 suggest that the model accurately preserves the CDFs for all variables at all sites. The deviations between the observed and synthetic curves, reflect the fact that the SWM is a Monte Carlo simulation. The wave directions and phases were randomly generated and therefore, the results are a reflection of the true randomness in the waves. The deviation of the synthetic ceiling CDF at DCA, appears to indicate a high percentage of low value ENDs for the synthetic raw ceiling data at DCA. This asymmetry may be the result of too few data points.

4.3.3. Modeling Additional Variables

The SWM used in the test runs can be expanded to model wind speed, wind direction and precipitation. The model would require correlation statistics and the CDFs for both wind speed and direction. The observed CDF for wind speed at BOS is shown in Figure 4.25.

![Figure 4.25: Observed Windspeed CDF for BOS](image)

The observed windspeed CDFs for LGA and DCA are similar. Preliminary analysis indicates a significant cross-correlation between wind speed and visibility, as would be expected.

Concerning wind direction, it is necessary to develop a technique to ascertain and model temporal correlation amongst vectors. For example, the high correlation between 359 degrees and 1 degree must be reflected in the correlation statistic. The modeling of wind can be accomplished by adding dimensions to the SWM or using a derivation of the breathing earth
technique. If a variable possesses no cross-correlation with other variables, yet spatial correlation is exhibited, the variable can be modeled by an independent SWM.

Precipitation can also be modeled in the SWM. One alternative is to model precipitation as a 0-1 variable with desired temporal, spatial and cross-correlation. To model a 0-1 variable, it is only necessary to determine a single threshold END value to model the presence or absence of precipitation.

4.4. Discussion and Conclusion

The test results and research on the SWM indicate that it is suitable to support an airport capacity model. Since the SWM is a standalone model, the process of converting weather observations and forecasts into capacity observations and forecasts can be developed independently. One simple technique would be to reference airport capacities in a table format as defined in the Engineered Performance Standards for runway configurations at major airports. Another technique would be to assign separation distances between arriving aircraft as a function of a capacity model. Separation intervals reflect the standards regarding different types of aircraft in an airport arrival queue. This technique would provide a higher fidelity capacity model, although the implementation would be more complicated.

The assumptions of the SWM do not appear to affect the models performance. Homogeneous spatial and temporal correlation are valid if the SWM is being applied to a local region. As an alternative approach, if the model is applied to a non-homogeneous area, it is possible to simulate a weighted average of spatial or temporal correlation, or a worst-case scenario. Another assumption is that the correlation curves are equivalent for all variables being modeled. This was not a problem in the simulation runs, since the temporal correlation curves are very similar for ceiling and visibility. Furthermore, Air Force literature suggests that the model can achieve varying forms of the correlation curves among sites by varying the lengths of the sawtooth waves in the space and time dimensions [26].

In conclusion, the basic SWM, as presented in this chapter, appears to preserve the characteristics of airport weather that are of interest to ATC FM research and simulation efforts. The test results indicate that the SWM is suitable for modeling airport weather over a geographical region with spatial, temporal and cross-correlation. The model could significantly enhance an FM simulation and assist in providing realistic assessments of different FM strategies. From an implementation viewpoint, the model is simple and flexible. It can be modified easily to model seasonal and regional variations in spatial and temporal correlation, by modifying the wavelengths appropriately.
Chapter 5

5. Conclusion

As demand for air traffic services increases in the future, there will be a corresponding need to improve the efficiency of the air traffic system and reduce flight delays. Factors such as safety, fuel efficiency, and passenger satisfaction will continue to motivate air traffic researchers to develop and analyze flow management (FM) strategies that could relieve routine congestion at major airports in the United States. The three areas of research covered in this thesis are motivated by the desire to provide air traffic flow management experts and policy analysts with some of the information and models needed to conduct effective research.

5.1. ATM Primer

The first area of research in this thesis involved the preparation of a two-part primer of ATM presented in Chapter 2. The first part provides an understanding of air traffic resources from an FM perspective. The second part of Chapter 2 provides insight into the causes and dynamics of air traffic congestion and the practice of FM, as it is currently conducted by controllers and traffic management specialists.

The ATM primer was developed through an iterative review process with controllers and traffic management specialists at multiple FAA facilities. Their comments and suggestions have been incorporated into the document.

5.2. Airport Demand Scheduling Model

The second area of research was the development of a commercial airport demand scheduling model, presented in Chapter 3. The demand model was created to provide researchers and policy analysts with the capability to generate hypothetical demand scenarios for a user-defined set of airports. With this capability, experts can readily analyze the effects that predicted demand changes will have on airports, airspace sectors and the overall efficiency of the system.

The demand model, known as the Pseudo-OAG Generator (POAGG) has been implemented and verified in a computer application. By solving the commercial schedule generation problem heuristically, POAGG is fast and efficient. POAGG has already
demonstrated its utility in support of several FM research projects. The main feature of POAGG is that it generates commercial demand schedules that contain user defined levels of flight connectivity and total airport traffic levels. POAGG also models shuttle flights and airline hubbing operations. Future applications of POAGG include continued flow management research and capacity and air traffic policy analysis.

5.3. Airport Weather and Capacity Model

The third and final research topic contained in this thesis, is an investigation into an airport weather model. Previous research had developed techniques for modeling weather at airports as independent processes. These models did not preserve the spatial correlation that exists in weather observations and impacts the capacity of many airports simultaneously. Nor did previous research address weather forecasts.

The airport weather model evaluated in Chapter 4 was developed by the Air Force. By superimposing many sawtooth waves in a multidimensional coordinate system, the Air Force model simulates weather forecasts and observations at multiple sites. The data generated by the model will possess spatial correlation over a geographical region and temporal autocorrelation, i.e., correlation between successive observations of the same variables. Included in the model is a technique to preserve cross-variable correlation among variables.

The evaluation and test results of the Air Force weather model indicate that it may be an appropriate model for FM simulation environments and airport capacity modeling.

5.4. Potential Applications

The demand and capacity models will enable researchers to answer some interesting questions concerning flow management strategies. Questions that may be addressed using the demand model include:

1. What effect do different levels of slack in scheduled layover times exert on flow management strategies?

2. What is gained by solving ground holding algorithms for larger networks of airports?

3. How do flow management strategies perform when applied to networks of airports which are close together versus airports which are more spread out?
4. What are the effects of delays on flights within an airline’s shuttle services, or an arrival or departure bank, versus regular connecting flights?

5. How robust are flow management strategies in relieving congestion for air traffic demand that reflects increased levels of flight connectivity?

6. What effect will a new airline hub at an airport exert on the entire system?

7. How do increases in total traffic levels at one or more airports affect the flow of traffic within the entire system?

Questions that may be addressed using an implementation of the airport weather model include:

1. How sensitive are flow management strategies to variations in weather/capacity forecasting skill levels?

2. What role does spatial correlation amongst airports play in the problem of air traffic congestion? How can spatial correlation be incorporated into flow management strategies?

3. How do different algorithms perform comparatively in regions of the country that possess different weather characteristics?

4. How do different algorithms perform given seasonal variations within regions of the country?

The initial test results of the demand and capacity models, and feedback concerning the ATM primer, indicate that the models and the primer may assist research in FM and other important areas.

5.5. Topics for Further Research

The models for airport weather and demand presented in this thesis are preliminary developments. While both models have undergone initial verification efforts, they each require additional verification and validation testing before they can be fully implemented in ATM research models. Moreover, the two models could be enhanced through further research.
The numerical method implemented in POAGG effectively models commercial demand for a network of specified airports. The method can potentially be extended to model general aviation (GA) and military air traffic. One alternative is to insert GA and military flights into the commercial schedule using a Monte Carlo method. GA and military flights could randomly be assigned arrival times into specified airports based upon historical or hypothetical distributions of GA and military airport demand.

Additionally, the POAGG numerical method can be extended to include a higher fidelity model of individual airline scheduling. Enhancements could include modeling flight times as airline and aircraft specific, and biasing flight departure times based on airline preferences. Furthermore, the process of assigning connecting flights could be modified to model connections at the airline level. This would require input parameters and constraints that describe the scheduling practices of the individual airlines.

The basic sawtooth wave model is a reasonable technique for simulating weather conditions with spatial and temporal correlation. One of the fundamental limitations, however, is the calibration technique. The basic model, as presented, is calibrated with a single reference distance in the space domain and a single reference separation time in the temporal domain. As a result, if the correlation formula does not accurately model spatial and temporal correlation for a desired region, the sawtooth wave model will inaccurately model correlation at distances and times that differ from the reference distance and time.

The Air Force has developed a more sophisticated calibration technique that makes it possible to tailor the correlation formula and resulting curves for regions that are not accurately represented by the basic sawtooth wave model. In the basic model, all waves have the same wavelength. A more sophisticated technique incorporates varying wavelengths amongst the randomly generated waves [26]. Wavelengths are selected in order to calibrate the model with multiple reference points in time and space. By adding together waves with different wavelengths, it is possible to produce correlation curves that deviate from the basic model.

In order to complete the weather model for airport simulation purposes, it is necessary to incorporate wind speed, wind direction, and precipitation modeling. The fundamental correlation properties of the sawtooth wave model extend to additional variables, however, further research is required in specific areas, such as modeling temporal autocorrelation for wind direction.
References


20. Lin, K-C., and Ng, H., Coordinate Transformations in Distributed Interactive Simulation, Simulation, Vol 6, No. 5, November 1993.


Technical interviews:


28. LeBrie, R., Traffic Management Unit, FAA Boston Center, Nashua, NH.

29. Lavimoniere, G., Airspace and Procedures, FAA Boston Center, Nashua, NH.

30. LaBrecque, J., Traffic Management Coordinator, Boston ATC Tower, Logan International Airport, Boston, MA
