MASTER'S THESIS

Modeling Demand Uncertainties during Ground Delay Programs

by Narender Bhogadi
Advisor: Michael O. Ball

NEXTOR MS 2002-2
(ISR MS 2002-7)

NEXTOR

National Center of Excellence in Aviation Operations Research

The National Center of Excellence in Aviation Operations Research (NEXTOR) is a joint university, industry and Federal Aviation Administration research organization. The center is supported by FAA research grant number 96-C-001.

Web site http://www.isr.umd.edu/NEXTOR/

ABSTRACT

Title of Thesis: MODELING DEMAND UNCERTAINTIES DURING GROUND DELAY PROGRAMS

Degree candidate: Narender Bhogadi

Degree and year: Master of Science, 2002

Thesis directed by: Professor Michael O. Ball
Institute for Systems Research, &
Robert H. Smith School of Business

Uncertainty in air traffic arrival demand creates difficulties for the Air Traffic Control (ATC) specialists in effectively planning Ground Delay Programs (GDPs). An inefficiently planned GDP leads to excessive flight delays and underutilization of the GDP airport. GDP optimization models that exist today may not generate the best strategies for planning GDPs, as, they consider demand as deterministic, when in reality, it is highly stochastic.

In this thesis, we identify Flight Cancellations, Pop-up Flight Arrivals, and Flight Drift, as the common sources of demand uncertainties. Two models - an optimization model and a simulation model - that generate effective planning

strategies for a stochastic demand and deterministic capacity scenario, are developed. These models incorporate uncertainty in demand by associating probabilities to the stochastic demand elements during GDPs.

The results from both the models suggest that setting Planned Airport Arrival Rates (PAARs) - the number of flights that are ordered to arrive in a time period at a GDP airport - that exhibit "staircase" pattern can effectively mitigate the detrimental effects of demand uncertainties during GDPs. This is a significant finding as it opposes the current policy of setting "flat" PAAR patterns by the ATC specialists.

MODELING DEMAND UNCERTAINTIES DURING GROUND DELAY PROGRAMS

by

Narender Bhogadi

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2002

Advisory Committee:

Professor Michael O. Ball, Chair
Associate Professor Jeffrey W. Herrmann
Dr. Robert L. Hoffman

DEDICATION

I dedicate this endeavor to my parents and to my brother, who have been the constant driving force in my pursuit of knowledge.

ACKNOWLEDGEMENTS

I thank Dr. Michael Ball for entrusting me a challenging problem in the area of Air Traffic Management. Working with him has been enjoyable and educational.

I thank Bob Hoffman for all the help he has given me whenever I needed. I thank Dr. Herrmann, for being a member of my thesis committee. I also extend my thanks to Dr. Mark Fleischer for lending me technical assistance during my thesis work.

Special thanks are due to Thomas Vossen, Jason Burke and Bala Chandran, for making my stay at NEXTOR a truly memorable experience.

This work was supported in part by the Federal Aviation Administration (FAA) through NEXTOR, the National Center of Excellence for Aviation Operations Research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Air Traffic Flow Management (ATFM)

Congestion remains air transport's biggest long-term challenge. It causes widespread system delays resulting in severe inconvenience to the passengers, high revenue losses to the airlines, and heavy workload on the Air Traffic specialists. In 1998 alone, the average delay among all U.S. carrier departures, attributable to Air Traffic Control, was 7.9 minutes per flight. With more than 8 million departures by the major and national US carriers during that year, this produced a total delay of over 1 million hours! The economic losses to airlines and their passengers that year stood at a staggering sum of US$4.5 billion.

High air transport growth combined with non-commensurate developments in airport and air traffic control infrastructure has led to *constraints* on the whole air transport system. Over the years, air traffic in United States grew more rapidly than the capacity could accommodate; thereby causing congestion at nation's busiest airports. The domestic passenger traffic segment in the U.S. alone is expected to grow at a rate of 2.5 % over the next two years taking the number of domestic passengers close to the 700 million mark by 2010. Serious congestion related situations could arise if capacity isn't enhanced to meet this anticipated demand.

Further, the *hub and spoke* system of flight operations adopted by the airlines aggravates the problem due to congestion. In this system, an airlines schedules a large number of its flights to arrive and depart from its hub airport in order that passengers might make convenient connections inbound flights to outbound flights. These spurts of activities result in uneven traffic distributions at the hub airports thereby taxing their resources. For example, the largest hub, Delta's at Atlanta, has over 600 daily jet departures, where banks of up to 60 arriving and departing flights are operated 11 times a day.

To *alleviate* congestion problems, three different approaches - long-term, medium-term and short-term - based on time span were proposed by Odoni [16]. In the long-term (5-10 years), the capacity-demand balance could be achieved by augmenting the infrastructure by constructing new airports and runways, by using larger aircraft, and by employing more efficient Air Traffic Control (ATC) technologies. In the medium-term (6 months to 2 years), the approach would be to alter the temporal flow of aircraft flow in the network; for example, by imposing time-varying landing fees and user charges at airports or by auctioning the available time slots in peak times. In the short-term (daily basis and with a planning horizon of at most 6-12 hours), the approach would be to address the *Air Traffic Flow Management (ATFM) Problem* or simply, *Flow Management Problem (FMP)* i.e to optimize operations such that a best match of demand and available capacity over the time horizon of consideration can be made. There are two types of actions - tactical and strategic - that could be taken to address the FMP. The most common strategic action is to issue ground holds to the flights before their departures so that costly airborne delay could be reduced in exchange with less-expensive ground delay.

Traditionally, the Federal Aviation Administration (FAA) was the sole authority making decisions related to ATFM. Most often, these decisions were solely aimed at improving the operational performance (throughput and utilization) of the National Airspace System (NAS) and neglected the economic performance of NAS users, namely airlines and general aviation. As it is, the NAS is difficult to coordinate given its size and complexity. As of today, it consists of more than 20 Air Route Traffic Control Centers (ARTCC), approximately 700 airspace sectors, 18,292 public and private airports, 171 Terminal Radar Approach Control Facilities and a vast amount of aircraft and airport related equipment in the contiguous United States. The airlines position was that the FAA neglected the economics of airlines in the decision-making process. This view, shared by most airlines, aroused distrust among the FAA and its users (airlines) for much of the 1970's. However, in the late 1980's, things improved as the airlines and the FAA started collaborating for mutual benefit. Over the years, this collaboration lead to the establishment of a Collaborative Decision Making (CDM) body that could bring together the FAA and all the participants of the NAS into a common decision-making environment. Since its inception in 1995, CDM efforts have resulted in successful development of many efficient and technologically advanced tools and procedures that could balance the needs of both, the system and its users. The enhanced Ground Delay Program (GDP) is one of the prominent outcomes of this collaboration.

The GDP is a key tool in the area of ATFM and serves as a short-term strategic tool to address the FMP. The principal intent of a GDP is to balance arrival demand and capacity at airports by delaying flight departures at origin airports so as to avoid serious capacity-demand imbalances (CDIs) that could otherwise occur. CDIs most commonly occur when the airport capacity is severely degraded due to bad weather, though there are variety of reasons like communication equipment failure, runway incursions, airport maintenance etc that cause a reduction in airport capacity. GDP demand consists of scheduled arrivals in a given time period at an airport. GDP capacity refers to the maximum number of flight arrivals that an airport can manage safely in any given time period. The capacity of any airport is primarily dependent on the runway system capacity consisting of runways, exits from and to runways, and taxi-ways associated with runways, airport ground space, airspace, airport design, and the regional weather at the airport. But, in most of the cases, it's the limited capacity of runways that restrict the airport capacity.

The purpose of GDPs is then to replace the fuel-consuming costly and unsafe airborne delay with less costly and safe ground delay. When the capacity at any airport is reduced drastically, the FAA issues a revised set of departure times called control times of departure (CTDs) for the flights bound to this affected airport. Essentially, at the beginning of a GDP, landing slots are created at the GDP airport for each of the incoming GDP flight. The landing slots are time-based, meaning, each flight has to arrive at the airport and take its allotted slot at its Control Time of Arrival (CTA), which is set by the GDP planners at the beginning of the GDP. These CTAs are calculated to spread out aircraft arrival times at the affected airport so that the demand is evened out over time and the balance between demand and capacity is restored.

However, one assumption the GDP planners make is that demand is deterministic. Of course, demand isn't deterministic, due to inherent uncertainties associated with flight arrivals , cancellations and other such factors. In spite of well planned ground delay, aircraft may still face some airborne holding at the destination airport due to the variability in the arrival process. These unpredictabilities could either result in an increase or a decrease in the initial projected demand, for which the GDP was originally planned. If the demand that actually materializes is more than the projected demand, the ground delay imposed on aircraft wouldn't be sufficient to prevent airborne delays. On the other hand, if the actual demand turns out to be less than the projected demand, the ground delay imposed on flight would be unnecessary and would lead to under-utilization of the airport resources (slots) during the GDP. Hence, an effective GDP should incorporate these demand uncertainties and should plan for them before the GDP has actually started.

This thesis studies the effects of demand uncertainties on the performance of ground delay programs. In planning ground delay for aircraft, one would expect trade-offs among airborne holding, ground holding, and airport utilization. Balance among these three performance measures should be achieved so that the overall costs are minimized. How demand uncertainties affect these costs is the main topic of this study. Two models - a Stochastic Mixed Integer Optimization Model (SMIO) and a Simulation Model- are developed. The SMIO model generates optimal strategies for planning the number of flight arrivals that are needed to meet desired utilization at minimum expected airborne holding during each time-period of GDP, given that uncertainties in demand exist. The Simulation Model helps in validating the results from the SMIO model. Both the models, however, are used for performing the sensitivity analysis of different parameters such as flight cancellations, flight pop-ups and flight drift (described later).

1.2 Motivation for Problem Studied

The performance of a GDP depends primarily on its inputs, namely capacity and demand at the affected airport. Capacity is highly sensitive to weather and hence is stochastic. Depending upon whether there is deterioration or improvement in weather conditions, airport capacity either degrade or increase during a GDP. Similarly, demand is also stochastic as it is affected by variety of factors such as last-minute flight cancellations, flight drifts and arrival of unknown flights (pop-ups) at the airport during the GDP. To predict the number of pop-up flights or the flights that could drift during a GDP is challenging, as these events are random in nature, and specific to an airport. For example, the number of pop-up flights in San Francisco can range anywhere from 0 to 6 per hour.

The current approach of GDP planners regarding demand prediction is to treat demand as deterministic. This means that to satisfy fully the airport capacity (or Airport Acceptance Rate, AAR) of 30 flights in a given hour, a Planned Arrival Rate (PAAR) of 30 is set. Here, we would like the readers to clearly understand the distinction between AAR and PAAR. An AAR is the number of flight arrivals that the airport can safely accommodate in a time period. A PAAR, on the other hand, is the number of arrival slots that are created by GDP planners at the airport in a time period. For deterministic demand, PAAR is set equal to AAR for any time period of consideration in a GDP.

In some cases, however, GDP planners treat demand as stochastic, and give some allowance with regard to the stochastic demand elements, namely, number of cancellations, flight drifts and pop-up flights. For example, the number of flight cancellations, number of pop-ups and number of flights drifting off into later time periods could be approximated to 2, 2, and 3 per hour respectively. In this case, to fully satisfy the airport capacity (AAR) of 30 flights per hour, a PAAR - of 30 + 2 - 2 + 3 = 33 flights per hour, would be set. Clearly, this approach is not accurate as these elements exhibit considerable variation from their estimated means. And again there are no standard estimates available for the GDP planners; meaning that the estimates of one planner could vary from the estimates of another. Further all the GDPs might not be identical, warranting special handling for each one. For example, see the graphs of PAAR vs Actual Arrivals generated for the San Francisco Airport for first four hours of GDP in Figure 1.1. It can be seen that decisions related to PAARs that are based solely on experience without mathematical analysis, do not necessarily guarantee the effectiveness of the decision-making process. For example, the actual traffic that can materialize may be higher or lower than what is being planned.

The graphs clearly show that there is a gap between the PAAR and the Actual Flight Arrival Rates, as recorded by Flight Schedule Monitor(FSM). The width of the gap shows the extent of unpredictability that is inherent in flight arrival process during GDP. Unpredictabilities arise from uncertainties in flight cancellations, arrival of pop-up flights and occurrence of flight drifts during a GDP. In Figure 1.1, sometimes, the PAAR was high meaning the actual traffic was less, in which case, the GDP airport was under-utilized. At other times, the PAAR turned out to be low, indicating that the excess demand that ma-

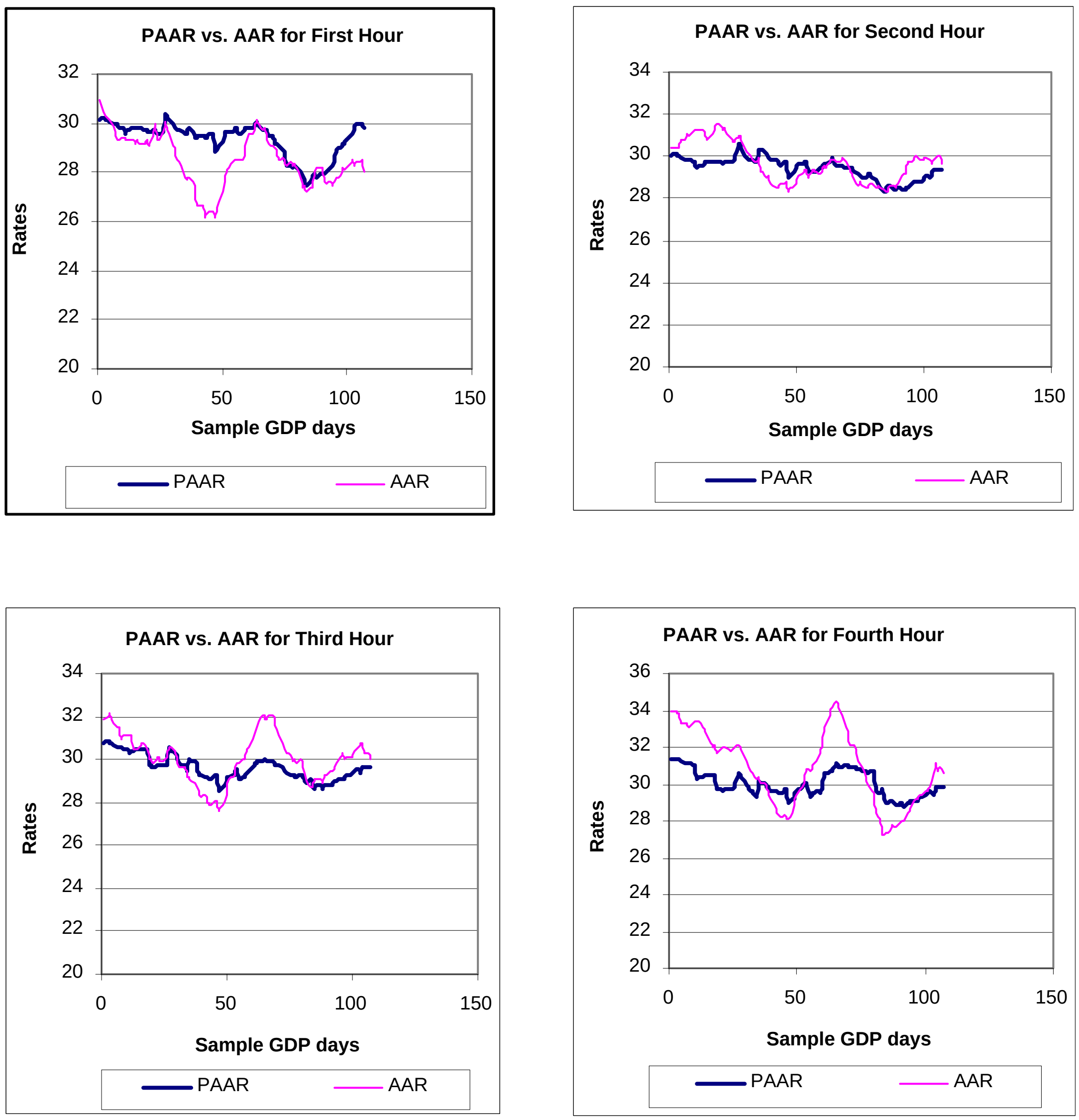

Figure 1.1: Graphs of PAAR vs Actual Arrivals at San Francisco Airport

terialized could have resulted in costly airborne holding. Thus, there is always a trade-off between airport utilization and airborne holding. Whether, a GDP planner pursues an aggressive PAAR policy by setting higher PAARs or follows a pessimistic policy by setting lower PAARs, the performance of a GDP is very much dependant on his/her decision-making capability related to PAARs.

Clearly, there is a need for a tool, which in the presence of demand uncertainties, could generate optimal strategies for planning a GDP. With an effective tool, the gap between the PAAR and Actual Arrivals could be reduced indicating that optimal balance between airborne holding and ground holding is achieved. Considering that the magnitude of airline costs from delays run into millions of dollars, this tool could bring potential savings for the airlines by enabling more efficient planning of GDPs.

1.3 Problem Description

In this section, we would state the formal definitions of the FMP and the generic FMP (also known as the Ground Holding Policy Problem GHPP) and then follow it up by defining our problem and its scope. Along the way, we would define and describe the key assumptions that go into formulating the various problems.

An excellent description of the FMP is given by Odoni in [16]. The FMP, when idealized as a network, has four essential components:

i. *Airports*, the sources and sinks of flows on the network.

ii. *Airways*, the arcs on which flows travel.

iii. *Waypoints*, the network's nodes at which airways intersect, merge or diverge.

iv. *Sectors*, collections of waypoints and contiguous segments of airways.

In most cases, airports constitute the principal bottlenecks of the ATC network. Hence, we can reasonably assume that the primary causes of congestion are the capacity-demand imbalances (CDIs) that occur at the origin and destination airports. Possible measures to restore capacity-demand balance on a short-term basis could involve actions such as delaying departure times of aircraft, imposing enroute speed control restrictions, traffic-metering, re-routing, high altitude holding and aircraft diversions. Thus, formally stated, the FMP is the short-term approach of designing a flow management system (collection of airports, airways, waypoints and sectors) to minimize the ATC delay costs, subject to operational constraints (physical and policy related).

The Generic FMP, also known as the GHPP, is a special case of the FMP when only strategic actions like assigning ground delays to aircrafts on an aggregate level are taken. Most often, a GDP serves as the means for issuing ground delays to the departing flights. Thus, the GHPP can be stated as developing optimal strategies for minimizing the airborne and ground delay costs during GDPs.

In this thesis, we tackle the GHPP under deterministic capacity and stochastic demand conditions. Our focus would be to develop optimal PAARs for each hour of a GDP that can minimize the expected airborne delays at the GDP airport. Note that ground delays are incorporated in the PAARs - the higher the PAARs, lower the ground delay, and vice-versa. Thus, formally our problem can be stated as follows:

> "Given that the demand D is stochastic, and the arrival capacity, AAR, is deterministic at an airport Z under GDP conditions, develop a model that can generate the optimal PAARs for each hour of the GDP so as to minimize the total expected airborne holding for the entire GDP duration T at a desired utilization level U for the airport Z."

In this thesis, we assume a simplified ATC network as shown in the following figure. The problem that we attempt to solve involves a single destination airport. Our focus is on the trade-offs between airborne holding and airport utilization at the GDP airport - more airborne holding in each hour of the GDP is necessary to satisfy higher airport utilization, and vice-versa. For the purposes of this problem, we assume the following:

- The only capacitated element of the ATC network is the arrival airport;

all other elements have unlimited capacity

- Travel times of aircraft between each origin and the destination airport Z is deterministic and known before the planning of GDP

- Airborne delays can occur only due to congestion at the airport Z

- Demand is stochastic for the entire GDP duration [0,T], and the parameters that characterize various stochastic elements are known in advance

- Capacity of airport Z is deterministic for entire GDP duration [0,T], and is known in advance

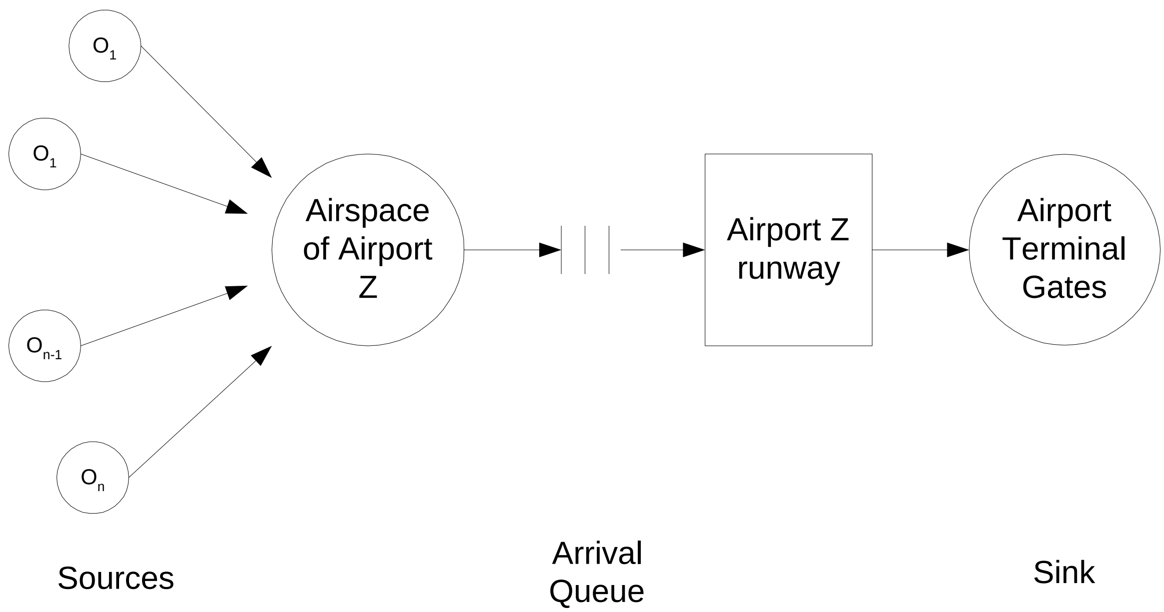

Figure 1.2: Simplified ATC Network with Airport as Main Element.

The above assumptions simplify the actual system to some extent; however, our model is still useful as it helps to generate good starting solutions (PAARs), which can provide the GDP planners with valuable insight in planning GDPs.

Specially, in scenarios, where airport capacities are predicted to a reasonable accuracy and where demands are highly stochastic, our model serves its purpose.

In concluding this section, we would like to add that the main purpose of our work is to develop a decision-making tool that could help in effective planning of GDP.

1.4 Literature Review

Many models for generating optimal strategies for minimizing ground and airborne delays during a GDP exist in the literature. Almost all of them consider demand as deterministic, and capacity as either deterministic or stochastic. These models can be divided into the following categories:

- Deterministic-Demand Deterministic-Capacity **(DDDC)**
- Deterministic-Demand Static, Stochastic-Capacity **(DDSSC)**
- Deterministic-Demand Dynamic, Stochastic-Capacity **(DDDSC)**

The **DDDC** models are particularly helpful when the capacity of an airport can be predicted with reasonable accuracy. A formulation for this model was given in Terrab [18]. The objective is to find optimal ground holding policy that minimizes the total ground delay costs:

where,

N is the number of flights scheduled to land;

T is the number of time periods for which the GDP is planned;

K_j is the capacity of the airport in period j;

DDDC formulation:

$$
\begin{aligned}
&\text{Min} && \textstyle\sum_{i=1}^{N} \sum_{j=P_i}^{P+1} C_{ij} X_{ij} \\
&\text{subject to:} && \\
& && \textstyle\sum_{j=P_i}^{P+1} X_{ij} = 1 \qquad \text{for all } i \in 1, \ldots, N \\
& && \textstyle\sum_{i=1}^{N} X_{ij} \leq K_j \qquad \text{for all } j \in 1, \ldots, P \\
& && X_{ij} \in \{0, 1\}
\end{aligned}
$$

Figure 1.3: IP Formulation of the DDDC Problem

X_{ij} is the decision variable; $X_{ij} = 1$, if aircraft i is assigned to land in period j, and $X_{ij} = 0$ otherwise;

C_{ij} is the cost incurred by aircraft i when assigned to land in period j;

P_i is the period of time during which aircraft i was originally scheduled to land.

This model can be solved very quickly using minimum cost flow or linear programming technology. The experimental results show significant savings in total delay costs given that capacity is taken to be deterministic. Further, it is shown that large savings could still be achieved even when different users are treated equitably.

DDSSC models for a *single-airport* are discussed by Andreatta and R.Jacur in [3]. In this paper, the authors propose an order of O(N^2) dynamic programming algorithm to generate optimal delay decision strategies for solving a single-period static-stochastic case of GHPP. Airport capacity is taken as a discrete random variable K which takes value $0, 1, \ldots, n$ with probabilities $p(0), p(1), \ldots, p(n)$ for the time-period of consideration. Another input, namely an optimal priority rule for flight landings, was derived by using the airborne delay costs of individual flights as a priority measure.

Later, Terrab [18] developed models that consider *multi-periods at a single airport* for DDDSC version. Here airport capacities are defined as discrete random variables that are given a probabilistic forecast that can be thought of as a number of scenarios, each scenario representing a particular instance of the random capacity vector with an associated probability. To solve small stochastic problems, a dynamic programming approach was used. For much larger problems, a greedy heuristic with some limited-look-ahead-capability was proposed. However, the authors were unable to prove that the formulation would yield an integer solution.

In 1993, Richetta and Odoni [17] used *stochastic linear programming* to solve the *single-airport* version of DDSSC. The authors extended a static-stochastic capacity model (DDSSC case) to obtain a dynamic-capacity model (DDDSC case) by overcoming the limitations of the dynamic programming formulation of Terrab. Here, the model considers Q alternative scenarios for airport capacities during the time period of interest; each scenario, q, having a probability of occurrence p_q. The model concentrates on aggregate flight groups rather than on individual flights as the GDP control mechanism for ground delays could be easily handled.

In 1999, Hoffman, Ball, Rifkin and Odoni [7] developed a polynomially solvable integer programming model for the *single-airport static stochastic GHPP (DDSSC case).* They improve on Richetta and Odoni's formulation by including fewer number of decision variables and exploit the network structure of the problem to generate optimal integer solutions. Another contribution of the authors is that their model fits with the current paradigm and procedures of CDM. During the GDP, the airport resources are divided into "slots" which are then distributed among the various airlines in an equitable manner. The fairness objective is achieved when slots are assigned to airlines based on their scheduled times of departures - the earliest slots are awarded to flights with earliest scheduled times (see section on CDM for more details). The model formulation is shown below:

DDSSC FORMULATION :

Min $\sum_{t=1}^{T} c_g G_t + \sum_{q=1}^{Q} \sum_{t=1}^{T} C_a p_q W_{q,t}$

subject to:

$$A_t - G_{t-1} + G_t = D_t \quad \text{for all } t \in 1,\dots,T+1$$
$$(G_0 = G_{T+1} = 0)$$
$$-W_{q,t-1} + W_{q,t} - A_t \geq -M_{q,t} \quad \text{for all } t \in 1,\dots,T+1$$
$$\text{for all } q \in 1,\dots,Q$$
$$(W_{q,0} = W_{q,T+1} = 0)$$
$$A_t \in Z_+, W_{q,t} \in Z_+, G_t \in Z_+$$

Figure 1.4: IP Formulation of the DDSSC Problem by Hoffman et al

Where:

A_t is the number of planes that should land in time period t,

D_t is the predicted demand(number of flights) for time period t at the airport,

G_t is the number of flights whose arrival times are adjusted from time period t to time period $t+1$ (or later) using ground delay

$W_{q,t}$ is the number of flights held in the air from time period t to $t+1$(or later) by an airborne delay under scenario q.

$M_{q,t}$ is the arrival capacity (AAR) of the airport during time t, if scenario q is realized;

p_q is the probability of occurrence of the q^{th} scenario during GDP;

c_g is the cost of ground holding a single plane for one time period;

c_a is the cost of one period of airborne delay for a single plane.

The inputs for the above model are D_t, c_g, c_a, $M_{q,t}$ and p_q. The decision variables, A_t values, can be viewed as planned airport acceptance rates (PAARs) in the sense that they represent the number of aircraft that should land in each time interval based on the planned departure times.

In 2000, Inniss, in her thesis [14],derives the probabilistic capacity scenarios for the San Francisco (SFO) airport during GDPs. These capacity scenarios (termed as Arrival Capacity Distribution ACD) serve as the inputs for the above DDSSC model developed by Hoffman et al. The ACDs are computed based on historical San Francisco GDP data.

Our work, in this thesis, does not fall under any of the above mentioned categories as the model assumes a stochastic demand. We believe our model is the first attempt to model demand as a stochastic variable and therefore represents a new category, Stochastic Demand Deterministic Capacity (SDDC). Although there has been some study in the area of demand uncertainties namely flight cancellations, pop-up flights and drift flights, by Metron , Inc., there has been no significant modeling effort in that area. In the report on the rolling spike problem [12] conducted by Metron, the drift flights (specifically, flights that depart later than their Control Times of Departures (CTDs)), and cancellations, are identified as two of the main causes that affect the uncertainty in forecasted demand for each hour of GDP. Another study [11] on pop-up traffic concludes that pop-up frequencies for an airport are highly variable from one GDP to another, and exhibit some seasonal and hourly trends, thus making them hard to predict on a hourly-basis.

In concluding this section, we note that very little work has been done to analyze the effects of demand uncertainties on GDPs. On a research level, no prior work exists. Since our work is the first in this challenging area, we would develop our thesis in a way that is meant to stimulate the reader's interest in this area and to motivate further research.

1.5 Organization of Thesis

In chapter 2, we describe the various classes of demand uncertainty in detail. Also, we introduce the readers to the philosophy of CDM and its operation. Specifically, the effects of CDM procedures on demand uncertainties are analyzed.

In chapter 3, we describe and develop the two models - SMIO and Simulation Models - for the problem we defined earlier. First, a non-linear formulation is developed for the SMIO model, then a linear formulation is developed. The Simulation Model is developed along the same lines assuming appropriate distributions for the occurrence of various uncertain elements. Towards the end of this chapter, a brief description of the data sources and probability distributions used for modeling purposes is given.

In chapter 4, we test the models under different scenarios and record the results. The SMIO model directly gives the optimal PAARs for any given scenario, while the Simulation Model requires the use of Pareto Optimality to select the optimal PAARs out of a number of different scenarios. Both the models are tested for marginal sensitivities of various demand elements on overall costs.

In chapter 5, we summarize the main contributions of this thesis. We also suggest some recommendations for policy changes in GDP planning based on our work. Finally, we provide some insights for future research in this area.

Chapter 2

Demand Uncertainty

2.1 Description of demand uncertainty

Elements or events that cannot be predicted with certainty, due to lack of complete information, are termed as uncertainties. Real-world systems almost always deal with uncertainties. Uncertainties in system processes lead to variability of the system responses to the environment, thereby degrading the system performance. Difficulties in planning and decision-making could arise in an unpredictable system. Usually uncertainties are proportional to the complexity of the system. For example, the Air Traffic Control (ATC) system, which is incredibly complex in size and scope, experiences innumerable uncertainties at different stages of planning and execution. When flight operations are being planned, uncertainties related to departures and arrivals of flights arise. Departures could be affected by variation in taxi-out times, which in turn are affected by miles-in-trail (MIT) restrictions and the inability to integrate into the overhead traffic stream. Arrival processes are affected by enroute times, and taxi-in times. Unpredictable block times - the time it takes an aircraft to travel from departure gate to arrival gate - affect the workload of pilots and hence, affect the planning of crew movements.

In the decision-making process, the ATC managers use information that is not up-to-date and often inaccurate causing difficulties for the airlines. In addition, the manner in which traffic problems are first tackled within a small local entity like an airport or a sector and then escalated to include a larger geographic area, result in inefficiencies in the decision making process. Here, each local entity optimizes its own objective function and hence, the collective performance seen for a wider region suffers. For example, in case of a GDP, inbound traffic to the afflicted airport is slowed down to reduce anticipated airborne holding. But, this action could propagate backwards over the network and could slow down overhead traffic flow that is far off from the local entity under consideration.

The best way to reduce uncertainty is to allow a better transfer of information among all the entities within the ATC framework. Specially, when limited resources are under contention, and choices regarding delay are required, economic insight can only come from the airlines. Hence the FAA should facilitate information exchange among airlines as well as various entities of ATC.

As described earlier, uncertainties are commonplace in the ATC system. Any uncertainty that could be reduced, would translate into enormous cost savings for the system, airlines and passengers. Therefore, in our work, we concentrate on a small, but significant area - demand prediction and flight scheduling during GDPs, where uncertainty is prevalent. Uncertainties in demand during GDP arise from three main elements - flight cancellations, pop-up flights and flight drift. These three elements have a combined effect of making the demand quite difficult to predict. In the later part of this chapter, we describe each of these elements in detail and analyze their effects on system performance.

2.1.1 Flight Cancellations

Airlines usually cancel their flights when they experience non-availability problems related to crew, maintenance and security personnel, ATC problems like runway breakdowns etc, and weather related problems that reduce the airport capacity. Before cancelling a flight, the airlines would weigh the economics - fuel costs saved for the cancelled flight versus cost incurred due to passenger delays and loss of goodwill - and then make a decision whether to cancel a flight or not. In most cases, decisions related to cancellations are affected by circumstances outside the control of airlines (e.g. weather problems and reduction of airport capacities). In some cases, airlines might face some operational problems and have to cancel their flights. However, there are some circumstances under which airlines cancel flights purely based on economics without any safety or other ATC problems. But whatever the reasons, the airlines have the responsibility to provide their updated flight plans to the ATC system so that airport resources can be better used in lieu of the flight cancellations.

Under the CDM framework, all participating airlines send their updated flight information through their Airline Operational Control centers(AOCs) to the hub site of Volpe National Transportation Systems Center, a federal organization within the U.S. Department of Transportation (DOT). Flight information from two other sources - NAS Monitoring Systems and Official Airline Guide(OAG) - are also supplied to Volpe. Volpe now responds by processing all the flight information and sending out CDM strings consisting of aggregate demand lists(ADLs) to each of the CDM participants through the CDMnet. Data is managed through the Enhanced Traffic Management System (ETMS) and Advanced Traffic Management System (ATMS) - ATMS has all the ETMS data and functions along with some additional functions. This flow of information within the ATC framework is shown in the figure 2.1.

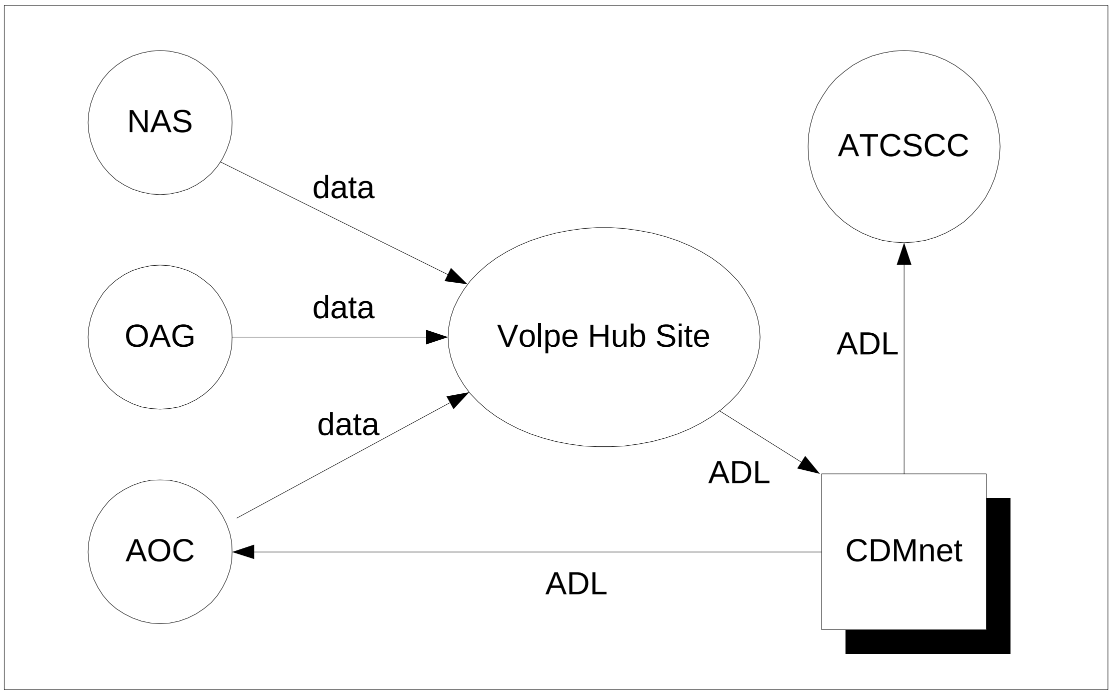

Figure 2.1: Information Flow Among Various Entities in ATC

The ADL file has approximately 61 data fields for every flight record, according to the 1999 version. In an ADL file, each record contains a comprehensive set of flight status information, including, arrival time, departure time and cancellation status of a single flight. Each flight record usually corresponds to a unique flight; if, however, two or more records for a single flight exist, the most recent record would give the accurate information about the flight. There are seven cancellation (CNX) fields or messages associated with each flight record:

- SI - Substitution Induced Cancellation

- FX - CDM airline cancellation
- RZ - NAS cancellation
- RS - OAG cancellation
- TO - timed out cancellation
- DV - diversion of destination type cancellation
- ID - the call sign of the flight has been changed causing cancellation

SI cancellations occur if the airlines choose to substitute that particular flight with another one. This often happens during GDPs - two smaller flights could be cancelled and one large flight could be substituted in their place. The *FX* message is the CDM message used by any CDM participant airline cancelling a flight. The *RZ* message is sent by airlines to indicate the cancellation of NAS flight plan for that particular flight. The *RS* message is an internal ETMS message generated when the ATC specialist takes an Official Airline Guide (OAG) flight out of the database.

The *TO* message indicates whether the flight is cancelled and timed out by the database. Flights are timed out when no activation message has been received within a certain time of predicted departure time. If the NAS or CDM messages have been received for a flight, then flight will be timed out one hour beyond its estimated departure time. In case only OAG data has been received for a flight, time out would be five minutes past OAG departure time. The DV message is given based on either NAS flight plan or CDM message, indicating that the flight would divert to an alternate destination. Finally, an ID cancellation occurs when the airlines change the flight identifying number (ID) of a flight.

For effective planning of a GDP, the timeliness of cancellation notices is crucial. If the cancellation notices are given well in advance, the ATC specialists will have sufficient time to get an accurate demand profile and can plan the GDP accordingly. But, on the other hand, tardiness in cancellation notices can distort the demand profile and disrupt the arrival sequence thereby causing wastage of slots allotted to these cancelled flights at the beginning of the GDP. The TO cancellation usually results in slots being wasted as the specialists, unaware that the flights are being cancelled, cannot risk substituting other flights into their slots. With the other type of cancellations, the system is usually able to dynamically adjust and effectively use the slot.

Clearly, SI, FX, RZ, RS, DV, and ID type of cancellations are not of a major concern as far as airport utilization is concerned. So, in this thesis, when we talk cancellations we mean TO cancellations as we are more interested in last-minute cancellations that could create holes in the arrival sequence that can't be filled (see Figure 2.2). To fill the vacant slots that appear due to flight cancellations,

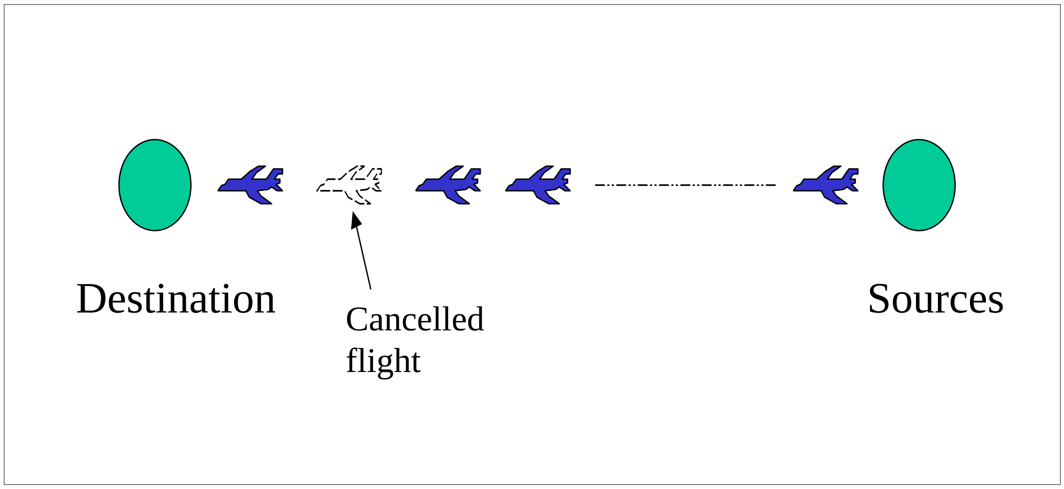

Figure 2.2: Cancellations in the Flight Arrival Stream

the ATC specialists can plan for more flight arrivals than the airport capacity can handle. The trade-off here is that if the buffer size of the flights turns out to be more than the number of vacant slots, then the additional flights undergo airborne holding. Estimating the possible number of vacant slots during GDP is difficult due to the variability in the cancellations. For example, on a single day (see Figure 2.3) that we analyzed, we found that the number of cancellations-induced vacant slots varied from 0 to 5 during the entire duration of the GDP.

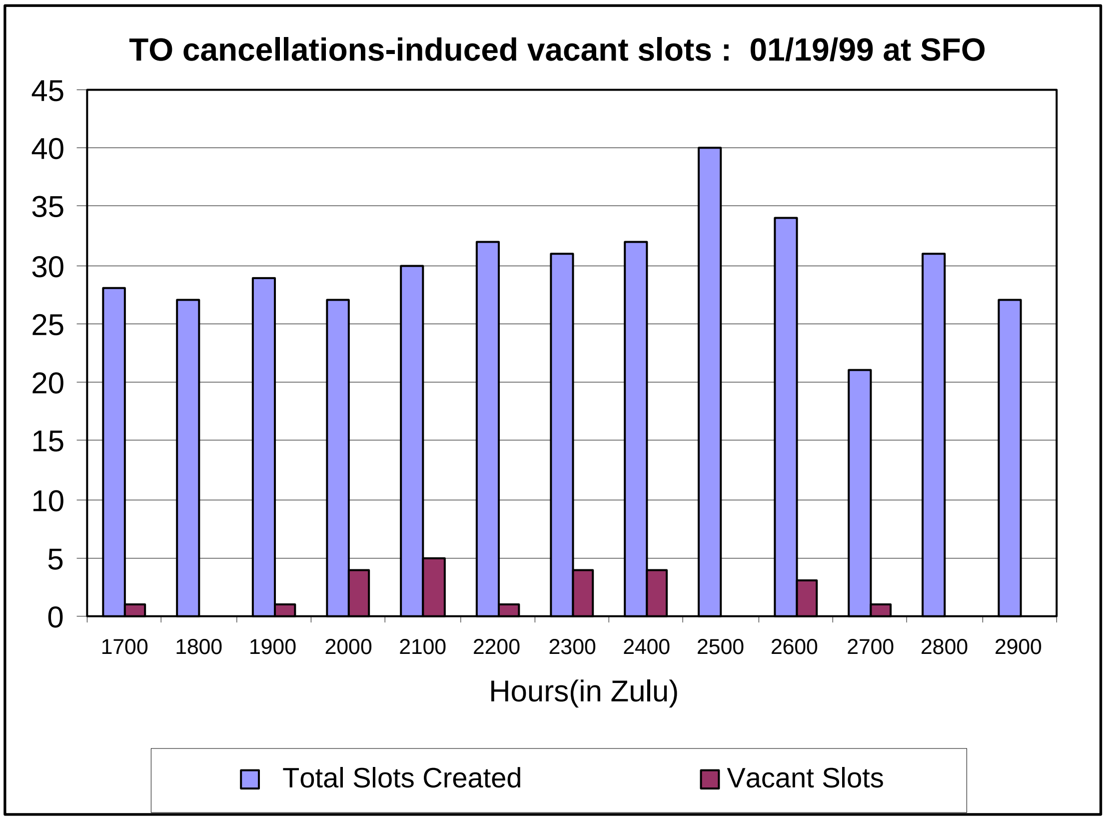

Figure 2.3: Graph of TO Cancellation-Induced Vacant Slots

In the worst cases, there could be at least 10% loss of slots. Also, the number of cancellation-induced vacant slots could be airport- and airline-specific, and season-dependent too. Hence, cancellations are to be incorporated in the GDP plans for a better airport utilization.

2.1.2 Pop-up Flights

Pop-up flights are defined as the unexpected flights that arrive during the Ground Delay Programs (see Figure 2.4). Pop-ups mainly consist of corporate jets, air-taxis, military aircraft, and last-minute flights created by airlines to accommodate overbooked passengers. Currently there are two definitions of pop-ups. They are

> Any flight that arrived during a GDP without schedule information in the ADL is defined as a Pop-up flight.

> Any flight that arrived during a GDP and that first appeared in the ADL after the GDP model time.

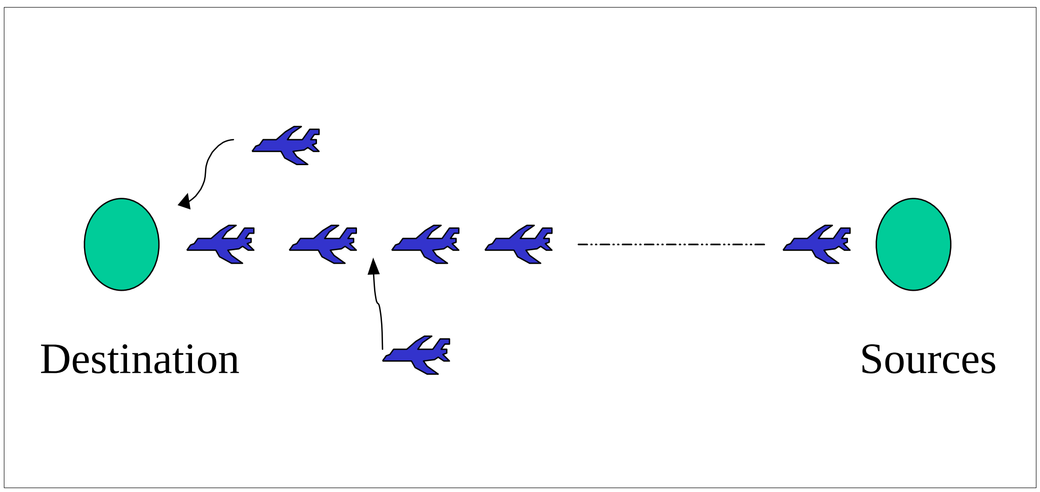

Figure 2.4: Pop-up Flights During GDP.

The schedule information of any flight is its information published in the Official Airline Guide (OAG). The GDP model time can be stated as the time when the GDP planners make their decisions about various GDP parameters, with the information available at that time.

By first definition of pop-up, it is possible that a flight that is known to the GDP planners before the model time, but which has no OAG information, might be treated as a pop-up. Clearly, this flight is not unexpected, at least, at the time of planning for the GDP. By second definition, however, this flight is not considered as a pop-up. Since information related to this flight is made available to the GDP planners, it is no longer an unexpected flight, as the GDP planners can take this flight into account and plan suitable actions.

Of course, there are always some flights that are unknown to the system until the last minute of their arrival at the airport. Such flights are definitely covered within the scope of the first definition. But, the second definition seems to incorporate such flights in addition to flights which are known to the system but have no OAG times. To illustrate what is called a pop-up flight, we use a simple example. Suppose that the GDP model time is 1500z hours and that the GDP starts at 1800z hours. Now, any flight that appears in the ADL after 1500z hours is treated as a pop-up according to second definition. This is because, when the GDP is planned, these flights were not known to the GDP planner and hence, not planned for. Also, by second definition, as stated earlier, it is possible that flights with no information in the OAG could still be not treated as pop-ups if they first appear in the system before the GDP model time. Thus, the second definition seems to differ from the first one in that all flights not known to the system at some vantage point of time are pop-ups. We should note that since GDPs are dynamically adjusted, e.g. through the compression and revision functions, it is possible that even the second definition doesn't capture the problem with complete accuracy.

Until now, there have been only two systematic studies [11] of pop-up traffic, both completed by Metron Aviation Inc. The first study was conducted in July 2000 and presented in the CDM meeting. This study made use of the first definition of pop-ups. The major conclusions of the July 2000 study were as follows.

- Pop-ups are more prevalent during GDPs than during non-GDPs (normal days).
- Pop-up flights are more likely to be cancelled than non pop-up flights.
- General Aviation flights were only 3% of all flights, but 30% of all pop-ups.

The first two results combined together reflect the dynamics of pop-ups during the GDPs. The last result indicates that the General Aviation (GA) category form a large portion of pop-up flights. Usually, most of the pop-ups are GA or military or sometimes last-minute creations of one of the airlines. If these flights are not planned for, they can seriously *disrupt the equitable distribution* of airport resources (slots) and also *displace the sequence* of the scheduled traffic. According to currently existing policies, the airport is supposed to provide landing slots for the pop-up flights as and when they arrive. If this is so, some of the allotted slots to airlines during the beginning of the GDP could be taken away from them and reassigned to pop-up flights, which is unfair. The other effect of pop-ups is to displace the sequence of aircraft which could, in general, lower airport utilization as there are certain factors like ground separation, MIT restrictions and other such separation rules that apply while the flights takeoff or land at the airports.

The second study was conducted in the summer of 2001 by Metron Inc [11]. This study was more extensive than the earlier one as it covers nine airports for a period of two years. It is to be noted that for this study pop-up was defined according to second definition. The pop-ups were classified into numerous categories based on airlines, aircraft sizes, CDM member status and many more.

Thus, this study gives more valuable insight into the pop-up phenomenon. Some of the important results of this study are as shown below:

- On an average for all airports, 7.0% of GDP arrivals were pop-ups.
- The air carriers account for 46% of pop-ups. However, since 79% of the GDP arrivals belong to air carriers, the air carriers produce less than their "fair share" of pop-ups.
- General aviation flights make up the second largest user type in pop-ups: 35%, and yet only 4% of the GDP arrivals. This over-representation is to be somewhat expected, since all general aviation flights during a GDP are pop-ups.

The 7.0 % figure for pop-up flights during a GDP is considered quite high by the community and demonstrates the need to control the pop-up phenomenon during GDPs. The second result shows that GA flights are 35% of all pop-ups indicating that any action to control pop-ups should start with control of GA flights.

Presently, the Flight Schedule Monitor (FSM), the CDM decision-support tool, allows the traffic specialist to compensate for pop-up flights by setting a "GA" factor when the GDP is planned. This means that a certain number of slots are set aside for pop-ups, while the remaining slots are included in the GDP planning for allocation to regular flights. The GA factor depends on the expected pop-up rate per hour at the airport. For example, if it were known that the average pop-up rate at SFO is 3 per hour when the capacity of the airport (slots) is 30 per hour, the GA factor would be set equal to 3 and the remaining 27 slots are allocated to regular flights in an equitable way. A shortcoming of this approach is the deterministic way of predicting the pop-up rate for a given capacity scenario. Studies show that pop-up rates per hour are highly variable, dependent on airport, seasons and airlines. In figure 2.5, it can be seen that average pop-ups per hour could vary anywhere from 0 to 10. Thus, setting a deterministic pop-up rate clearly has limited effectiveness.

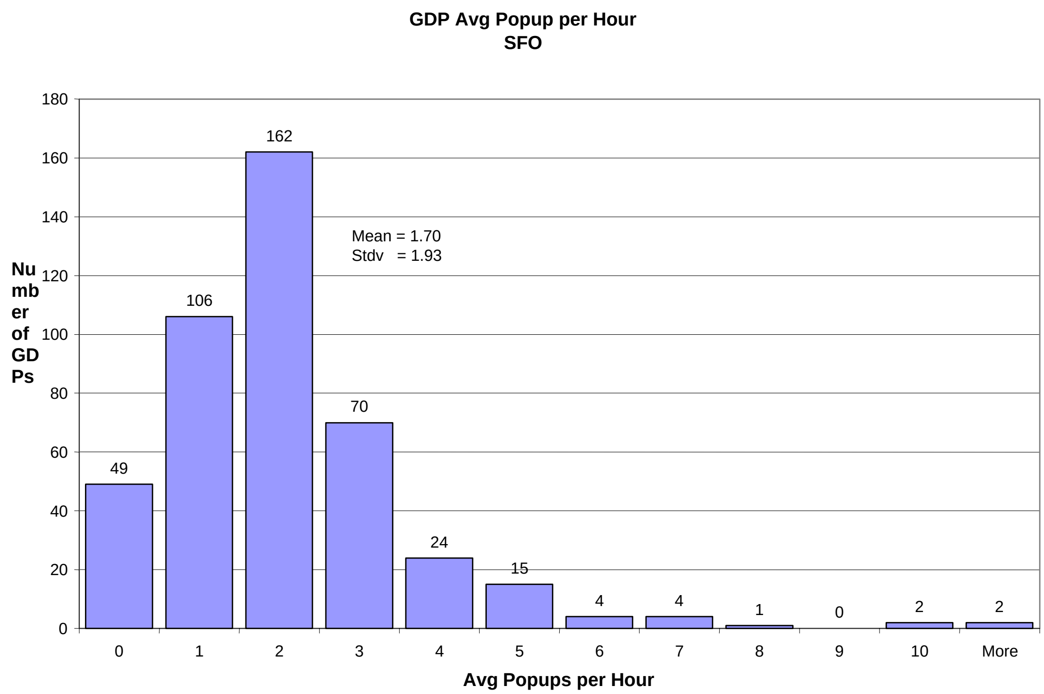

Figure 2.5: Average Pop-up Flights per Hour (Courtesy : Metron Inc.)

Finally, the effect of pop-up flights on the performance of GDP is significant, both in magnitude and nature of impact. Equity and predictability of flight arrivals are two issues that are directly affected by pop-ups.

2.1.3 Flight Drift

Flights are given Control Times of Arrivals(CTAs) and Control Times of Departure (CTDs) during GDPs. Essentially, each flight has to takeoff at a particular time and arrive at a particular time to take its assigned slot at the destination airport. Now, there are always some flights that drift with time and land in either earlier or later arrival slots than their actual slots. This phenomenon is called Flight Drift.

Flight Drift result from CTA non-compliance of flights, meaning, flights are not arriving at the appointed slot time. CTA non-compliance could occur in two cases:

- CTD non-compliance, where $ARTD \neq CTD$
- CTD compliance (i.e. $ARTD = CTD$) but $AETE \neq OETE$

In first case, the actual runway time of departure (ARTD) of flights could be either sooner or later than their CTDs. This drift is termed Ground Drift. Therefore,

$$\text{Ground Drift} = ARTD - CTD$$

In the figure 2.6, the flight with CTD of 1715z hours could depart either before 1715z hours (at 1645z, 1700z etc hours) or after 1715z hours (at 1730z, 1745z etc hours) resulting in ground drift.

In most of the cases, the ground drift of flights is unrecoverable. The only way to recover ground drift is to increase air speed. However, this is not common due to various enroute restrictions like MIT restrictions, Sector loads etc.

In the second case, even assuming that flights depart at their CTDs, they incur some drift in air before they arrive at the destination airport. This drift

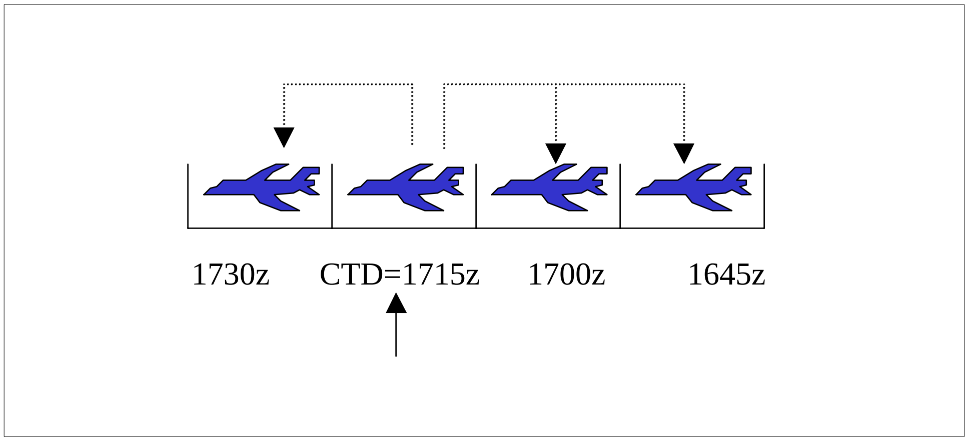

Figure 2.6: Ground Drift Due to CTD Non-compliance

results when Actual Enroute Time(AETE) is either more or less than OETE. This is termed as Enroute Drift. Hence,

$$\text{Enroute Drift} = AETE - OETE$$

The most common reason for enroute drift can be attributed to variability in overhead air traffic. If a flight arrives later than its CTA, it is termed Forward Drift; alternatively, if the flight arrives before its CTA, it is termed Backward Drift. In Figure 2.7, the flight with CTD=1700z and CTA=1915Z hours could incur enroute drift and end up taking an arrival slot either before its CTA (at 1900 z hours or before) or after its CTA (at 1930z hours or beyond).

Thus, each flight has a ground drift and an enroute drift. Hence, the net drift for each flight could be given as:

$$\text{Net Drift} = \text{Ground Drift} + \text{Enroute Drift}$$

Drift during GDPs can lead to delays, unpredictable flows in air traffic and under-utilization of airports. In a study [10] by Metron, drift was identified

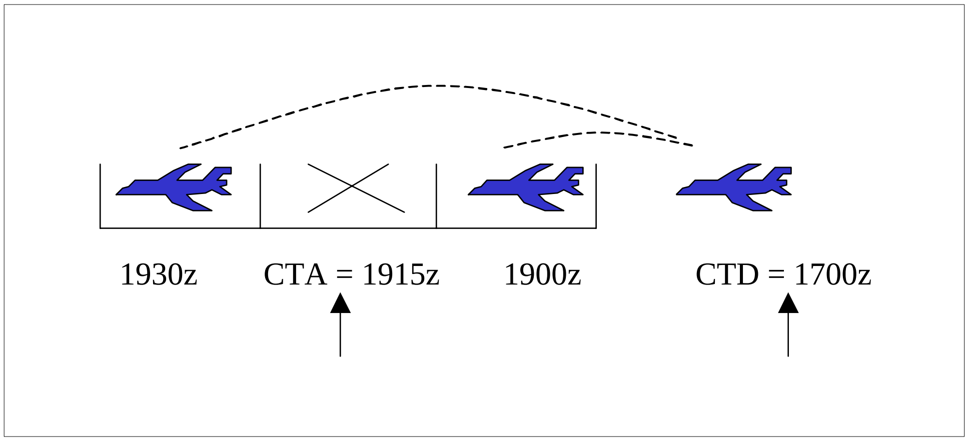

Figure 2.7: Enroute Drift

as one of the main causes for *under-delivery of airports* during GDPs. In this study, they analyzed GDPs at four different airports and found that drift was very common during GDPs and had a significant impact on the performance of an airport. Early drift in the first hour and late drift in the last hour leads to reduced demand at the airports causing under-utilization. Drift that occurs during middle hours have little impact as each hour loses and gains roughly the same amount of flights; hence, there is little or no under-delivery for those hours. The statistics for the number of flights that drifted (missed their CTAs) for the selected GDP days at four different airports namely Atlanta (ATL), Chicago O'Hare (ORD), Boston (BOS), and Philadelphia (PHL) airports are as shown in Table 2.1.

Clearly, the statistics demonstrate that the forward drift is more common than background drift. The worst case statistics belong to the Atlanta (ATL) airport, wherein the number of flights that drifted, roughly, come to 3 per hour of GDP. This is significant as each flight that drifts could potentially result in

Table 2.1: Analysis of Drift on Sample GDP days (Courtesy: Metron Inc)

Airport	**GDP day**	**Duration**	**Backward Drift**	**Forward Drift**
ATL	06-05-00	1800-0059Z	4	18
ORD	08-10-00	1700-0259Z	3	15
BOS	05-26-00	2000-0259Z	2	3
PHL	05-13-00	1800-2359Z	2	1

an unused slot during that hour, or unwanted airborne holding in some other hour.

Departure compliance of aircraft is the best way to reduce delays and to increase predictability of arrivals. Currently, the departure compliance window during GDP is [-5, +15] meaning a flight could depart no earlier than 5 minutes and no later than 15 minutes. However, this window was considered too wide to effectively mitigate the impact of drift. Recently, the FAA, on an experimental basis, tested the effect of tightening the departure compliance window to [-5, +5] minutes. It was found that flight drift was significantly low under reduced time window.

2.2 CDM and Demand Uncertainties

2.2.1 CDM Philosophy

Collaborative Decision Making (CDM) is an effort to improve air traffic management through information exchange, procedural improvements, tool development, and common situational awareness (see [21] and [4]). It serves as an efficient approach in allocating the scarce resources such as airport runways, airport terminal gates, and air traffic flow management takeoff slots to all the users of NAS in an equitable manner. Originally conceived within the Federal Aviation Administration (FAA) - Airlines Data Exchange (FADE) project, it proved that with real-time submission of airlines operational information to the FAA, decisions that directly impact the NAS users and the Air Traffic system, could be made better.

Before CDM, there was a notion among various NAS participants, specially within the airline community, that the GDPs were excessively controlled, not giving enough flexibility to the airlines. The FAA was seen as making economic decisions for the airlines. For example, assume that FAA assigns two flights of a particular airlines to slots A and B during GDP. Suppose that the first flight can not make it to slot A or if, for some reason, the economic impact of the second flight is more significant than the first flight, then the airline would ideally like to re-assign slot A to the second flight and slot B to the first flight. Under the stringent FAA policies at that time, this substitution might not be possible. Thus, prior to CDM, the conditions were not conducive to co-operation and trust between the system (FAA) and its users (airlines and others).

In 1993, the FADE experiment showed there could be large scale benefits from incorporating dynamic schedule information from the airlines into decision making. It is clear to see that the airline schedules are not static as published in Official Airline Guide (OAG) but are dynamic owing to the weather and other conditions. Hence, the decisions made by the FAA prior to CDM were based on less accurate data. The FAA and the airlines were quick to realize the potential benefits of CDM and came forward to promote this concept.

Today CDM has about 47 airlines as active participants along with the FAA and works towards developing tools and procedures that can benefit every user of the system. In its eight years of implementation [5] and [6], CDM has

- established a communications infrastructure to supply both the FAA and participating airlines with a common arrival demand picture at every major airport in the United States.

- removed (unintentional) disincentives for the airlines to report up-to-date flight status and intention information(e.g. During January through May, flight cancellation notices received under CDM, on average, were at least 63 minutes earlier at SFO, than they would have been without CDM).

- developed a mechanism (the Compression algorithm) to perform dynamic, inter-airline slot swapping that utilizes arrival slots vacated by cancelled or delayed flights;

- provided traffic flow managers with the ability to revise program parameters during a GDP that are dependent upon stochastic conditions (e.g., airport acceptance rate);

- disseminated accurate aggregate forecasts of arrival demand at all major airports in the US to all traffic flow managers and to all airline operational control centers (AOCs)

- distributed to the Air Traffic Control System Control Center (ATCSCC) and to all AOCs a uniform set of decision support software and airport demand monitoring tools (the Flight Schedule Monitor) for formulating and analyzing GDPs.

In the study on pop-ups [11], it was found that CDM participants make up 78% of the GDP arrivals, but only 32% of the pop-ups, while non-CDM participants make up 22% of the GDP arrivals, but 68% of the pop-ups. Therefore, CDM participants are under-represented in the pop-up class while non-CDM participants are dramatically over-represented. This result demonstrates that CDM has had a definite impact on improving information quality.

In all, CDM helped in reducing the airborne delays of traffic by helping flow of information among various NAS users. Before CDM, accurate information was not made available to the ATC specialists leading to an inefficient use of airport resources (slots). However, with the advent of CDM, the possibility of slots going unutilized, reduced to a great extent due to the underlying incentives and effectiveness of various CDM methodologies and tools for all NAS users.

2.2.2 Effects of Compression Algorithm on Demand Uncertainties

Most often it happens that if airlines were cancelling flights without substituting any of their flights in the vacant slot, the slot would go unused during the GDP. A new mechanism called Compression was developed to adjust flights delays to fill in the vacant slots "holes".

The Compression Algorithm essentially is a dynamic tool designed to move flights up in the arrival hierarchy during a GDP in order to fill slots vacated by cancelled flights. The algorithm associates an owner (airlines) with each arrival slot in the GDP duration. In the event that a slot is vacated due to a flight cancellation, then compression seeks to move a feasible flight of the owning airline as close as possible to that slot. If none of the flights of that airlines is available to fill the empty slot, only then, a flight of another airlines(possibly competitors) would be considered for slot allocation. Thus, Compression ensures not only that the vacant slots are filled but also that equity is maintained while filling the vacant slots. This way, an airline receives a benefit by trading in a slot it cannot otherwise use. Other (competing) airlines also receive a benefit (to the extent necessary to provide a usable slot to the airline that freed up the original slot). This is viewed as a win-win situation for all.

A brief illustration of Compression is shown below. The Table 2.2 gives the initial allocation of slots at the beginning of a GDP. Notice that the total delay imposed on the flights is 69 minutes.

Table 2.2: Initial Slot Allocation to Flights During GDP

Airlines	**Flight-ID**	**ETA(hh:mm)**	**CTA(hh:mm)**	**Delay(min)**
Delta	DAL1	9:45	10:00	15
United	UAL1	9:57	10:05	08
American	AAL1	10:03	10:10	07
Delta	DAL2	10:09	10:15	06
USA	USA1	10:13	10:20	07
American	AAL2	10:14	10:25	11
United	UAL2	10:15	10:30	15
			Total Delay	69

Now assume that UAL1 is cancelled. This cancellation creates a vacancy of a slot as shown in Table 2.3. Now suppose if no compression was applied, meaning, the flights below UAL1 are not pushed up in the hierarchy, then the total delay would be 61 minutes as shown in table 2.3 and also, a slot gets wasted. However, if some form of compression is used by just pushing up the flights below UAL1, then the delay reduces to 36 minutes and no slot is wasted. However, the delay incurred by the UAL2 flight is 10 minutes. In this case, no explicit incentive is given to the United Airlines for relieving its slot; though, system wise, the total delay achieved was minimum. This is shown in the table 2.4.

Table 2.3: Delay and Slots Assignment in Lieu of a Cancellation

Airlines	**Flight-ID**	**ETA(hh:mm)**	**CTA(hh:mm)**	**Delay(min)**
Delta	DAL1	9:45	10:00	15
United	UAL1	cancelled	10:05	-
American	AAL1	10:03	10:10	07
Delta	DAL2	10:09	10:15	06
USA	USA1	10:13	10:20	07
American	AAL2	10:14	10:25	11
United	UAL2	10:15	10:30	15
			Total Delay	61

Now, we apply compression with equity considerations, given that the flight UAL1 is cancelled. Since UAL1 belongs to United Airlines, this particular airline should get some incentive for vacating the UAL1 slot. The Compression algorithm provides some sort of bartering among the various airlines for the slot exchanges in an equitable manner (refer [9] and [20] for more details). In this particular instance, United Airlines exchanges its vacant slot with Delta Airlines, securing in return a convenient slot for its immediate flight beneath UAL1 - namely , slot 10:15 for UAL2. After the final exchange of slots, the flights are

Table 2.4: Slots Assignment with Compression in Absence of Equity

Airlines	**Flight-ID**	**ETA(hh:mm)**	**CTA(hh:mm)**	**Delay(min)**
Delta	DAL1	9:45	10:00	15
American	AAL1	10:03	10:05	02
Delta	DAL2	10:09	10:10	01
USA	USA1	10:13	10:15	02
American	AAL2	10:14	10:20	06
United	UAL2	10:15	10:25	10
-	-	-	10:30	-
			Total Delay	36

moved up in the hierarchy as shown in the table 2.5.

The total delay of 36 minutes is the same in both the cases of compression. However, notice the delay of United Airlines. UAL2 flight reduced its delay from 15 minutes to 0 minutes. Thus, Compression with an element of equity is much more reasonable and attractive for the airlines, as they get their incentives even if they disclose their cancellations in time. As seen from the simple example, no airline that is reporting its cancellation is losing to any other competing airlines. In addition to equity and delay savings, timely notices of cancellations is also one of the biggest benefits of Compression.

It was reported [9] that between January 20,1998 and July 15,1999, the percent delay savings for EWR and SFO airports were 13.0% and 9.7% respectively upon execution of compression algorithm. The savings mentioned here are the assigned ground delay savings and do not necessarily reflect airborne holding savings (or increase). In a different study [5] by National Center of Excel-

Table 2.5: Revised Slots Assignment and Delays After Compression

Airlines	**Flight-ID**	**ETA(hh:mm)**	**CTA(hh:mm)**	**Delay(min)**
Delta	DAL1	9:45	10:00	15
American	AAL1	10:03	10:05	02
Delta	DAL2	10:09	10:10	01
United	UAL2	10:15	10:15	0
USA	USA1	10:13	10:20	07
American	AAL2	10:14	10:25	11
			Total Delay	36

lence in Aviation Operations Research (NEXTOR), one-half of the compression benefits are realized as destination delay savings, with the rest being offset by airborne holding delays. One more benefit of compression is that FAA is now able to deliver a *constant smooth arrival rate* at the airports. That is to say compression has improved *'predictability' of arrival stream.* As this thesis will show, this will translate into savings in airborne delay.

2.2.3 Effects of RBS Algorithm on Demand Uncertainties

The purpose of the Ration-by-Schedule (RBS) algorithm [9] is to ration arrival slots to the airlines according to the scheduled times of the flights, where "scheduled Time" are the published times in Official Airline Guide (OAG). We note that since the allocation of slots is done based on arrival times that are created long before the GDP is run, the allocation process is independent of any delay information that the air carriers might submit.

To understand the significance of scheduled-times-based allocation, we should first look at the algorithm used prior to the existence of RBS. This algorithm is called Grover Jack and the slot allotment was done on a first-come-first-serve basis [9]. That is to say flights were assigned slots based on their "Estimated Time of Arrivals (ETAs)" rather than OAG times. There were two major objections to using ETAs in slot allocations, namely

- Flights that are already delayed prior to slot allocation at the afflicted airport, received an even greater delay as their ETAs are used for slot allocation.This is known as a double-penalty.

- If a flight was cancelled just prior to allocation, no compensation was given to the airline that released its slot. Moreover, the competing airline that secured the slot benefitted at the expense of the "donor" airline.

The double penalty-issue can be explained based on the table below. Table 2.6 gives the initial slot ownership of airlines based on their OAG times. Now, since the Grover-Jack algorithm looks at only the ETA for slot allocation, the reassigned slots are as in Table 2.7. The delay incurred by UAL1 flight is 7 minutes (including initial delay of 2 minutes). Thus, flight UAL1 incurs a double penalty.

When the Grover-Jack was used for slot allocation, the airlines generally objected to the two issues of double-penalty and non-compensation for forsaking their slots. They argued that sending accurate information about their flights might actually was detrimental to the position of the airline in question. Thus, airlines refrained from providing up-to-date information on their flights to the FAA as they found no incentive to do so. This proved to be a serious impediment to the flow of information, especially, in times of bad weather, when the need for the FAA to manage airspace efficiently was most critical.

Table 2.6: Initial Slots Assignment Based on OAG Times

Airlines	**Flight-ID**	**ETA(hh:mm)**	**CTA(hh:mm)**
Delta	DAL1	10:01	10:00
United	UAL1	10:08	10:05
American	AAL1	10:04	10:10
Delta	DAL2	10:07	10:15
United	UAL2	10:20	10:20
American	AAL2	10:25	10:25

Table 2.7: Slot Assignment Based on Grover-Jack

Airlines	**Flight-ID**	**ETA(hh:mm)**	**CTA(hh:mm)**	**Delay (min)**
Delta	DAL1	10:01	10:00	00
American	AAL1	10:04	10:05	01
Delta	DAL2	10:07	10:10	03
United	UAL1	10:08	10:15	07
United	UAL2	10:20	10:20	00
American	AAL2	10:25	10:25	00
			Total Delay	11

To remedy the drawbacks of the Grover Jack algorithm, the CDM working group developed the RBS algorithm which removes the disincentives just discussed. The RBS algorithm is illustrated in Table 2.8. Clearly, United Airlines retains the slot allotted to UAL1 prior to the beginning of program, no matter what the ETA of UAL1 is. Further, United Airlines exchanges its UAL1 slot with American Airlines AAL1 slot, thereby accruing only 2 minutes of delay for UAL1 flight i.e. no additional delay on top of its initial delay.

Table 2.8: Slots Assignment Based on RBS

Airlines	**Flight-ID**	**ETA(hh:mm)**	**CTA(hh:mm)**	**Delay (min)**
Delta	DAL1	10:01	10:00	00
American	AAL1	10:04	10:05	01
United	UAL1	10:08	10:10	02
Delta	DAL2	10:07	10:15	08
United	UAL2	10:20	10:20	00
American	AAL2	10:25	10:25	00
			Total Delay	11

In concluding this section, we note that the main benefits of RBS is that airlines now can provide updated information with trust and knowledge that their sharing of information does not hamper their business goals, but only improves the performance of the system. Additionally, RBS provides airlines with full control of their slots and the decision to trade them with other airlines or to substitute their own flights, is left entirely to their discretion.

Chapter 3

Modeling Demand Uncertainties

3.1 Stochastic Mixed-Integer Optimization (SMIO) Model

In Chapter 2, we have seen how uncertainties in demand affect the performance of GDP with respect to airport utilization and airborne delays. We have also stated that better information exchange among various ATC entities would help reduce the degree of demand uncertainty. However, uncertainty can only be reduced but not eliminated altogether. Hence, the system should consider uncertainty as a part of everyday process and respond in an effective manner to mitigate the effects of uncertainty.

In this section, we discuss the two versions of the Stochastic Mixed-Integer Optimization (SMIO) model that we developed to help the ATC specialists in planning effective GDPs. To remind the readers of the formal statement of our problem, we state it again:

> "Given that the demand D is stochastic, and arrival capacity AAR is deterministic at an airport Z under GDP conditions, develop a

model that generates the optimal PAARs for each hour of GDP so as to minimize the total expected airborne holding for the entire GDP duration T at a desired utilization level U for the airport Z"

Both the versions of SMIO model generate optimal PAARs for every hour of GDP in the presence of the uncertain demand elements. However, the SMIO model incorporates only two types of uncertainty namely - uncertainty associated with pop-up flights and cancellations. Drift could not be incorporated as the structure of SMIO model was not amenable due to the presence of a Markovian requirement, which is not satisfied by drift.

3.1.1 Model Description

The Stochastic Mixed-Integer Optimization (SMIO) Model is formulated as a 0-1 IP that consists of an objective function and five types of constraints (see [22] for in-depth description of Mixed-Integer Programs). The objective function to be minimized represents the expected size of the airborne queue at the airport for the entire GDP duration. A set of 0-1 decision variables X_{paar}(k,t) is specified for each time period t. Time is discretized into one hour blocks.

Initially, we develop a non-linear formulation for this model, then follow it up with a linear formulation that can be more easily solved. Essentially, our model associates probabilities to each event - pop-ups and cancellations - to compute an expected number of pop-ups and cancellations during the GDP. The formulation and explanation of the model is elaborated in the next section.

3.1.2 Non-Linear Formulation

Let $t \in 1, \ldots, T$ be a discrete set of continuous one hour time intervals spanning the entire duration of a GDP at airport Z. Airport capacities (also known as airport acceptance rates, AARs) are taken as deterministic for all time periods $t \in 1, \ldots, T$ and are given at the start of the GDP. When the GDP specialists plan for a GDP, they plan for a certain level of utilization depending upon various criteria. Since, the airport resources are distributed into hypothetical slots during the GDP, utilization could be measured as the ratio of the number of utilized slots to the total available slots during the entire GDP. Let ε be the target number of unutilized slots during the entire GDP. The total available slots in any period is equal to the AAR in that period. Hence, utilization can be calculated given ε, and AARs for each period of GDP.

The Planned Airport Arrival Rate (PAAR) is the number of flights that are ordered to be sent to the airport Z when the GDP is issued. Now, we assume MinPaar and MaxPaar as the upper and the lower bounds for PAARs respectively. These bounds are restricted to a small but reasonable interval around the AARs so that the solution search space could be manageable.

Finally, we take into account the airspace capacity of an airport. Every airport has its own limitation on the number of flights that can occupy an airborne queue. The general reasons are related to safety and space requirements. To capture the airspace capacity constraint in our model, we assume a maximum allowable airborne queue size $MaxQ[t]$ for every interval t of the GDP at the airport Z.

The non-linear formulation has the following variables and coefficients:

- $X_{paar}(k,t) \in \{0,1\}$; $X_{paar}(k,t) = 1$, if PAAR $= k$ in time period t, and 0 otherwise.

- $Y(j,t) \in [0\ ,\ 1]$; $Y(j,t)$ is the probability that there are j flights in the queue at the end of time period t.

- $q(i,j,t) \in [0\ ,\ 1]$; ; $q(i,j,t)$ represents the transition probability that there will be j flights in the airborne queue at the end of period t, given there are i flights in the queue at the end of period $t-1$.

- $v_e(i,t) \geq 0$; $v_e(i,t)$ is the expected number of unutilized slots in period t given that there are i flights in airborne queue at the end of period $t-1$

The objective function of the model, which minimizes the expected airborne queue size, can be formulated as shown below:

$$\textbf{Minimize} \sum_{t=1}^{T} \sum_{j=0}^{MaxQ[t]} j\ Y(j,t)$$

The various constraints of the model are explained below:

- For any time period t, PAAR can take only a single value in the interval [MinPaar , MaxPaar]. For example, if the PAAR for time period $t = 2$ is set at 27, then $X_{paar}(27,2) = 1$ and hence, we have $X_{paar}(k,2) = 0 \quad \forall \quad k \neq 27$, and $k \in$ [MinPaar , MaxPaar]. This constraint is formulated as shown below:

$$\sum_{k=MinPaar}^{MaxPaar} X_{paar}(k,t) = 1 \qquad \forall \quad t \in 1,\ldots,T+1 \tag{3.1}$$

- The airborne queue at the beginning of a GDP, i.e., in time period $t = 0$ would be zero. Hence, we have the constraints:

$$Y(0,0) = 1 \tag{3.2}$$

$$Y(j,0) = 0 \qquad \forall \;\; j \in 1, \ldots, MaxQ[0] \tag{3.3}$$

- Let the airborne queue size at the end of time period t be j, queue size at the end of time period $t-1$ be i, the number of flight arrivals during time period t be $N_{arr}[t]$ and finally, AAR[t] be the airport acceptance rate in time period t. Then the following equality holds in our model setting:

$$j = i + N_{arr}[t] - AAR[t]$$

This equality states that since $N_{arr}[t]$ and AAR[t] are independent and random in nature, the airborne queues exhibit a Markovian Property [19]. This means the queue j in time period t is only dependent on the queue i in time period $t-1$. The queue probabilities are modeled as shown below:

$$Y(j,t) = \sum_{i=0}^{MaxQ[t-1]} q(i,j,t) \; Y(i,t-1)$$
$$\forall \;\; j \in 0, \ldots, MaxQ[t], \quad t \in 1, \ldots, T \tag{3.4}$$

These transition probabilities are determined in the following way.

$$q(i,j,t) = \Pr\{(j,t) \mid (i,t-1)\}$$

We know that $j = i + N_{arr}[t] - AAR[t]$. So, we can formulate $q(i,j,t)$ as follows:

if $j = 0$, then

$$q(i,j,t) = \sum_{k=MinPaar}^{MaxPaar} \Pr\{N_{arr}[t] \leq AAR[t] - i \mid X_{paar}(k,t) = 1\} \;\; X_{paar}(k,t)$$

else if, $1 \leq j < MaxQ[t]$, then

$$q(i,j,t) = \sum_{k=MinPaar}^{MaxPaar} \Pr\{N_{arr}[t] = j{-}i{+}AAR[t] \mid X_{paar}(k,t) = 1\} \quad X_{paar}(k,t)$$

else,

$$q(i,j,t) = \sum_{k=MinPaar}^{MaxPaar} \Pr\{N_{arr}[t] \geq j{-}i{+}AAR[t] \mid X_{paar}(k,t) = 1\} \quad X_{paar}(k,t)$$

The exact form of the transitional probabilities depend on the probability distributions for cancellations and pop-up arrivals. $X_{paar}(k,t)$ is included in the transitional probability because the number of flight arrivals is directly dependent on the PAAR. Note that the presence of $X_{paar}(k,t)$ variable in the transitional probability gives non-linearity to constraint (3.4).

- We discussed earlier the notion of slots during a GDP and how airport utilization could be measured. In the presence of random events, we can model the utilization in terms of the expected number of utilized slots during the GDP. This is shown in the constraint below:

$$\sum_{t=1}^{T} \sum_{i=0}^{MaxQ[t-1]} v_e(i,t)\, Y(i,t-1) \;\leq\; \varepsilon \tag{3.5}$$

We state that the expected number of unutilized slots during the entire duration of GDP is not more than a certain number given by ε. $v_e(i,t)$ can be computed as follows:

$$v_e(i,t) = \sum_{k=MinPaar}^{MaxPaar} \sum_{h=0}^{R} (R-h) Pr\{(N_{arr}[t] = h) | X_{paar}(k,t) = 1\} X_{paar}(k,t)$$

$$\forall \;\; i \in 0, \ldots, MaxQ[T-1] \quad \text{and} \quad \forall \;\; t \in 1, \ldots, T$$

where, the variable R is given as $R = AAR[t] - i$.

Constraints (3.2), (3.3), (3.4), and (3.5) along with the non-linear objective function formulated in the beginning of this section constitute non-linear formulation.

The above optimization problem can be solved, but solving it efficiently for large instances may be challenging. For the problem at hand, it is highly desirable to find a linear formulation in order to take advantage of the significant advances in mixed IP solvers.

3.1.3 Linear Formulation

Linear Formulation: Version 1

To generate a linear formulation from the non-linear formulation just described, we introduce a new set of variables:

- $Y(k,i,t) \in [0\ ,\ 1]$; $Y(k,i,t)$ represents the probability that there are i flights at the end of time period $t-1$ if the planned arrival rate in period t is k, and equals 0 otherwise. In some way, $Y(k,i,t) = Y(i,t-1)X_{paar}(k,t)$

- $q(k,i,j,t) \in [0\ ,\ 1]$; $q(k,i,j,t)$ represents the transition probability that there are j flights in the queue at the end of period t; given that there are i flights in the queue at the end of period $t-1$, and the planned arrival rate (PAAR) in period t is k.

- $v_e(k,i,t) \geq 0$; $v_e(k,i,t)$ represents the expected number of unutilized slots in period t, given the planned arrival rate (PAAR) in period t is k, and there are i flights in queue at the end of period $t-1$.

The objective function for this version of linear formulation can be written as:

$$\textbf{Minimize} \sum_{t=1}^{T+1} \sum_{k=MinPaar}^{MaxPaar} \sum_{j=0}^{MaxQ[t]} j \;\; Y(k,j,t)$$

Constraint (3.1) remains. However, constraints (3.2), (3.3), (3.4), and (3.5) are no longer valid upon introduction of the new set of variables. The new constraints, therefore, are formulated in the following way:

$$\sum_{k=MinPaar}^{MaxPaar} Y(k,0,1) = 1 \tag{3.6}$$

$$\sum_{i=0}^{MaxQ[t-1]} Y(k,i,t) \;\; \leq \;\; X_{paar}(k,t)$$

$$\forall \;\; t \in 1,\ldots,T+1 \qquad \forall \;\; k \in MinPaar,\ldots,MaxPaar \tag{3.7}$$

$$X_{paar}(MinPaar, T+1) = 1 \tag{3.8}$$

$$\sum_{k'=MinPaar}^{MaxPaar} Y(k',j,t+1) \; - \sum_{i=0}^{MaxQ[t-1]} q(k,i,j,t)\, Y(k,i,t) \;\; \geq \;\; X_{paar}(k,t) - 1$$

$$\forall \;\; j \in 0,\ldots,MaxQ[t] \quad \forall \;\; t \in 1,\ldots,T \quad \forall \;\; k \in MinPaar,\ldots,MaxPaar \tag{3.9}$$

$$\sum_{k'=MinPaar}^{MaxPaar} Y(k',j,t+1) \; - \sum_{i=0}^{MaxQ[t-1]} q(k,i,j,t)\, Y(k,i,t) \;\; \leq \;\; 1 - X_{paar}(k,t)$$

$$\forall \;\; j \in 0,\ldots,MaxQ[t] \quad \forall \;\; t \in 1,\ldots,T \quad \forall \;\; k \in MinPaar,\ldots,MaxPaar \tag{3.10}$$

$$\sum_{t=1}^{T} \sum_{k=MinPaar}^{MaxPaar} \sum_{i=0}^{MaxQ[t-1]} v_e(k,i,t)\, Y(k,i,t) \;\leq\; \varepsilon \tag{3.11}$$

The various constraints stated in this version of linear formulation are explained as follows:

- Constraint (3.6) replaces (3.2) of non-linear model. At the beginning of a GDP, the queue size is zero and the PAAR in the first hour could be any value. Note that constraint (3.3) of the non-linear model is actually redundant and hence, not included in the linear model.

- In formulating constraint (3.7), we coupled the $X_{paar}(k,t)$ variable with the $Y(k,i,t)$ variable. If $X_{paar}(k,t) = 1$, then summation of $Y(k,i,t)$ over all possible queues would be equal to 1 as $Y(k,i,t)$ is dependent on k; else, if $X_{paar}(k,t) = 0$, the summation of $Y(k,i,t)$ over all possible queues would be equal to 0. Thus, the constraint is valid.

- Constraint (3.8) is formulated to ensure the validity of constraints (3.9) and (3.10).

- Constraints (3.9) and (3.10) are a pair of coupling constraints that are generated after decomposing constraint (3.4). However, the Markovian feature of the model is still intact. If $Xpaar(k,t) = 1$, then the right term of the inequalities (3.9) and (3.10) reduce to 0 - implying that the probability that there is queue j at the end of period t is equal to summation of all probabilities that there are queues of sizes 0 to MaxQ[t-1] at the end of period $t-1$. If $X_{paar}(k,t) = 0$, then (3.9) and (3.10) become redundant as $Y(k,i,t)$ equals 0, and we get $\sum_{k'} Y(k^{'},j,t+1) \leq 1$ and $\sum_{k'} Y(k^{'},j,t+1) \geq -1$, which is always true.

We would now show how to compute the transition probability $q(k,i,j,t)$. Using the same logic as we did in non-linear case, we could arrive at the following formulation for $q(k,i,j,t)$.

If j=0, then

$$q(k,i,j,t) = Pr\{(N_{arr}[t] \leq j - i + AAR[t])|(X_{paar}(k,t) = 1)\}$$

else if, $1 \leq j < MaxQ[t]$

$$q(k,i,j,t) = Pr\{(N_{arr}[t] = j - i + AAR[t])|(X_{paar}(k,t) = 1)\}$$

else,

$$q(k,i,j,t) = Pr\{(N_{arr}[t] \geq j - i + AAR[t])|(X_{paar}(k,t) = 1)\}$$

To compute the probability that $N_{arr}[t]$ flights arrive in time period t, given that $X_{paar}(k,t) = 1$, we need to make some assumptions. All flights are treated as homogenous commodities and all flight arrivals are assumed to be binomially distributed. The regular flights whose arrivals are planned for, during the GDP, follow a binomial distribution with the probability of arrival $(1 - P_{cnx})$, where P_{cnx} is the probability of cancellation for each flight. Pop-up flights are assumed to follow a different binomial distribution with arrival probability of $(1 - P_{popcnx})$, where P_{popcnx} is the probability of cancellation for each pop-up flight. Since, pop-up flights are independent of PAAR, we assume a maximum number of pop-ups during any period to be equal to "N_p" flights per hour. The transition probability can now be computed as follows.

If j=0, then

$$q(k,i,j,t) = \sum_{y=0}^{U_b} \sum_{x=0}^{N_p} \binom{k}{y-x} (1-P_{cnx})^{y-x} P_{cnx}^{(k-y+x)} \binom{N_p}{x} (1-P_{popcnx})^{x} P_{popcnx}^{(N_p-x)}$$

$$\forall\ i \in 0,\ldots,MaxQ[t-1] \quad ,\forall\ j \in 0,\ldots,MaxQ[t],$$

$$\forall\ k \in MinPaar,\ldots,MaxPaar, \quad \forall\ t \in 1,\ldots,T$$

where,

k is the PAAR in time period t,

y is the number of flight arrivals in time period t,

x is the number of pop-up flight arrivals in period t, and

U_b is the upper bound for y and equals to '$j + AAR[t] - i'$ and we assume $U_b \geq N_p$.

Similarly, when $1 \leq j < MaxQ[t]$, then

$$q(k,i,j,t) = \sum_{x=0}^{N_p} \binom{k}{y-x} (1-P_{cnx})^{y-x} P_{cnx}^{(k-y+x)} \binom{N_p}{x} (1-P_{popcnx})^{x} P_{popcnx}^{(N_p-x)}$$

$$\forall\ i \in 0,\ldots,MaxQ[t-1] \quad ,\forall\ j \in 0,\ldots,MaxQ[t],$$

$$\forall\ k \in MinPaar,\ldots,MaxPaar, \quad \forall\ t \in 1,\ldots,T$$

where,

y represents the number of arrivals and equals to '$j + AAR[t] - i'$ and we assume $y \geq N_p$

$y - x$ represents the number of arrivals of regular (planned) flights during period t

x represents the number of arrivals of pop-up (unexpected) flights during period t

Finally, for $j > MaxQ[t]$,

$$q(k,i,j,t) = \sum_{y=LB}^{UB} \sum_{x=0}^{N_p} \binom{k}{y-x} (1-P_{cnx})^{y-x} P_{cnx}^{(k-y+x)} \binom{N_p}{x} (1-P_{popcnx})^{x} P_{popcnx}^{(N_p-x)}$$

$$\forall \ i \in 0,\ldots,MaxQ[t-1] \quad ,\forall \ j \in 0,\ldots,MaxQ[t],$$

$$\forall \ k \in MinPaar,\ldots,MaxPaar, \quad \forall \ t \in 1,\ldots,T$$

where,

LB represents the lower bound of y and equals '$j + AAR[t] - i$', and

UB represents an upper bound of y, which we fixed at a value equal to '$j + k - AAR[t] + 2N_p$'

Thus, $q(k,j,i,t)$ can be computed and used in the model.

- Constraint (3.11) was formulated by modifying constraint (3.5) of non-linear model. $v_e(k,j,t)$ can be calculated in the same way as in case of non-linear version. This is shown below:

$$v_e(k,i,t) = \sum_{s=0}^{AAR[t]-i} (AAR[t] - i - s) \quad Pr\{(N_{arr}[t] = s) | (X_{paar}(k,t) = 1))\}$$

$$\forall \, i \in 0,\ldots,MaxQ[t-1] \quad ,\forall \, k \in MinPaar,\ldots,MaxPaar, \quad \forall \, t \in 1,\ldots,T$$

where, the arrival probabilities can be computed as

$$Pr\{(N_{arr}[t] = s) | (X_{paar}(k,t) = 1))\}$$

$$= \sum_{x=0}^{N_p} \binom{k}{y-x} (1-P_{cnx})^{y-x} P_{cnx}^{(k-y+x)} \binom{N_p}{x} (1-P_{popcnx})^{x} P_{popcnx}^{(N_p-x)}$$

where,

y represents the number of arrivals and equals to '$j + AAR[t] - i$'

$y - x$ represents the number of arrivals of regular (planned) flights during period t

x represents the number of arrivals of pop-up (unexpected) flights during period t

The size of this formulation is polynomial in the problem data. We found that for larger problems, the constraint set becomes unmanageable. Consequently, we constructed another version which has a lower number of constraints.

Linear Formulation: Version 2

The computational burden in the linear-formulation of version 1 is mainly attributable to constraints (3.9) and (3.10). So, we remodified constraints (3.9) and (3.10) as follows:

- The variables $X_{paar}(k,t)$ are de-coupled from the coupling constraints (3.9) and (3.10) and the new formulation is shown below:

$$\sum_{k'=MinPaar}^{MaxPaar} Y(k', j, t+1) = \sum_{k=MinPaar}^{MaxPaar} \sum_{i=0}^{MaxQ[t-1]} q(k,i,j,t)\, Y(k,i,t)$$

$$\forall j \in 0, \ldots, MaxQ[t] \quad \forall\ t \in 1, \ldots, T \quad (3.12)$$

The structure of this set of constraints essentially takes the same form as the set of constraints (3.4) in the non-linear formulation. That is to say the probability that there are j flights in queue at the end of period t (the left term in (3.12)) is equal to the product of transition probability, and the probability that there are i flights at the end of period $t-1$ (the right term in (3.12)).

So, version 2 formulation consists of constraints (3.1), (3.6), (3.7), (3.8), (3.12), and (3.11) with the same objective function as in version 1 of linear formulation.

The number of constraints in both the versions of the linear formulation is computed below by assuming $MaxQ[t] = MaxQ \quad \forall \quad t \in 1, \ldots, T$.

$$\text{Version } 1 = (T+3) + (T)(MinPaar - MaxPaar + 1)(1 + 2(MaxQ + 1)$$

$$\text{Version } 2 = (T+3) + (T)(MinPaar - MaxPaar + 1) + (T)(MaxQ + 1)$$

The reduction in the number of constraints in version 2 when compared to version 1 would evaluate to

$$(MaxQ + 1)(T)(2(MaxPaar - MinPaar) + 1)$$

To give an idea of the amount of computation saved, assume that MaxQ=15, MaxPaar = 40, MinPaar = 20, and T = 8. The total number of constraints in version 1 and version 2 are 5555 and 307 respectively. Hence, the reduction in the number of constraints in version 2 when compared to version 1 comes to 5248, which is approximately 94% reduction in this particular case. With respect to computational time, version 2 could take few hours to solve while version 1 could solve in minutes. This is further discussed in chapter 4 of this thesis.

3.2 The Simulation Model

The main purpose of the simulation model is to validate the SMIO model. The other purpose of this model is to measure true sensitivities of various uncertain elements on the overall delays and utilization. Recall that we could not incorporate drift in the SMIO model. The simulation model makes up for the deficiency in the SMIO model by incorporating all uncertain demand elements including drift.

In this section and the next one, we describe the model as well as the data analysis underlying it.

3.2.1 Model Assumptions

We make the following assumptions for modeling a GDP:

- we assume the airport to have only a single runway and model the airport system as a single server queuing system. In reality, there could be many runways at the destination airport facilitating multiple flight arrivals at the same time. However, to simplify our model, we assume a single runway at the GDP airport. This assumption implies that at any given time, the server (airport) has only one customer (flight) that is engaged in service. Until that flight completes its service and departs, the next flight in queue has to wait.

- The airborne queue starts only at the runway of the destination airport. This simplifies the queuing model but in no way affect the trade-offs among various performance measures. In practice, flights are usually put in some sort of sequence much ahead of the local GDP airport owing to the enroute airspace restrictions like miles-in-trail and sector loads.

- Both, the number of time periods T, and the start time of the GDP are known. Our model uses Monte-Carlo simulation technique, which requires all the input data at the start of the GDP.

- The airport acceptance rate (AAR) is deterministic and assumes only one value per period. The period can be quarterly or hourly. As said in our problem statement, we are modeling deterministic AAR scenario for a GDP. The AARs usually take a single value for a time period as they are governed by runway configuration, runway capacity and wind factors. Sometimes, due to some sudden weather changes or temporary runway incursions, AARs may take two or more values for a time period.

- The GDP flights are homogenous (of same sizes). In practice, the utilization of an airport is not precisely defined; as the airport capacity depends on a variety of factors, such as aircraft arrival mix and the sequence of arrivals. For example, a sequence of flights with a large flight followed by a small flight may deliver different utilization when compared with a sequence of flights with a large flight followed by a large flight and a small flight followed by a small flight, though the total number of flight arrivals in both the cases may be the same. One of the reasons is that a small flight that follows a large aircraft has to maintain a good separation in air, to cope with the turbulent wake left behind by the large flight. Hence, by making an assumption that all flights are of same sizes, we can measure utilization as the ratio of actual arrivals / airport arrival capacity for each hour (or for entire GDP duration), without loss of generality.

3.2.2 Model Description

To plan a GDP, GDP specialists should have information of the GDP airport arrival capacity, AAR, and the GDP airport demand. Given these two parameters, specialists set a planned arrival acceptance rate (PAAR) such that the overall delays are minimized, and the airport capacity is properly utilized. Therefore, the three main performance measures for a GDP are ground delay, airborne delay and utilization. Ground delay is controllable while airborne delay and utilization are dependent on the stochastic nature of the flight arrivals. For a GDP to be effective, the right balance among all the three measures is required.

The whole GDP process can be viewed as assigning revised times of departures (namely Control Time of Departures CTDs), and assigning landing slots to all the GDP flights by specifying the Control Time of Arrivals (CTAs). This is shown in the figure 3.1

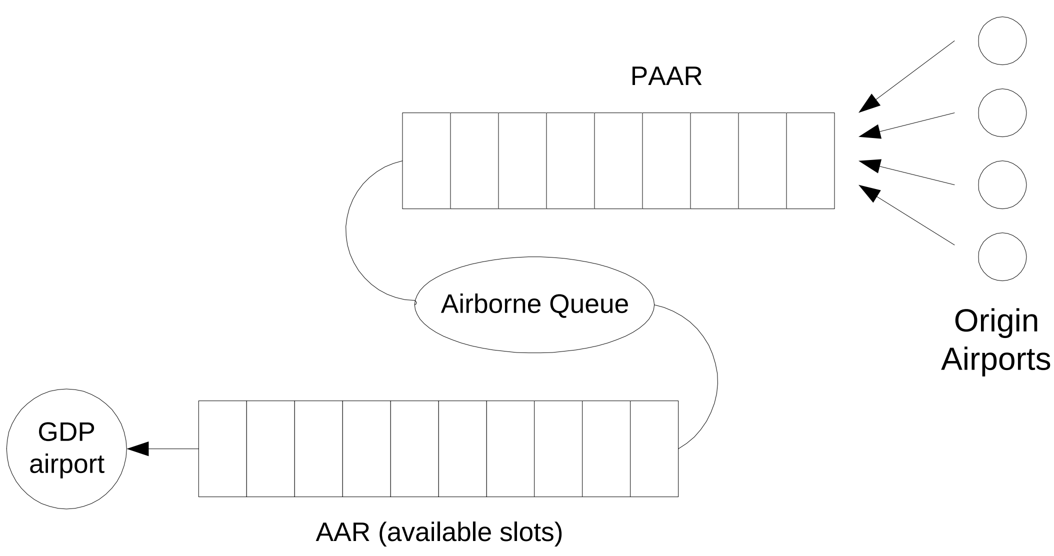

Figure 3.1: Slot Representation of GDP Process

The following information characterizes the simulation model for a GDP:

- Start Time S, and the number of Time Periods T, each of width W_t
- Planned Sequence of Flight Arrivals $f_0, \ldots, f_n$
- Planned Inter-Arrival Times $a_1, \ldots, a_T$
- Service Times $\mu_1, \ldots, \mu_T$
- Performance Measures

The planned sequence of flights $f_0, \ldots, f_n$, scheduled to arrive at the GDP airport, is assumed to be known apriori; and, n is a sufficiently large number to warrant a GDP. This sequence must be known to assign ground delays to the flights before the start of a GDP. Once the GDP starts, the actual sequence of flight arrivals may differ from planned sequence due to stochastic demand elements such as drift, cancellations and pop-ups.

The planned inter-arrival time a_t between any two flights that are planned to arrive in time period t is given by:

$$a_t = W_t / Paar_t$$

where, W_t is the width of the time period t, and $Paar_t$ is the PAAR for the time period t. Technically speaking, PAAR is a rate i.e., it has units of flights per unit time; however, for modeling purposes, it has units of flights. Typically, PAARs are set on a hourly basis during GDPs, and hence, the time periods are mostly hourly blocks. Similarly, the service time for all flights arriving in the same time period t is assumed to be the same. Hence, service time in time period t is given by:

$$\mu_t = W_t / AAR_t$$

where, AAR_t is the AAR for time period t and W_t is the width of the time period t.

The first performance measure, *ground holding* gh_i assigned to any GDP flight f_i, is given by:

$$gh_i = CTD_i - OGTD_i$$

where, CTD_i is the Controlled Time of Departure, and $OGTD_i$ is the Original Gate Time of Departure at the origin airport. In the simulation model, we represent ground holding gh_i as :

$$gh_i = S + ia - OGTD_i$$

where, S represents the start time of the GDP, i represents the slot assigned to the flight f_i, and a represents the planned inter-arrival time between any two flights.

Similarly, the second performance measure, expected *airborne holding* ah_i for any flight f_i, is given by:

$$ah_i = ARTA_i - CTA_i$$

where, $ARTA_i$ is the Actual Runway Time of Arrival, and CTA_i is the Controlled Time of Arrival at the origin airport. In the simulation model, airborne holding for any flight is viewed as the time spent by the flight waiting in the queue to be served. This is shown mathematically below:

$$ah_i = \text{Actual Service Time}_i - \text{Actual Arrival Time}_i$$

where, Actual Service Time_i is the time at which the service for the flight f_i begins. This measure would be zero if there were no uncertainties in the arrival process. However, uncertainties do exist, and lead to variability in the actual arrival times. We represent the actual arrival time of flight f_i as

$$\text{Actual Arrival Time}_i = S + ia + \Delta_i$$

where, Δ_i is the deviation suffered by the flight f_i from its planned arrival time. These deviations in flight arrival times are caused by three elements - Flight Cancellations, Pop-up Flights, and Drift. Each of these elements is modeled in the following way:

Cancellations In our model, each flight will be canceled with a probability p_{cnx}. If cancellation were the sole source of uncertainty, the inter-arrival times would follow a geometric distribution with a mean $p_{cnx}/(1-p_{cnx})$ and variance $p_{cnx}/(1-p_{cnx})^2$. Based on our analysis of real data, we estimated that cancellations were indeed following a geometric distribution with $p_{cnx} = 0.05$.

Pop-ups We modeled pop-ups as a Poisson process, meaning, their inter-arrival times are exponentially distributed with mean λ. Based on our analysis of data from SFO airport, we determined an estimate of λ.

Drift We computed empirical distributions for drift based on GDP data analysis. More specifically, flight f_i which incurs drift will arrive at the airport at time

$$t_i = S + ia + D_i$$

where, D_i is the net drift encountered by the flight (See subsection 2.1.3 of chapter 2 for full explanation of drift).

Finally, the third performance measure, expected *utilization* u, for a GDP airport, is given by:

$$u = \text{(Actual Number of Flight Arrivals)/(Available Airport Capacity)}$$

where, the numerator is the summation of flight arrivals for all periods of the GDP, and the denominator is the summation of AARs for all the periods of the GDP. Upon converting the numerator and the denominator into time units, the utilization measure looks like:

$$u = \text{(Total Server Busy Time)/(Total Server Available Time)}$$

The simulation model for GDP is implemented in C language. The discrete events that trigger a change in the system state - Arrival of Planned Flights, Arrival of Pop-up Flights, and Departure of Flights - are implemented as different event functions outside the 'main' of the C program. The random numbers generated by the main program, and its functions, is based on Marse and Roberts' Fortran random-number generator UNIRAN [15]. The implementation logic and structure of the C program is shown in Figure 3.2.

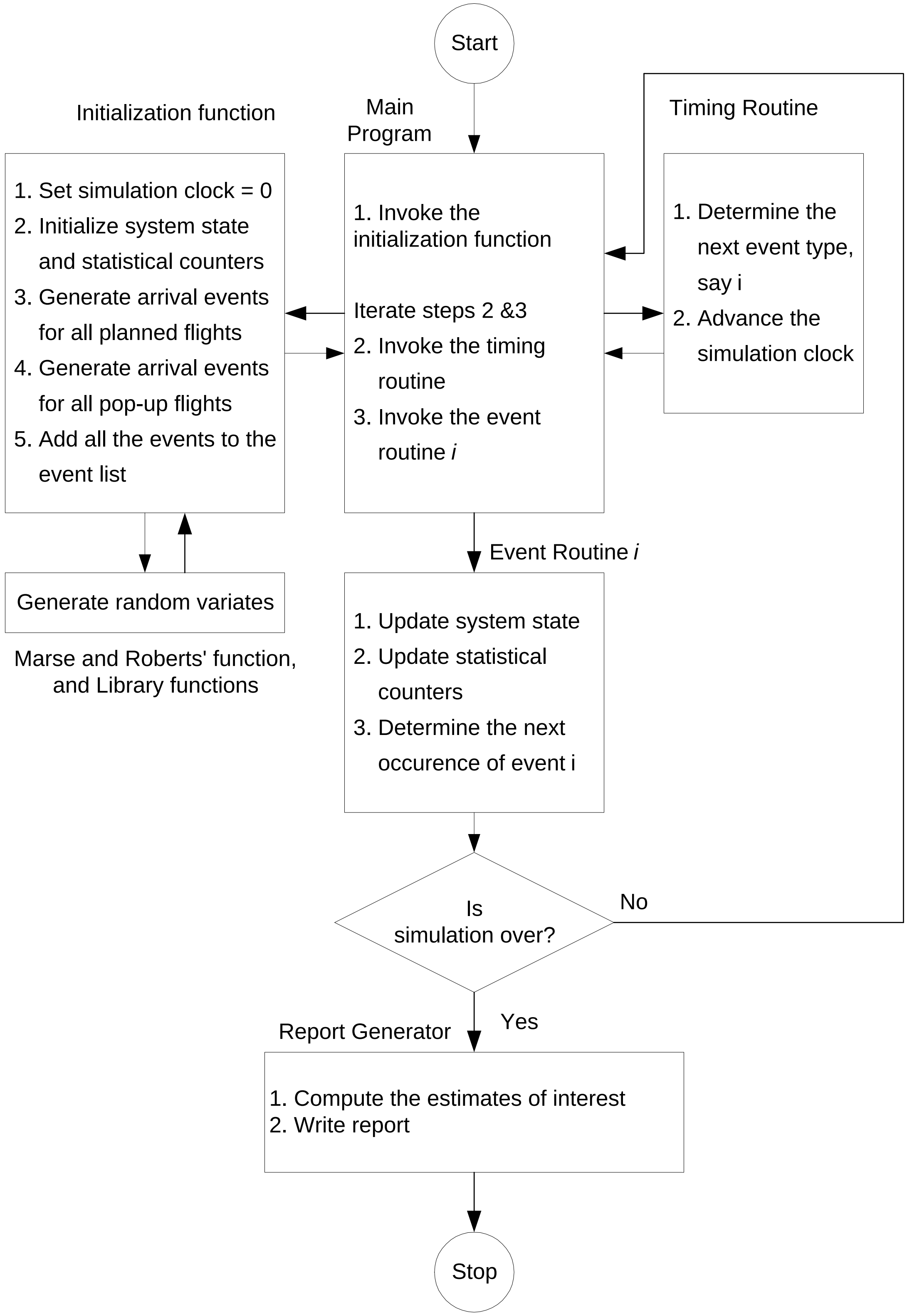

Figure 3.2: Flow Chart for GDP Simulation Model

3.3 Description of Data : Sources and Preparation

3.3.1 Sources (ADL files and Metron Database)

The principal sources of data used here are ADL files and the Metron database - both archived by Metron, Inc. The data collected from these sources provided required ground delay program variables including:

- Airport Acceptance Rates (AAR)
- Planned Airport Arrival Rates (PAAR)
- Flight related information (actual and controlled arrival and departure times, cancellation info, etc)

The process of generating ADL files was explained in section 2.1.1. An ADL file contains all data relevant to running a ground delay program [13]. A particular ADL contains data for only one airport. For our analysis, we made use of ADL files of SFO for the years 1999, 2000 and 2002. The Metron Database refines the ADL data files by extracting only validated data to be archived.

For the purposes of statistical analysis, GDP days had to be carefully selected. The following criteria were used to pick an appropriate set of days

- The GDP should run to completion and must be of at least 4 hours time duration. GDPs canceled before completion don't reflect the dynamics of ideal GDPs. A GDP should last for at least four hours in order that there are strong interactions among various GDP elements.

- No ground stops prior to the start of GDP. Ground stops would affect the hourly arrival rates causing low influx of flights in some hours and heavy influx in other hours.

- Only Morning GDPs at SFO were considered. Since traffic flows would be different at different times of the day, this approach gives a uniform demand profile.

To extract data from the text-based ADL files, scripts in GAWK language - a Unix utility tool - were developed. Metron Database files are stored in MS Excel format and hence, were directly available for our work.

3.3.2 Fitting Probability Distributions

The probability distributions for flight cancellations, arrival of pop-up flights, and the occurrence of flight drift during GDP, was generated from the available data for use as input parameters in the simulation model.

For generating a cancellation probability distribution, we used ADL files of SFO for the year 1999. We hypothesized that cancellations during GDP follow a geometric distribution with a mean $(1-p_{cnx})/(p_{cnx})$ and variance $(1-p_{cnx})/{p_{cnx}}^2$, where p_{cnx} is the probability of cancellation of a flight. The maximum likelihood estimator (MLE) is given by $MLE = 1/(\bar{X}+1)$, where $\bar{X}$ is the mean of the sample population. We conducted Lexis ratio to test our hypothesis. The theoretical and experimental values are shown in Table 3.1.

Table 3.1: Hypothesis Testing for Cancellations Distribution

Tests	Observed Values	Theoretical Values
Lexis Ratio τ	21.90	19.40

In the above table, Lexis Ratio $\tau = S^2/(\bar{X})$, where $\bar{X}$ is the mean and S^2 is the variance of the sample population. Clearly, $\tau \geq 1$ indicates that the distribution is geometric in nature (if $\tau < 1$, it is binomial distribution, and if $\tau = 1$, it is poisson distribution [15]). Further, the observed values closely match the theoretical values of the Lexis Ratio test. Hence, we accept the hypothesis that cancellations follow a geometric distribution. The probability distribution for flight cancellations during GDP is shown in Figure 3.3

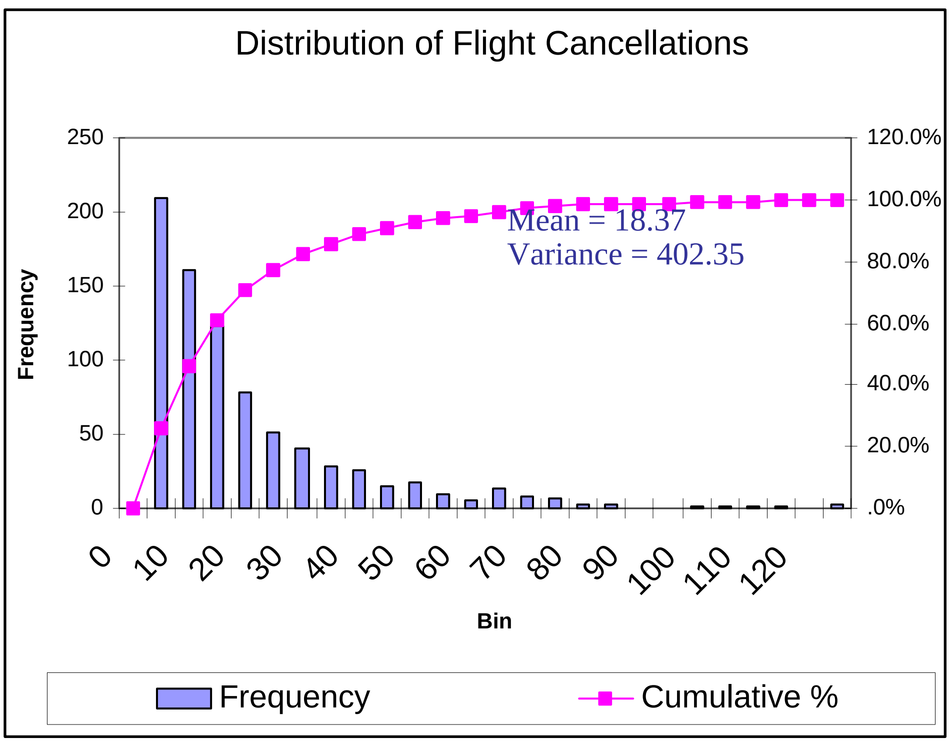

Figure 3.3: Probability Distribution for Flight Cancellations During GDP

The frequency distribution of pop-up flights is shown in figure 2.5. However, pop-ups were highly variable on a hourly basis and hence, the probability distribution of pop-ups by each hour of a GDP is assumed to be exponential with mean pop-up rate λ of about 3 to 5 flights per hour [11].

Empirical distributions for ground drift and enroute drift were generated as the tests indicated that they do not follow any theoretical distributions. The ADL data files for SFO for the year 2002 were used in analyzing the distributions for drift. Figures 3.4 and 3.5 represent the empirical distributions for the flight drift. It can be seen that, the distribution for ground drift is slightly skewed to

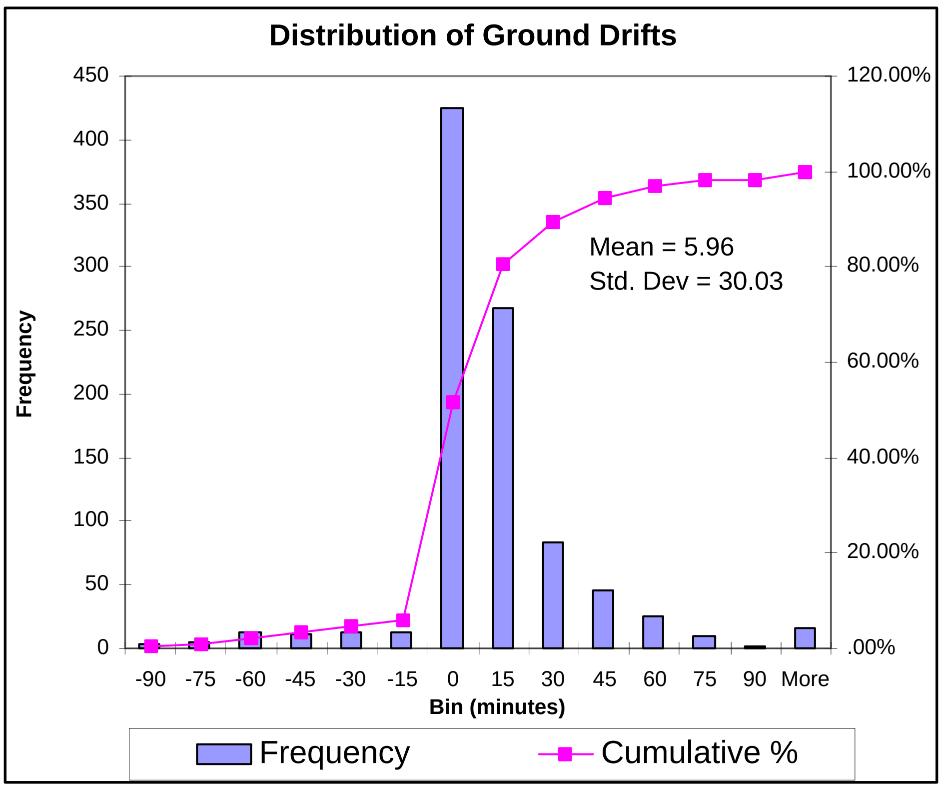

Figure 3.4: Relative Frequency Distribution for Ground Drifts During GDP.

the right meaning that, on an average, flights were departing later than their controlled departure times. It can also be seen that flights are departing as early as 60 minutes prior to the controlled departure times, and as late as 90 minutes after the controlled departure times.

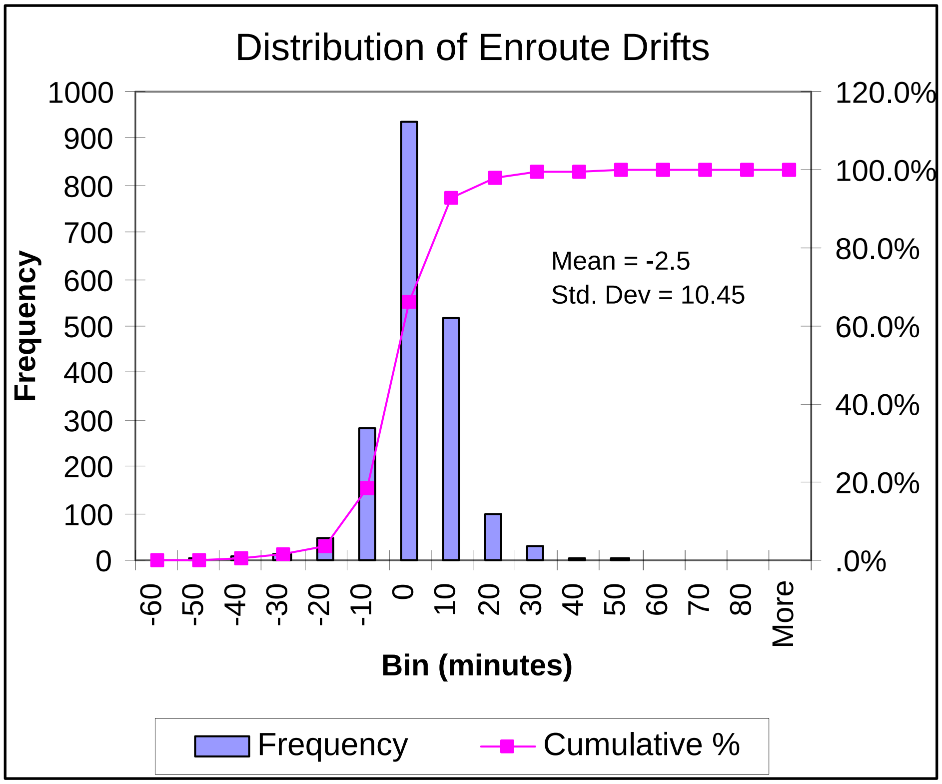

Figure 3.5: Relative Frequency Distribution for Enroute Drifts During GDP

From figure 3.5 it seems that flights take less enroute time than expected as the distribution has a mean of -2.5 minutes. The standard deviation of 10.45 minutes also shows that the enroute drift is not as highly variable as ground drifts, and that they mostly concentrate around the time window [-10,10]. Thus, the effect of ground drift is relatively more profound than that of enroute drift.

Chapter 4

Results

4.1 SMIO Model

The required inputs for the SMIO model are the predicted airport capacities (AARs) for each hour of a GDP, the expected number of unutilized slots (ϵ) in the GDP, the probabilities of flight cancellations (P_{cnx}) and pop-up arrivals (P_{pop}), the maximum allowable airborne queue size in any period ($MaxQ$) for each hour, and finally the duration of the GDP (T). The outputs from the model are the expected airborne holding (ABH) size and the optimal PAARs for each hour of a GDP that optimize the airborne holding size for the given GDP scenario.

In the next two sub-sections we present and analyze results of SMIO model applied to some of the most common GDP scenarios at San Francisco (SFO) Airport.

4.1.1 Optimal PAAR Structures

Airport Capacity Distributions (ACDs) at SFO airport based on different seasons of a year were derived by Inniss [14]. Table 4.1 shows three most common ACDs prevalent at SFO.

Table 4.1: GDP ACDs for SFO Airport

Hours of Reduced Capacity	**ACD**
1	30 45 45 45 45 45
2	30 30 45 45 45 45
6	30 30 30 30 30 30

Our model (version 2 of linearized SMIO model) is tested for the above three scenarios using a Sun Microsystems machine, with SunOS 5.7 version, virtual memory of 1.1GB, and RAM of 128MB. The results are shown in Table 4.2, Table 4.3, and Table 4.4. For testing purposes, the common inputs for all the scenarios are: $P_{cnx} = 0.05$, mean pop-up rate $(N_{pop} * P_{pop} = 20 * 0.05) = 1$ flight per hour, maximum allowable queue size in all periods$(MaxQ)$ = 12, MinPaar = 20, MaxPaar = 40, and duration of the GDP (T) = 6 hours.

Table 4.2: Output for 6-Hour Reduced ACD at SFO

Epsilon (ϵ)	**ABH Size (flights)**	**Optimal PAARs**
1	16.3319	34 29 30 31 31 31
2	11.3828	34 28 30 31 30 30
3	8.5886	34 27 30 30 30 31
4	6.4084	30 32 28 32 28 30
5	4.9728	30 31 28 30 32 28

Table 4.3: Output for 1-Hour Reduced ACD at SFO

Epsilon (ϵ)	**ABH Size (flights)**	**Optimal PAARs**
1	13.6153	31 47 44 45 45 45
2	8.3301	31 44 46 45 44 45
3	5.6774	30 44 46 44 45 45
4	3.5437	30 44 45 44 45 44
5	2.6379	29 44 45 44 45 44

Table 4.4: Output for 2-Hour Reduced ACD at SFO

Epsilon (ϵ)	ABH Size (flights)	Optimal PAARs
1	13.7760	33 28 45 46 44 46
2	8.7957	31 28 46 44 46 46
3	5.6684	31 28 45 44 45 46
4	3.5309	29 30 44 45 44 45
5	2.6914	30 28 44 44 45 45

The above results consistently show that optimal PAARs follow a "staggered pattern" in almost all of the scenarios tested so far. Specially, at higher airport utilization levels, this staggered pattern is more prominent. This clearly shows that the current policy of GDP planners to set "Flat" or "Uniform" PAAR rates may not be effective to mitigate the effects of demand uncertainties (see Figures 4.1 and 4.2).

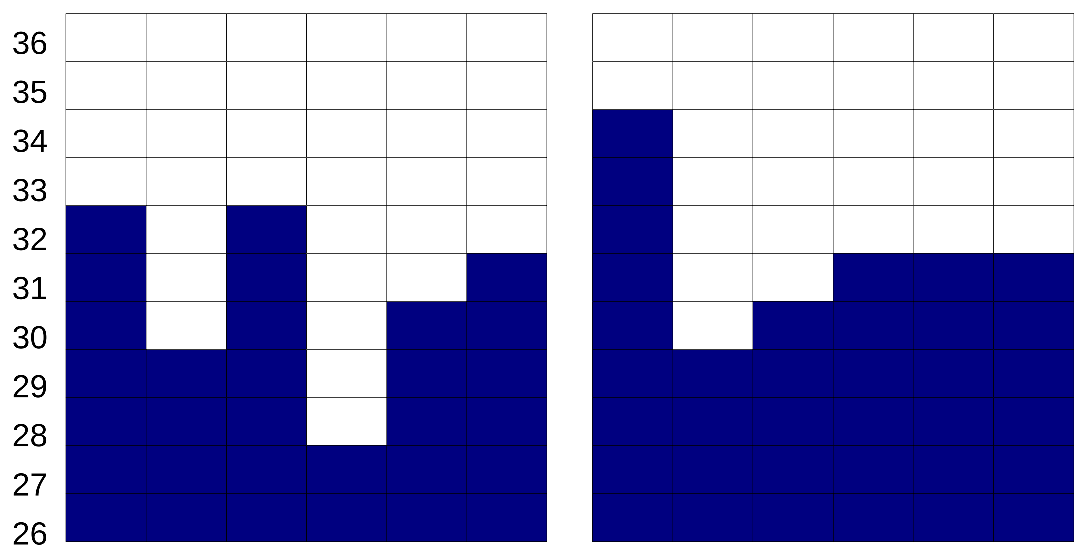

Figure 4.1: Optimal PAARs Generated by SMIO model

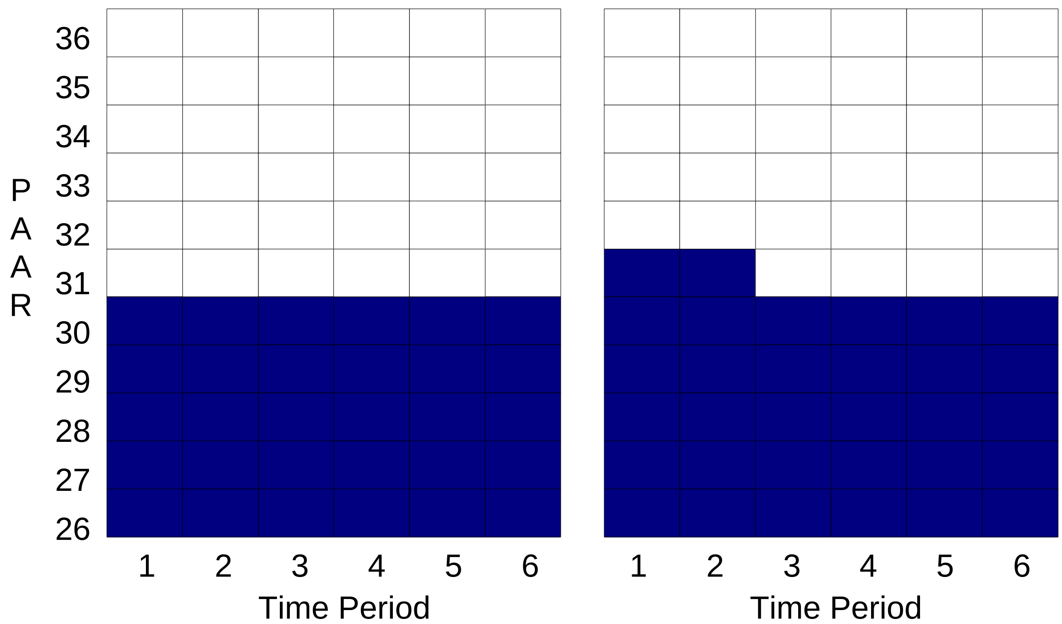

Figure 4.2: Typical PAARs Used by GDP Planners

Typically, the GDP planners assume deterministic demand and set PAARs that are the same as the AARs. However, sometimes, they plan for stochastic elements and set higher PAARs for the first few periods of the GDP to cope with flight cancellations and pop-ups. This buffer, known as Managed Arrival Reservoir (MAR), helps in putting a constant pressure on the airport resources, thus effectively utilizing the airport, but at an expense of higher airborne holding.

Our results, intuitively, show that setting a series of small MARs at periodic intervals during the GDP is better than setting very high MARs at the beginning of GDP. This is the same working principle of inventory-stock models - when inventory falls below the re-order level, then an order for stock is placed. Similarly, when an airport exhausts a MAR, then a new MAR can be ordered. Thus, the best way to mitigate the effects of demand uncertainties is to send flights in periodic bursts at certain intervals.

The results discussed so far are generated by Version 2 of the SMIO model. Version 1 produces almost the same results as Version 2, but with small differences, which are most likely due to inaccuracies in floating point computations. However, it is important to note that the computational burden on version 1 is remarkably higher when compared with that of version 2. To illustrate the computational aspect of the two versions, we tested a 6-hour reduced ACD at SFO airport, with input parameters same as those used for the above tests. The objective function values and the solution times for both the versions is shown in Table 4.5:

Table 4.5: Comparison of Results from Version 1 and Version 2 of SMIO model

	Version 1		Version 2	
ϵ	**ABH Size (flights)**	**Solution Time (sec)**	**ABH Size (flights)**	**Solution Time (sec)**
1	16.3319	7946.95	16.3319	367.11
2	11.3828	3590.4	11.3828	452.31
3	8.3304	2507.07	8.5886	323.13
4	5.43	1485.38	6.4084	298.04
5	4.6184	1562.51	4.9728	215.13

Table 4.5 clearly shows an enormous difference in computational complexity between the two versions. At higher utilizations and for longer GDP programs (say, 8 hours duration), the difference in solution times could be much more. Since, GDP planners need a quick and practical model to work with for effectively planning GDPs, version 2 is more useful than version 1.

4.1.2 Sensitivity Effects of Uncertainties

The effects of flight cancellations and pop-up flights are analyzed in this section. All of our analysis is based on 6-hour reduced ACD, namely, AAR = 30 for each of the GDP hours.

Effects of Flight Cancellations

The effect of varying the probability of flight cancellations (P_{cnx}) on the expected airborne holding (ABH) size, at constant utilization, is studied. The other parameters that are constant for this analysis are: mean pop-up rate ($N_{pop} * P_{pop} = 3 * 0.3$) = 0.9 flight per hour, Maximum allowable queue size for all periods (MaxQ) = 12, MinPaar = 20, MaxPaar = 40, and Duration of the GDP (T) = 6 hours. Tables 4.6, 4.7, 4.8 and 4.9 show all the results.

Table 4.6: Effect of Flight Cancellations, at Constant $\epsilon = 1$

P_{cnx}	**ABH Size (flights)**	**Optimal PAARs**	**Solution Time (sec)**
0.025	10.0458	31 30 30 30 29 30	150.74
0.05	14.0887	32 31 31 31 30 31	222.5
0.075	18.7267	34 32 32 31 31 31	643.9
0.10	21.6711	36 31 33 33 33 32	985.6
0.125	23.8625	36 34 33 34 34 34	2529.11
0.150	27.1095	37 36 34 36 33 36	3812.05
0.175	28.937	39 36 36 36 36 35	1261.66
0.200	30.9402	40 37 38 37 37 38	375.65

Table 4.7: Effect of Flight Cancellations, at Constant $\epsilon = 2$

P_{cnx}	**ABH Size (flights)**	**Optimal PAARs**	**Solution Time (sec)**
0.025	5.763	30 30 30 29 30 29	35.82
0.050	8.3686	32 30 30 30 31 31	63.12
0.075	11.663	32 31 32 32 31 32	319.38
0.100	13.964	33 33 33 31 33 33	205.41
0.125	16.1413	35 34 32 33 34 34	674.26
0.150	18.6471	36 36 32 35 34 36	1328.85
0.175	20.0918	38 35 35 35 36 37	1774.98
0.200	21.9061	39 37 37 36 36 37	1348.55

Table 4.8: Effect of Flight Cancellations, at Constant $\epsilon = 3$

P_{cnx}	**ABH Size (flights)**	**Optimal PAARs**	**Solution Time (sec)**
0.025	3.467	30 29 29 30 29 30	12.49
0.050	5.3549	31 30 30 30 31 30	30.93
0.075	8.7341	32 31 32 32 29 31	151.41
0.100	10.0799	33 33 31 33 31 32	173.58
0.125	11.9837	34 34 32 32 34 34	481.57
0.150	13.9674	36 33 36 32 34 35	909.68
0.175	14.8708	37 35 35 35 35 35	1069.97
0.200	16.7439	38 36 37 36 35 38	1077.3

Table 4.9: Effect of Flight Cancellations, at Constant $\epsilon = 4$

P_{cnx}	**ABH Size (flights)**	**Optimal PAARs**	**Solution Time (sec)**
0.025	2.4963	30 29 29 29 29 30	10.37
0.050	3.6113	30 30 30 30 30 31	10.46
0.075	6.6717	32 31 29 31 32 32	91.43
0.100	7.3258	33 31 31 33 31 33	58.66
0.125	8.9477	34 32 32 34 32 34	278.66
0.150	10.8737	36 32 34 35 32 35	704.05
0.175	11.5283	36 35 35 34 35 35	443.36
0.200	12.9412	37 37 35 36 36 36	942.05

The graph shown in Figure 4.3 is based on the results documented in Tables 4.6, 4.7, 4.8 and 4.9.

From Figure 4.3, it can be observed that, at constant airport utilization and at constant pop-up traffic levels, as uncertainty in flight cancellations increase, the expected airborne queue sizes increase. For lower airport utilization levels, there appears to be a linear correlation between probability of flight cancellations and expected airborne queue size; however, at higher utilization levels, the trend seems to be slightly non-linear. As flight cancellations increase, the variability in flight arrival process also increases; thus, to guarantee that an airport is utilized to the desired level, GDP planners have to set higher PAARs to counter the danger of under-utilization at the expense of airborne holding. Figure 4.4 shows expected airborne queue sizes as a function of airport utilization - airborne queue size decreases as utilization decreases, given that the level of uncertainty

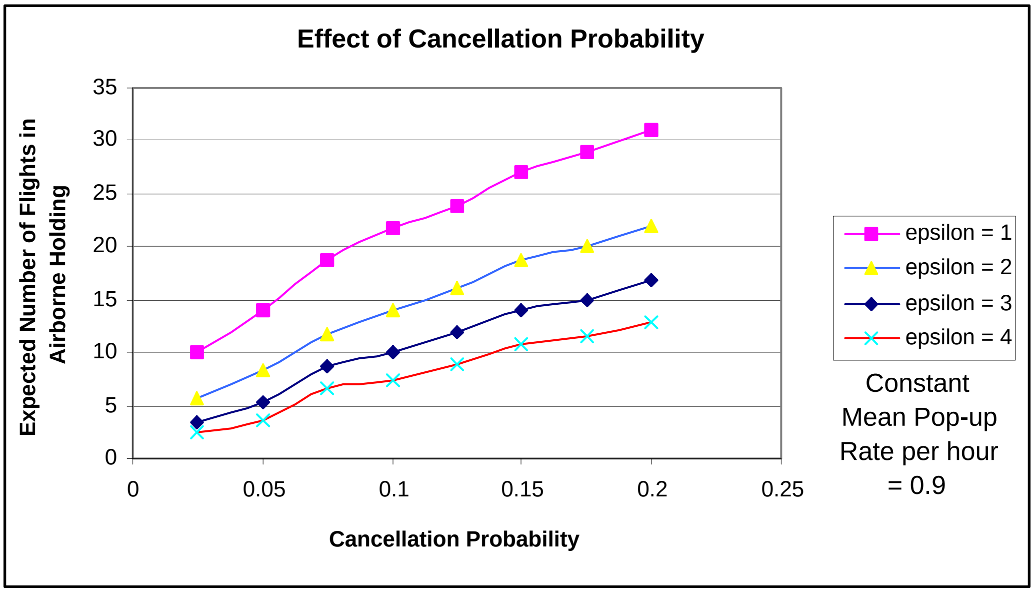

Figure 4.3: Marginal Effects of Flight Cancellations on Expected Airborne Queue Sizes

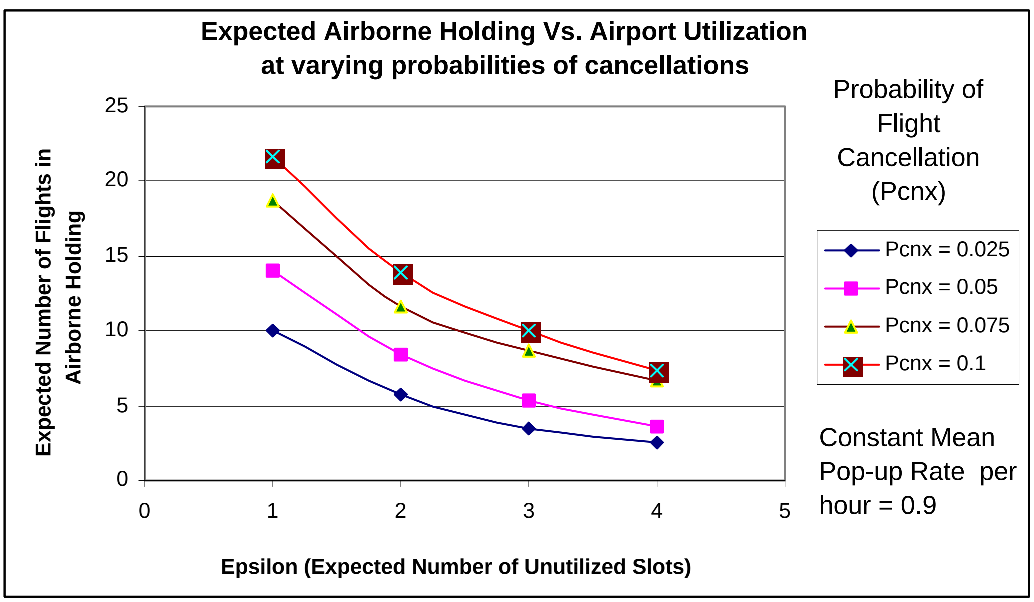

Figure 4.4: Expected Airborne Queue Sizes Versus Airport Utilization at Varying Probability of Cancellations

in flight cancellations is constant.

Effects of Pop-up Flight Arrivals

The effect of varying the mean pop-up arrival rate per hour "MPR" ($P_{pop} * N_{pop}$) on the expected airborne holding (ABH) size, at constant utilization, is studied. The other parameters that are constant for this analysis are: $P_{cnx} = 0.05$, MaxQ = 12, MinPaar = 20, MaxPaar = 40, and Duration of the GDP (T) = 6 hours. Tables 4.10, 4.11, 4.12 and 4.13 show all the results.

Table 4.10: Effect of Pop-up Flight Arrivals, at Constant $\epsilon = 1$

MPR	**ABH Size (flights)**	**Optimal PAARs**	**Solution Time (sec)**
1	16.3319	34 29 30 31 31 31	323.65
2	19.8651	32 29 32 27 30 31	761.14
3	21.081	30 30 28 29 29 29	966.53
4	23.1092	30 28 28 27 28 28	1302.57
5	25.715	29 28 25 28 26 27	1688.95
6	26.9839	27 27 26 25 27 26	1546.5
7	28.369	29 23 25 24 25 25	1889.04
8	28.1956	27 23 24 23 24 24	411.36

Table 4.11: Effect of Pop-up Flight Arrivals, at Constant $\epsilon = 2$

MPR	**ABH Size (flights)**	**Optimal PAARs**	**Solution Time (sec)**
1	11.3828	34 28 30 31 30 30	475.93
2	14.1223	32 27 32 28 29 30	449.52
3	14.7193	30 29 27 30 27 29	566.91
4	15.5648	29 27 28 27 28 27	451.65
5	18.0854	28 26 28 25 27 26	812.02
6	18.4903	27 25 26 25 25 27	670.01
7	19.9188	25 25 25 24 25 25	232.85
8	20.6673	27 21 24 23 23 24	317.2

Table 4.12: Effect of Pop-up Flight Arrivals, at Constant $\epsilon = 3$

MPR	**ABH Size (flights)**	**Optimal PAARs**	**Solution Time (sec)**
1	8.5886	34 27 30 30 30 31	359.07
2	11.1667	32 27 28 32 28 29	246.58
3	11.1331	30 27 27 29 29 29	229.87
4	11.5746	28 27 27 27 28 28	235.55
5	13.9839	28 25 28 24 27 26	344.05
6	14.0155	27 25 24 25 26 25	523.03
7	15.0396	25 25 23 25 24 24	76.15
8	14.9237	24 23 23 23 24 23	107.12

Table 4.13: Effect of Pop-up Flight Arrivals, at Constant $\epsilon = 4$

MPR	ABH Size (flights)	Optimal PAARs	Solution Time (sec)
1	6.4084	30 32 28 32 28 30	257.42
2	8.2663	32 28 28 28 30 28	84.62
3	8.7412	30 27 27 29 27 29	344.93
4	8.6639	28 26 27 27 27 28	117.42
5	11.1591	28 25 25 28 24 26	304.11
6	11.0204	27 24 25 24 25 25	203.63
7	11.959	25 23 25 23 25 23	81.96
8	11.6084	24 22 23 23 23 23	103.61

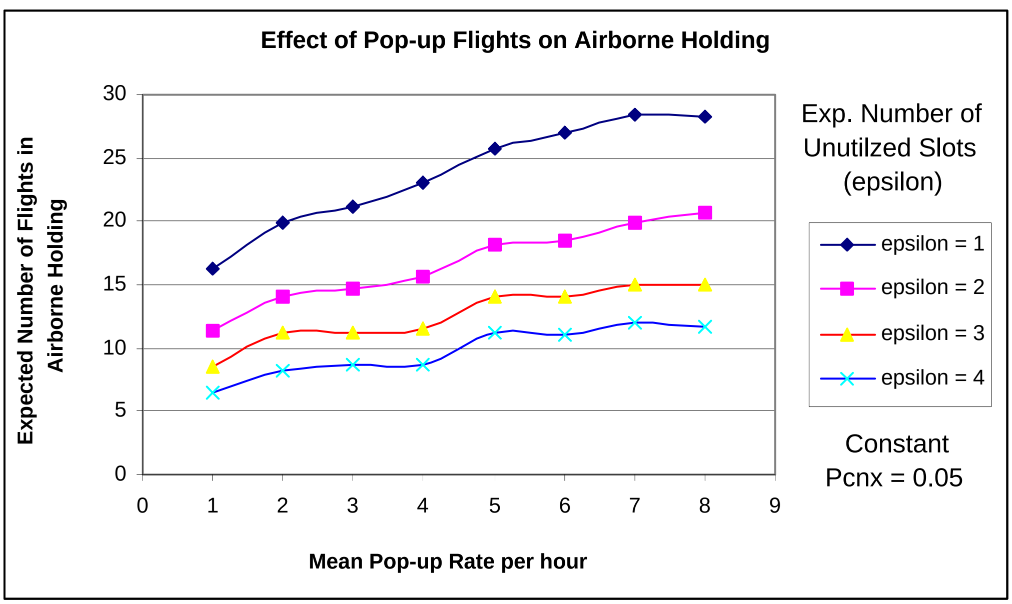

Figure 4.5: Marginal Effects of Pop-up Flight Arrivals on Expected Airborne Queue Sizes

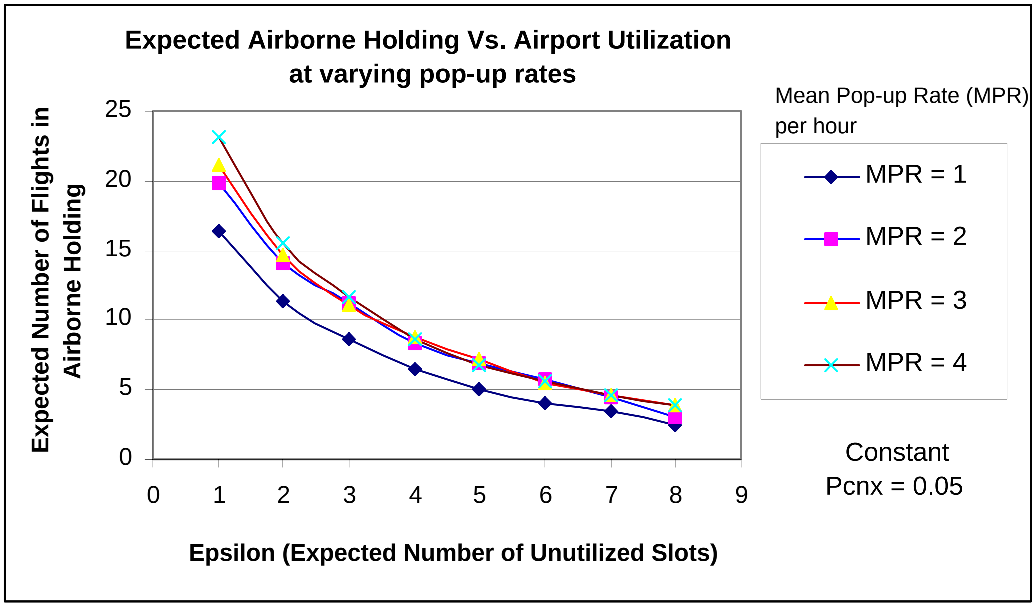

Figure 4.6: Expected Airborne Queue Sizes Versus Airport Utilization at Varying Pop-up Arrival Rates per Hour

Figure 4.5 illustrates that, at constant airport utilization and at constant uncertainty of flight cancellations, as uncertainty in arrival of pop-up flights increases, the expected airborne queue size increases. There appears to be a linear correlation between mean pop-up arrival rates and expected airborne queue size. Pop-up flights add up to the GDP arrival demand; thus, causing excess airborne holding. Figure 4.6 strengthens the earlier observation that expected airborne queue size increases as a function of airport utilization.

4.2 Simulation Model

The Simulation Model incorporates all three forms of demand uncertainties, namely, flight cancellations, arrival of pop-up flights and flight drift. The performance measures are airborne delay, ground delay and airport utilization.

4.2.1 Pareto Optimal PAARs

We have seen that the SMIO model generates some peculiar PAAR patterns that oppose the current policy of setting flat PAARs. To validate that the optimal PAARs indeed follow these "Staircase" kind of patterns, we employed our simulation model to derive the Pareto Frontier of PAARs that optimize airport utilization and airborne delay.

Definition 1 (Pareto Optimality) *A state* A *(a set of object parameters) is said to be Pareto optimal, if there is no other state* B *dominating A with respect to a set of objective functions. A state A dominates a state B , if A is better than B in at least one objective function and not worse with respect to all other objective functions.*

Definition 2 (Object Parameters) *The input parameters that minimize or maximize the objective function are called Object Parameters.*

In our case, the object parameters are the input parameters of the simulation model, namely, PAARs, airport arrival capacity, distributions for flight cancellations, pop-up arrivals and flight drift. The objective functions are 'airborne delay (minimize)', 'ground delay (minimize)', and 'airport utilization (maximize)' functions. Since the SMIO model generates optimal PAARs based on two explicit criteria - airborne delay and airport utilization, the pareto frontier of PAARs will also be constructed with only two objective measures by factoring out ground delay so that the comparison of results is valid.

The pareto optimal PAARs are generated for the following GDP scenario. To model a realistic scenario, we chose average values for the parameters that are involved.

- Airport Capacity = 30 per hour, and GDP duration = 6 hours.
- PAARs are varied from 28 to 34 for each hour of a GDP
- Cancellations follow a geometric distribution with probability $P_{cnx} = 0.052$.
- Arrival times of Pop-ups is exponential with mean rate = 3 per hour.
- Flight Drift follows an empirical distribution (derived in Section 3.3.2).

The simulation model was run for 10 replications and a 95% confidence interval was considered for the output data. The final pareto curve is shown in Figure 4.7. From this figure, two important conclusions can be made:

- Optimal Paars indeed follow a staggered pattern. At lower utilizations, the staggered pattern may not be very noticeable, but at high utilizations (≥ 0.88), the rise and fall of PAARs is very visible. This is the same observation we noted from the results of the SMIO model. In addition, it seems that the first hour of the GDP should be more heavily loaded. Hence, we believe that the most effective PAARs should indeed follow staircase patterns that have high first hour loads.
- Airport Utilization and Airborne Holding share a non-linear positive correlation, given that the level of demand uncertainty remains constant. Again,

we noted the same observation earlier from our results of the SMIO model. A significant observation is that, to achieve very high utilization (close to 1.0), the airborne holding incurred by the flights can become very large. For example, for an increase in utilization from 0.97 (PAAR scenario 12) to 0.98 (PAAR scenario 13), the increase in airborne holding is 4 minutes per flight.

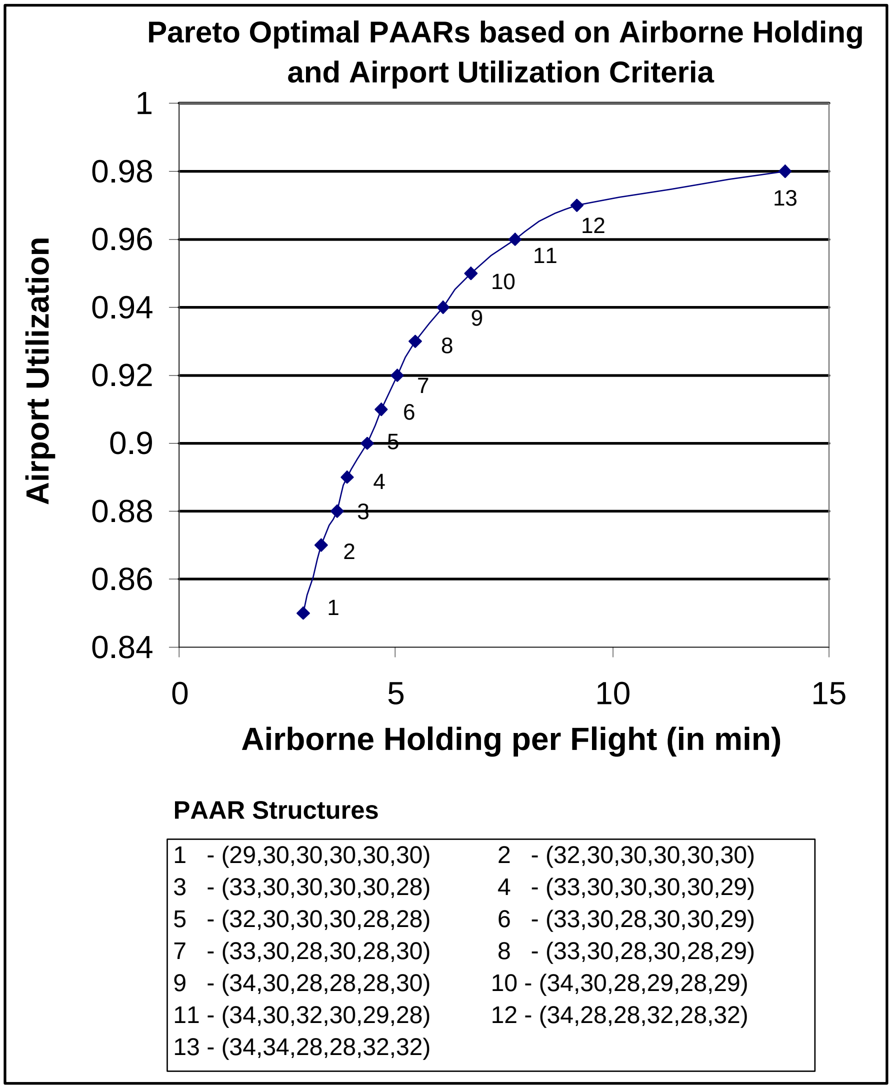

Figure 4.7: Pareto Optimal PAARs Based on Airport Utilization and Airborne Delay Criteria

To understand how "Flat" PAARs deviate from the "Staircase" or "Pareto Optimal" PAARs with respect to performance, we constructed a scatter plot of some of the Flat PAAR scenarios which is shown in Figure 4.8.

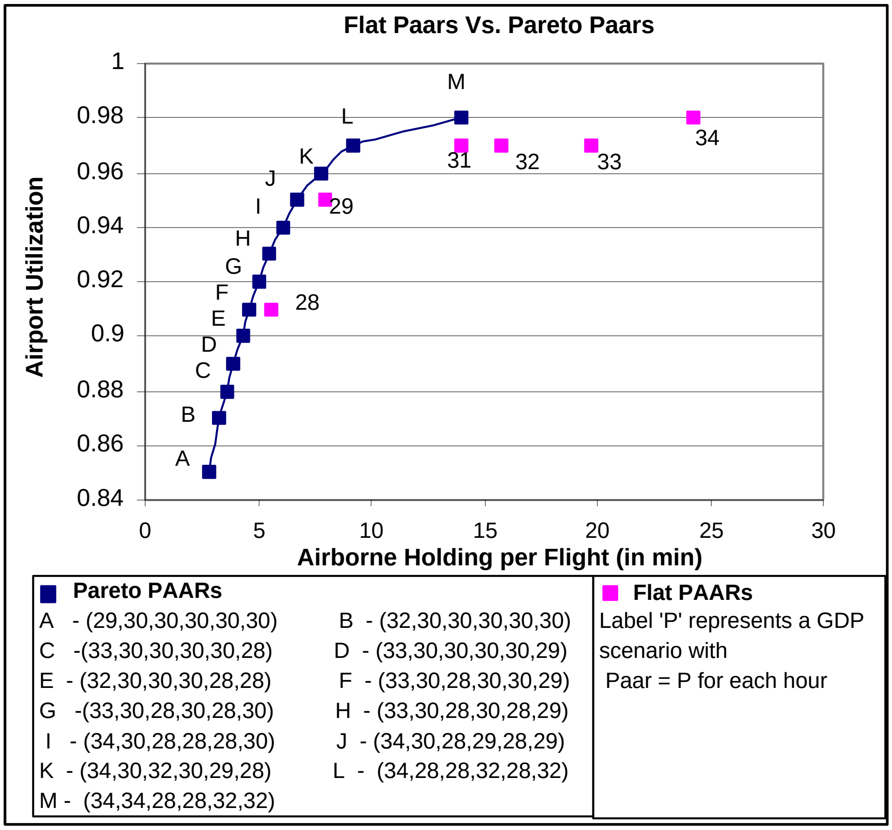

Figure 4.8: Performance comparison : Flat PAARs Vs. Pareto PAARs

From the figure, it can be observed that pareto optimal PAARs definitely show improvement in performance when compared with flat PAARs, as all flat PAARs lie below the pareto curve. At higher utilization, there is significant difference in the performances delivered by staircase PAARs and Flat PAARs.

4.2.2 Sensitivity Effects of Uncertainties

In this section, we study the effect of varying certain input parameters. The simulation model was replicated 100 times and a 95% confidence interval was considered for all the results shown in this section.

PAARs

Figure 4.9 and Figure 4.10 show the effect of varying PAARs on the output parameters, namely, airport utilization and delays (ground and airborne). The GDP scenario has constant arrival capacity of 30 per hour, and it lasts for 6 hours. Other input parameters are same as in the above section. Clearly, from Figures 4.9 and 4.10, it is visible that setting high PAARs would deliver a high airport utilization, though at the expense of airborne delays. The GDP planners must carefully weigh the trade-offs between ground delays and airborne delays when setting PAARs for a GDP.

Flight Cancellations

At this point of analysis, it would be interesting to quantify the effects of various uncertainties, including flight cancellations, in terms of dollars. With this idea in mind, a cost function is formulated as shown below:

$$\text{Operating Cost} = \text{Airport Under-utilization Cost} + \text{Airborne delay Cost}$$

i.e.,

$$C_f = C_u * X_u + C_{abh} * X_{abh}$$

where,

C_u is the cost of one open (wasted) slot during a GDP,

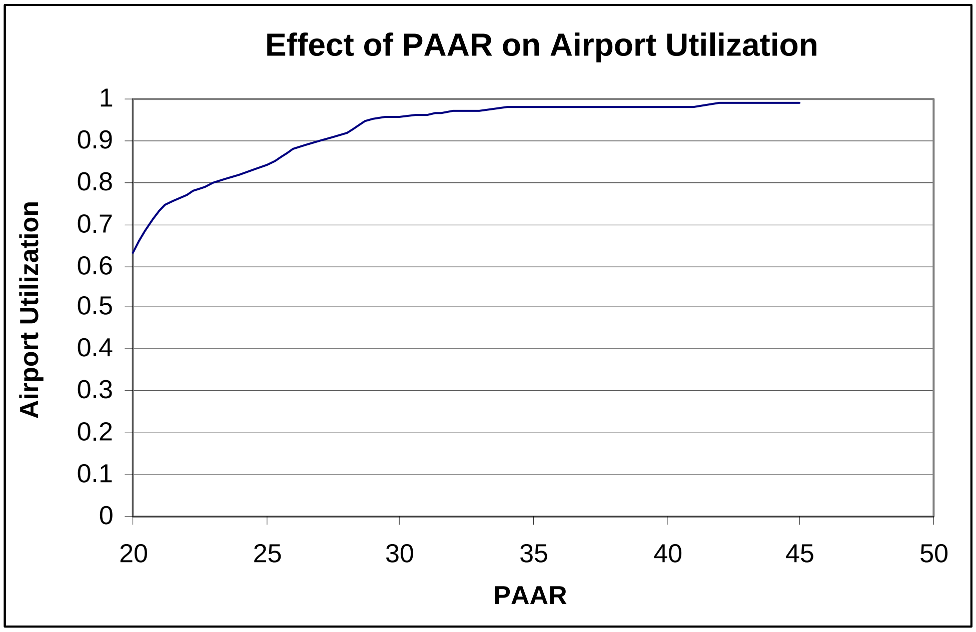

Figure 4.9: Effect of PAARs on Airport Utilization

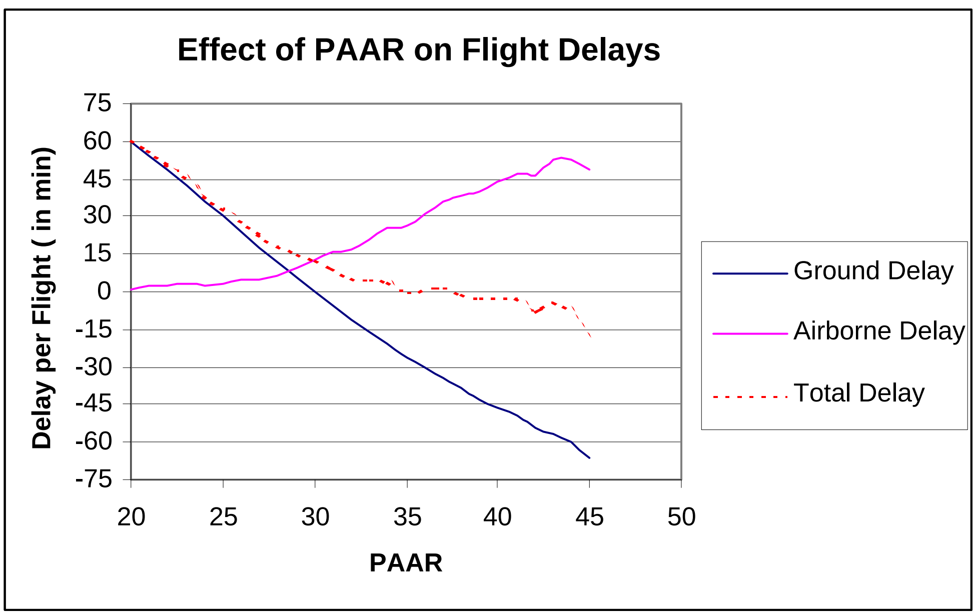

Figure 4.10: Effect of PAARs on Flight Delays

C_{abh} is the cost incurred by an airline due to airborne holding per minute,

X_u is the total number of open slots during a GDP, and

X_{abh} is the total airborne holding (abh) in minutes (abh per flight * total flights) during a GDP.

We need to justify why under-utilization cost is incurred by airlines. The reasoning is this. Arrival slots, during a GDP, are wasted mostly due to occurrence of cancellations, drifts and others; however, cancellations are the most dominant of all with respect to costs and frequency of occurrence. Hence, it would be equivalent to say that, during a GDP, one last-minute flight cancellation can lead to one wasted slot (this is mostly true in case of Timed-Out (TO) cancellations, as each TO cancellation potentially wastes an airline's slot (see Section 2.1.1)). Thus, airport under-utilization cost is an approximation of the cancellation costs incurred by airlines.

One more important thing to note about the cost function is that it excludes ground holding costs. We can argue that in a GDP scenario, where an airport constantly operates at a specific PAAR, the ground delay is always constant. What we mean by this is that, suppose that we pick a pareto optimal PAAR, say , scenario 12 in Figure 4.7, then if we vary flight cancellations or pop-up mean rates of arrivals, the only parameters that are affected are airport utilization and airborne delay of flights. Hence, ground delay costs need not be a criterion in our cost function.

Since the simulation model does not output the number of slots that are being wasted during a GDP, we use the following transformation to determine, approximately, the total number of vacant slots during a GDP:

$$X_{cnx} = (1 - \rho) * N_{hours} * \sigma$$

where,

ρ is the airport utilization during a GDP,

N_{hours} is the number of hours for which a GDP lasted, and

σ is the number of available slots per hour during a GDP

X_{abh} is the total airborne holding (abh) in minutes (abh per flight * total flights) during a GDP.

A reasonable estimate of cost incurred by an airline per cancelled flight (C_{cnx}) is given by \$20,000 [1]. The airborne delay cost per minute (C_{abh}) is estimated to be \$47.64 [2].

To study the sensitivity effects of flight cancellations on the overall operating costs of the airlines, we selected scenario 12 with PAAR vector (34,28,28,32,28,32) from the pareto curve (see Figure 4.7). Assuming that an airport is operating at the chosen PAAR scenario, we measure the costs incurred by airlines as a function of flight cancellations during a GDP. Table 4.14 gives the sensitivity results and Figure 4.11 shows the plot of probability of flight cancellation P_{cnx} vs. operating cost of airlines per GDP.

It is clear from Figure 4.11 that at a constant PAAR setting, if uncertainty in flight cancellations increases, then the airport utilization could drastically decrease, thereby creating vacant slots and also, high airborne holding per flight. Thus, the costs incurred by the airlines increase proportionally with an increase in probability of flight cancellation.

Table 4.14: Effect of Flight Cancellations on Operating Costs of Airlines

P_{cnx}	ρ	X_u	X_{abh}	**Costs**
0.025	0.97	5.4	1966.504	$201,684
0.050	0.96	7.2	1584.824	$219,501
0.075	0.95	9	1241.688	$239,154
0.1	0.92	14.4	1085.69	$339,722
0.125	0.91	16.2	912.296	$367,462
0.15	0.88	21.6	767.725	$468,574
0.175	0.87	23.4	545.528	$493,989
0.2	0.84	28.8	461.04	$597,964

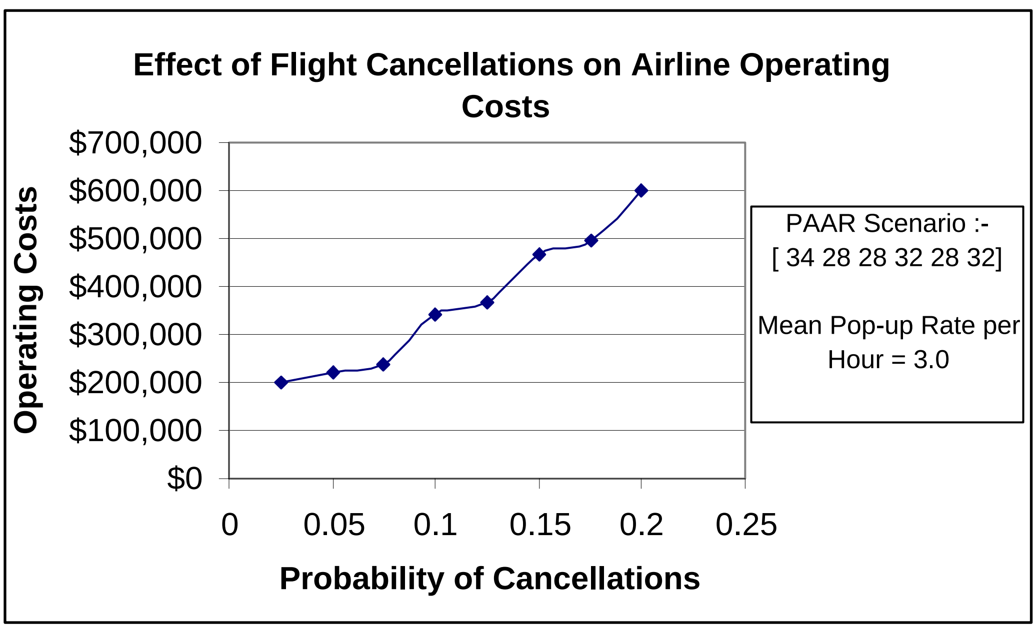

Figure 4.11: Effect of Flight Cancellations on Overall Airline Costs.

Pop-up Flight Arrivals

We use the same PAAR scenario as used in the above section. The Mean Pop-up Rate (MPR) per hour is varied at constant drift and constant probability of cancellation, and the costs are recorded as shown in Table 4.15. Figure 4.12 shows the graph of MPR versus costs.

Table 4.15: Effect of Pop-up Flight Arrivals on Operating Costs of Airlines

MPR	ρ	X_u	X_{abh}	**costs**
3	0.97	5.4	1457.53	$177,437
4	0.97	5.4	2003.06	$203,426
5	0.98	3.6	2600.43	$195,885
6	0.98	3.6	3051.60	$217,378
7	0.98	3.6	3522.61	$239,817
8	0.98	3.6	3979.53	$261,585

Pop-up flights also seem to have the same kind of impact on the costs as the flight cancellations. However, the effect of pop-ups is quite linear as opposed to slight non-linear correlation of flight cancellations on the overall airline costs. Pop-ups displace the actual flights in the arrival sequence; hence, the utilization may not suffer much, but the airborne holding incurred by the regular flights increases significantly.

Flight Drift

To study the sensitivity effects of drift, we make certain assumptions. Firstly, we factor out the enroute drift as it is insignificant with respect to ground drift; hence, we vary only the ground drift to study sensitivity of drift. Secondly, we

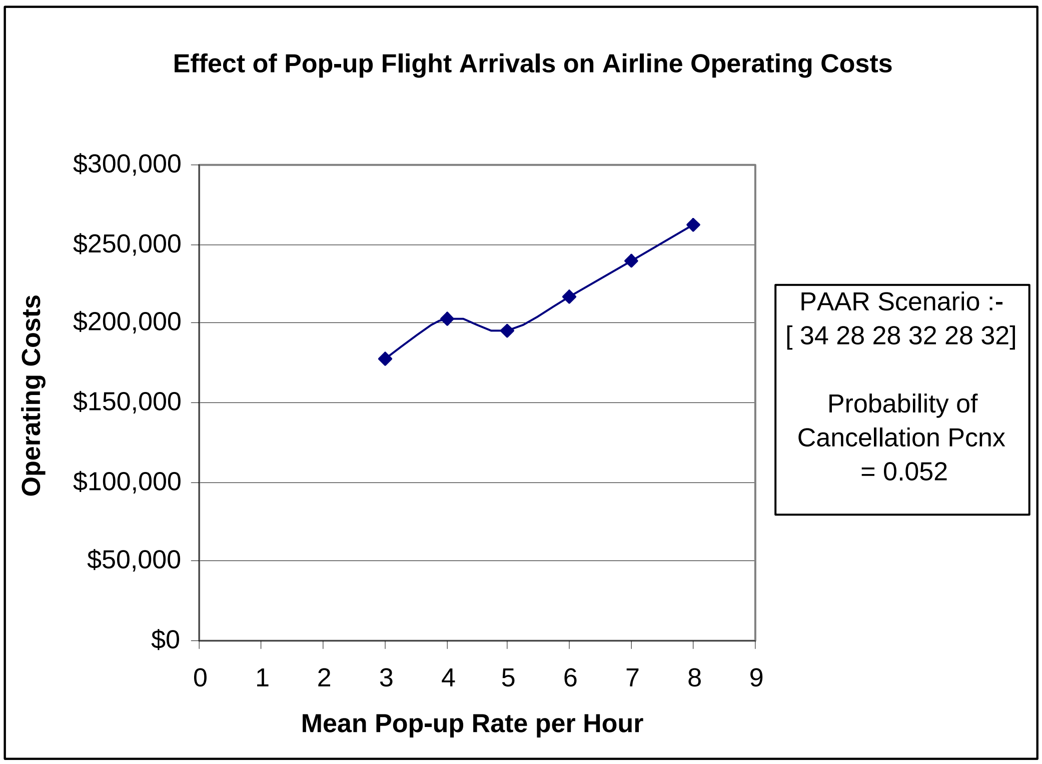

Figure 4.12: Effect of Pop-up Flight Arrivals on Overall Airline Costs

need some way to control the drift variable to study its sensitivity. Recall that we derived an empirical distribution of drift in Section 3.3.2; all our results generated so far made use of the same distribution. The disadvantage of using an empirical distribution is that it cannot be controlled by its mean and standard deviation, but controlled by empirical probabilities (histogram relative frequencies). Hence, to vary the distribution, we use the following approach shown in Figure 4.13.

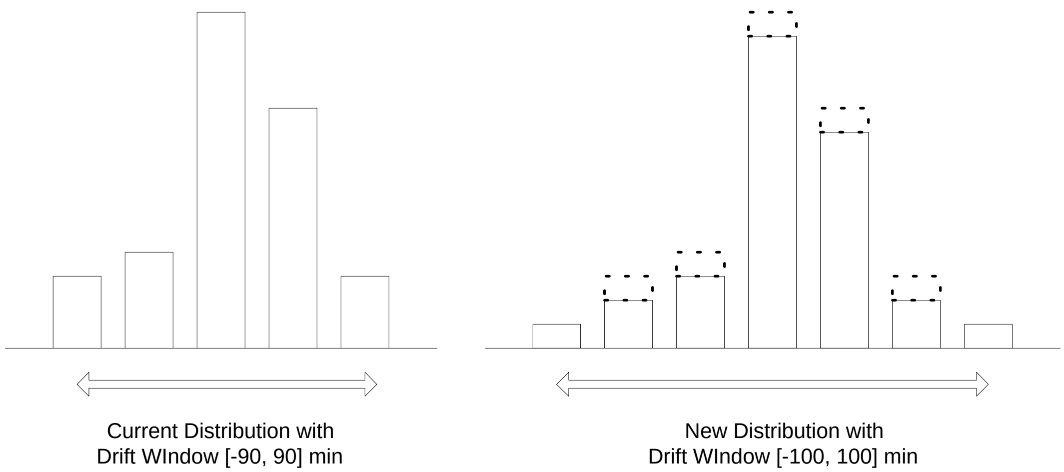

Figure 4.13: Generation of New Distribution From a Given Empirical Distribution For Drift

Figure 4.13 illustrates how we stretch an empirical distribution (with drift window [-τ, τ]) to generate a new distribution (with drift window [-($\tau + \delta$), ($\tau + \delta$)]). We slice off an equal percentage from all the histograms within the original distribution, and add new histograms of width δ each. Now, we assign percentage of drift to the newly created histograms in such a way that the mean of the whole distribution remains the same. The assumption here is that the drift within the window [-($\tau + \delta$), τ] and [τ, ($\tau + \delta$)] follows a *uniform distribution*. One more assumption here is that the cumulative percentage of flights with drift in the window [-($\tau+\delta$), τ] and [τ, ($\tau+\delta$)] will be 10% of all the flights that drift during a GDP, which means that the probability of any flight drifting into this window is 0.1.

Using the above assumptions, we performed sensitivity analysis of drift. Three different distributions are being generated from the original distribution (with mean 4.95 minutes and standard deviation 21.32 minutes). The results are shown in Table 4.16. The plot of cost vs. drift window is shown in Figure Figure 4.14.

Table 4.16: Effect of Flight Drift on Operating Costs of Airlines

Drift Window	**Std. Deviation**	ρ	X_u	X_{abh}	**costs**
[-120,120]	38.7	0.95	9	1480.9	$250,550
[-110,110]	37.2	0.95	9	1385.1	$245,987
[-100,100]	35.9	0.95	9	1272.4	$240,617
[-90,90]	21.32	0.96	7.2	1680.5	$224,057

Clearly, the results indicate that drift can result in considerable operating losses for the airlines. Each flight that drifts alters the planned arrival sequence, creating unnecessary airborne holding and wastage of slots. Controlling the variance in flight arrivals will help reduce the operating costs of the airlines. In Figure 4.14, by decreasing the drift by 10 minutes per flight on an average (from 35.9 to 21.32 minutes), the airlines can save approximately $16,000 per GDP. Hence, small drift windows will curtail the amount of uncertainty in flight arrivals and produce a smooth and predictable arrival sequence, that can be effectively controlled.

From the sensitivity analysis of flight cancellations, pop-up flights and drift,

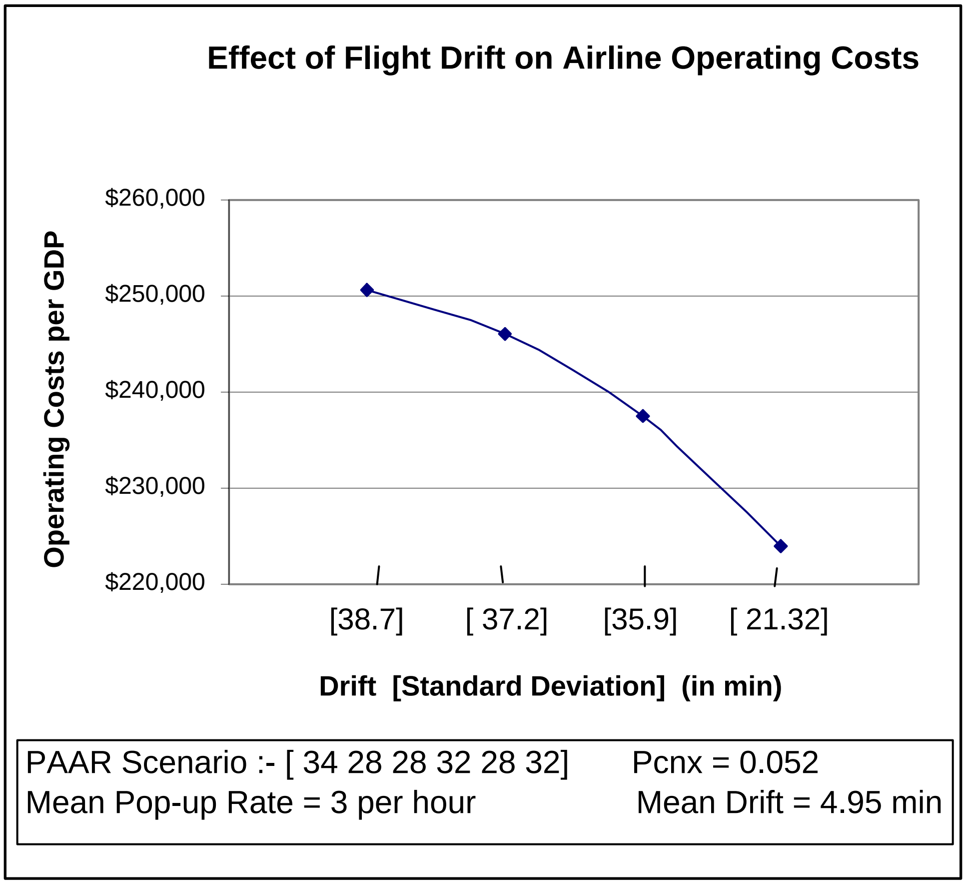

Figure 4.14: Effect of Drift on Airline Operating Costs.

it can be concluded that all of them affect the airline economics in a significant way. Thus, minimizing the uncertainty in demand would help reduce the airline costs by a fair margin.

Chapter 5

Conclusions

5.1 Main Contributions

The main purpose of our thesis is to model and analyze demand uncertainty in the context of Ground Delay Programs (GDPs). To this extent, we have developed two models - the Stochastic Mixed-Integer Optimization (SMIO) Model and the Simulation Model. The SMIO model produced surprising results as they indicated that the Planned Airport Arrival Rates (PAARs) that optimize the performance of a GDP exhibit non-conventional patterns, that are probably unknown to GDP planners until today. We believe that this is our biggest contribution. The results from the Simulation Model also reinforced the PAAR structures exhibited by the SMIO model. Hence, we suggest the GDP planners to rethink their policy of setting PAARs during a GDP - "Staircase" PAARs for "Flat" PAARs.

For a specific GDP scenario, we applied our simulation model and developed a pareto optimal curve with airborne holding and airport utilization as performance measures. This pareto curve could be employed by the GDP planners as it provides them with insight into the structure of optimal PAARs that achieve equilibrium between airborne holding and airport utilization, under the given GDP conditions. Similar such pareto curves for a variety of GDP scenarios can expedite the decision-making process involved in planning a GDP.

Both models can serve as good strategic tools to be used by airlines in optimizing their objectives. Specially, the simulation model provides estimates as well as sensitivity of demand uncertainty cost during GDPs. The SMIO model is special in its own way because it accommodates most of the current procedures and paradigms of air traffic management developed by Collaborative Decision Making (CDM) working group. The output from this model is the number of flights that should be ordered to arrive at the destination (affected) airport so that the airborne holding is minimum at desired airport utilization. Once the aggregate number of flights that needs to be sent to the GDP airport is determined, the CDM procedures like Ration-By-Schedule (RBS) and Compression are then applied to determine the individual flights that should be assigned ground delay. Thus, this model has a potential to evolve as a CDM decision-support tool for the common use of all CDM members.

5.2 Directions for Future Research

Both of the models that we developed in this thesis produced significant results. However, these models can be further enhanced to meet more realistic GDP situations. Specifically, the SMIO model can be extended by incorporating flight drift and by devising a better formulation to lighten the computational burden.

The GDP process is becoming more and more dynamic due to the rapid development of tools and technologies from within the CDM group. For example, GDP revisions and the use of Ration-By-Schedule (RBS++) and Compression algorithms make GDPs inherently a dynamic process. A dynamic programming model may be more suitable for approximating a GDP than a static or a static stochastic model. Hence, a possible extension of our models can be in this area.

Finally, it would be considered a major breakthrough if an optimization model were to be constructed that incorporates both stochastic demands and stochastic capacities in a GDP context.

Bibliography

[1] Personal Communication with Experienced Personnel at Metron Inc., 2002.

[2] Aircraft Operating Costs by Stage of Flight and Associated Air Traffic Control Delay Costs Based on a Typical Flight. Air Transport Association, Washington, D.C., 2000.

[3] G. Andreatta and G. Romanin-Jacur. Aircraft Flow Management Under Congestion. *Transportation Science*, 21(4):249–253, November 1987.

[4] M. Ball, C. Chen, R. Hoffman, and T. Vossen. *Collaborative Decision Making in Air Traffic Management: Current and Future Research Directions*, pages 17–30. In Bianco et al. [8], 2001. (a preprint appeared in Proceedings of ATM'99, Workshop on Advanced Technologies and Their Impact on Air Traffic Management in the 21st Century, 1999).

[5] M. Ball, R. Hoffman, W. Hall, and A. Muharremoglu. Collaborative Decision Making in Air Traffic Management: A Preliminary Assessment. Technical Report RR-99-3, NEXTOR, September 2001.

[6] M. Ball, R. Hoffman, D. Knorr, J. Wetherly, and M. Wambsganss. *Assessing the Benefits Of Collaborative Decision Making In Air Traffic Management*, pages 239–250. American Institute of Aeronautics and Astronautics,

Inc., Reston, VA., USA, 2001. (a preprint appeared in Proceedings of 3rd USA/Europe air Traffic Management R&D Seminar, 2000).

[7] M. Ball, R. Hoffman, A. Odoni, and R. Rifkin. The Static Stochastic Ground Holding Problem with Aggregate Demands. Technical Report RR-99-1, NEXTOR, UC Berkeley, Jan 1999.

[8] L. Bianco, P.D. Olmo, and A.R. Odoni, editors. *New Concepts and Methods in Air Traffic Management*. Springer-Verlag, Berlin, 1999.

[9] C. Brinton, R. Hoffman, K. Kollmann, N. Shaw, and M. Feldman. The equity practices of collaborative decision making in air traffic management. Technical report, Metron, Inc., June 2000.

[10] D. Gallus, C. Ermatinger, and M. Venanzi. Airport under-delivery project, 2000. Presented at the September 2000 CDM meeting.

[11] R. Hoffman. Pop-up Traffic Study: Phase 1. Technical report, Metron, Inc., September 2001.

[12] R. Hoffman, L. Shishler, K. Howard, M. Klopfenstein, and M. Ball. Report of the Rolling Spike Committee. Technical report, Metron, Inc., Volpe and University of Maryland, October 1999.

[13] K. Howard. ADL Description, Version 5.0, 2001. In a Memorandum Prepared by John A. Volpe National Transportation Systems Center Addressed to the CDM Working Group. Available at http://vivaldi.metsci.com/cdm/archives.html.

[14] T. Inniss. *Stochastic Models for the Estimation of Airport Arrival Capacity Distributions*. PhD thesis, University of Maryland, College Park, MD 20742, December 2000.

[15] A.M. Law and W.D. Kelton, editors. *Simulation Modeling and Analysis*. McGraw-Hill, Inc., USA, 1991.

[16] A. Odoni. *The Flow Management Problem in Air Traffic Control*, pages 269–288. Springer-Verlag, Berlin, 1987.

[17] O. Richetta and A. Odoni. Solving Optimally the Static Ground Holding Policy Problem in Air Traffic Control. *Transportation Science*, 27(3):228–238, August 1993.

[18] M. Terrab. *Ground Holding Strategies for Air Traffic Control*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, February 1990.

[19] H. C. Tijms. *Stochastic Models : An Algorithmic Approach*. John Wiley & Sons, Inc., United Kingdom, 1994.

[20] T. Vossen and M. Ball. Optimization and mediated bartering models for ground delay programs. Available at http://bmgt1-notes.umd.edu/faculty/km/papers.nsf.

[21] M. Wambsganss. *Collaborative Decision Making in Air Traffic Management*, pages 1–15. In Bianco et al. [8], 1999.

[22] L. A. Wolsey. *Integer Programming*. John Wiley & Sons, Inc., New York, USA, 1998.