ABSTRACT

Title of Document: AVIATION CONGESTION MANAGEMENT IMPROVEMENTS IN MODELING THE PREDICTION, MITIGATION, AND EVALUATION OF CONGESTION IN THE NATIONAL AIRSPACE SYSTEM

Kleoniki Vlachou, Doctor of Philosophy, 2014

Directed By: Professor David J. Lovell, Department of Civil and Environmental Engineering
Institute for Systems Research

The air transportation system in the United States is one of the most complex systems in the world. Projections of increasing air traffic demand in conjunction with limited capacity, that is volatile and affected by exogenous random events, represent a major problem in aviation system management. From a management perspective, it is essential to make efficient use of the available resources and to create mechanisms that will help alleviate the problems of the imbalance between demand and capacity. Air traffic delays are always present and the more air traffic increases the more the delays will increase with very unwanted economic impacts. It is of great interest to study them further in order to be able to more effectively mitigate them. A first step would be to try to predict them under various circumstances. A second step would be to develop various mechanisms that will help in reducing delays in different settings.

The scope of this dissertation is to look closer at a threefold approach to the problem of congestion in aviation. The first effort is the prediction of delays and the development of a model that will make these predictions under a wide variety of distributional assumptions. The work presented here is specifically on a continuum approximation using diffusion methods that enables efficient solutions under a wide variety of distributional assumptions. The second part of the work effort presents the design of a parsimonious language of exchange, with accompanying allocation mechanisms that allow carriers and the FAA to work together quickly, in a Collaborative Decision Making environment, to allocate scarce capacity resources and mitigate delays. Finally, because airlines proactively use longer scheduled block times to deal with unexpected delays, the third portion of this dissertation presents the assessment of the monetary benefits due to improvements in predictability as manifested through carriers' scheduled block times.

AVIATION CONGESTION MANAGEMENT
IMPROVEMENTS IN MODELING THE PREDICTION, MITIGATION, AND EVALUATION OF CONGESTION IN THE NATIONAL AIRSPACE SYSTEM

By

Kleoniki Vlachou

Dissertation 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
Doctor of Philosophy
2014

Advisory Committee:
Professor David J. Lovell, Chair
Professor Michael O. Ball
Professor Paul Schonfeld
Professor Cinzia Cirillo
Dr. Robert Hoffman

Dedication

To my beloved parents Georgios and Christina Vlachou.

Acknowledgements

This has been a long journey and it wouldn't be feasible to reach the end without the help of many valuable people. Most of all, I would like to thank my advisor Pr. David Lovell for his guidance, help and support all these years. He is a very knowledgeable educator and a great mentor. I admire his ingenuity and his approach to problem solving, and he will always be a role model for me.

I would also like to thank Pr. Michael Ball, it was an honor working with him and being guided by him. I would also like to thank the members of my committee, Pr. Paul Schonfeld, Pr. Cinzia Cirillo and Dr. Bob Hoffman, for their valuable feedback and direction.

I would also like to thank my NEXTOR office mates with whom we shared ideas, fun times and I got to learn a lot from them: Andy, Alex, Moein, Nasim, Charles, Kennis, Carina, Xenia and James.

Of course, this journey would not be the same without being surrounded by good friends, who made these past 7 years fun. I had the pleasure to meet and befriend so many people along the way. Friends, that we shared interests, concerns and experiences. I want to thank Kostas, Nikos, Giannis, Kostas, Laoura, Alex, Evripidis, Vasilis, Jason, Thodoris, Maya, Sasha, Udayan, Prem, Aram, David, Alice.

I want also to thank some wonderful friends that I hold in a special place in my heart: Konstantinos, Anastasia, Myra, Rama, Pauline, Gianluca, Meggy.

A special thanks goes to some friends that were my biggest support in the most difficult times and the best company for the wonderful times. My wonderful Giota Andrakakou who although thousands of miles away, was always there for me when I needed someone to talk; and I am so lucky to know her. I love you. My beloved Georgia Vergadou who is one of the very few people I could open up to when I need it the most. I love you. Christos Vergados who knows me so well and would always understand me even if I said nothing. I love you. Yolanda Mahnke, my "sister" who is an amazing woman and cook (!) and it is a blessing to have her in my life. I love you. Konstantinos Zampogiannis who doesn't know how amazing he is, but I do, and I am so fortunate to call him my friend. I love you.

I want also to thank my family. My parents Georgios and Christina Vlachou, who always supported, inspired and believed in me. I love you. My brother Nikos and sister-in-law Giouli and of course my niece Elisavet and nephew Aggelos. I love them dearly. They have given me so much strength to keep trying and never give up.

Last but not least, I want to thank Haytham, who gave me the courage to finish this journey well. He is my biggest supporter. He was there in the craziest final moments, to cheer me up and make me believe that I have the strength to reach my goal. Enta habibi, behabak ktir.

Table of Contents

Chapter 1: Introduction

The air transportation system in the United States is one of the most complex systems in the world. Every day approximately 60,000 flights of commercial, military and general aviation aircraft occupy the National Airspace System (NAS). Air traffic volume has proven to be quite volatile and susceptible to outside political and economic pressures, including most recently the terrorist attacks of September 11th and the recent economic recession. These events of course caused air traffic to decrease, and continue to hamper its recovery. Despite the economic downturn, the long range forecast remains positive. Even in a pessimistic scenario, the Federal Aviation Administration (FAA) has predicted that the average passenger enplanement growth for the following 10 years will be about 1.5 percent per year (FAA, 2013a). In Figure 1.1 we can see the 2013 forecast for the system enplanements for the years 2013-2022 under different scenarios. In the optimistic scenario, the growth rate of the passenger traffic could reach 3.4 percent annually.

Another forecast for the next 20 years, which we can see in Figure 1.2, shows that the passenger enplanements could increase beyond 1 billion passenger enplanements in the next 20 years. This would be a result of an average annual increase of 2.2 percent.

1.1 Problem Description

Projections of increasing demand only represent one side of the problem in aviation system management, however. The capacities of various system resources can also be volatile and affected by exogenous random events, such as fluctuating weather conditions and equipment outages. From a management perspective, it is essential to make efficient use of the available resources and to create mechanisms that will help alleviate the problems of the imbalance between demand and capacity. Even with good forecasts of demand growth and capacity evolution, it is still complicated to model the specific consequences to be expected at individual airports. Day-to-day airline and airport operations are also quite complex, with their own uncertainties, plus un-knowable factors such as the proprietary actions of air carriers.

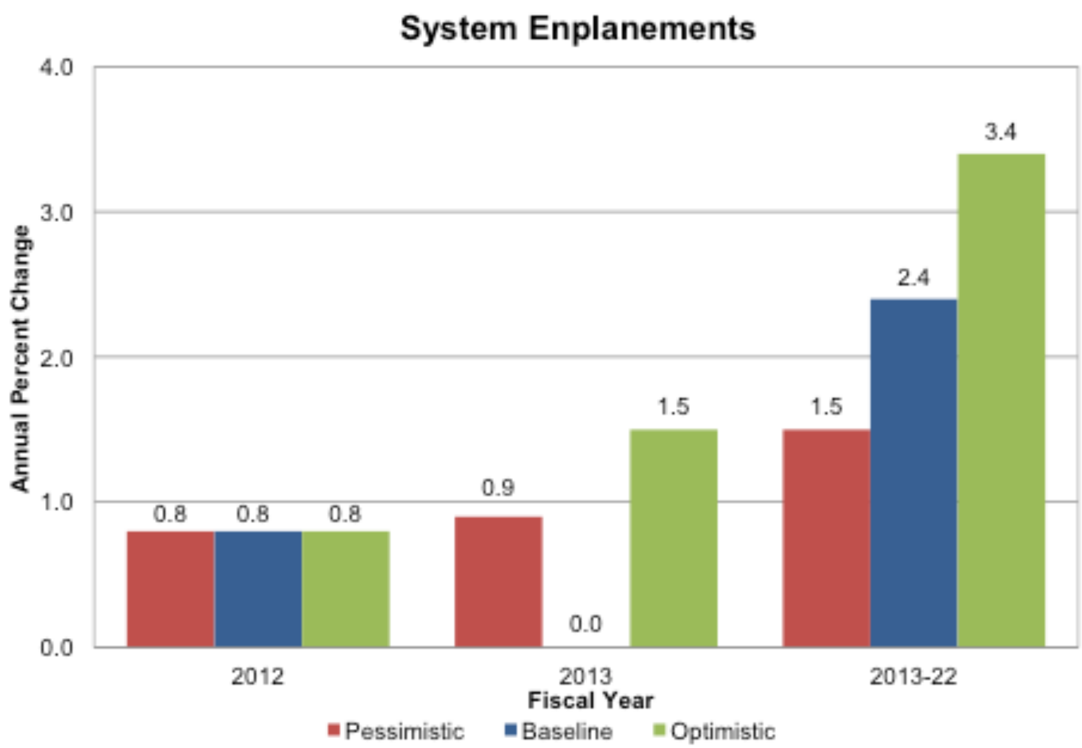


Figure 1.1 2013 FAA Forecast for the system enplanements for the years 2013-2022 (Source: 2013 FAA forecast)

The performance of the NAS is too complicated to be represented directly by its “inputs” demand and capacity. Together, these things conspire to produce other metrics, chief amongst them delay. Not surprisingly, when air traffic demand increases, system delays tend to do the same. Similarly, delays increase with decreases in resource capacity, either systemic or impromptu, such as under the influence of adverse weather.

Figure 1.3 shows the actual flight operations throughout the NAS for the years 2004-2013. On the left portion of the figure, we can see evidence of the continued slow recovery of NAS traffic after September 11, 2001. The economic downturn that began in 2007-2008 is also reflected in a commensurate decrease in air traffic. There is some hint, on the right portion of the figure, that traffic is beginning to rise again.

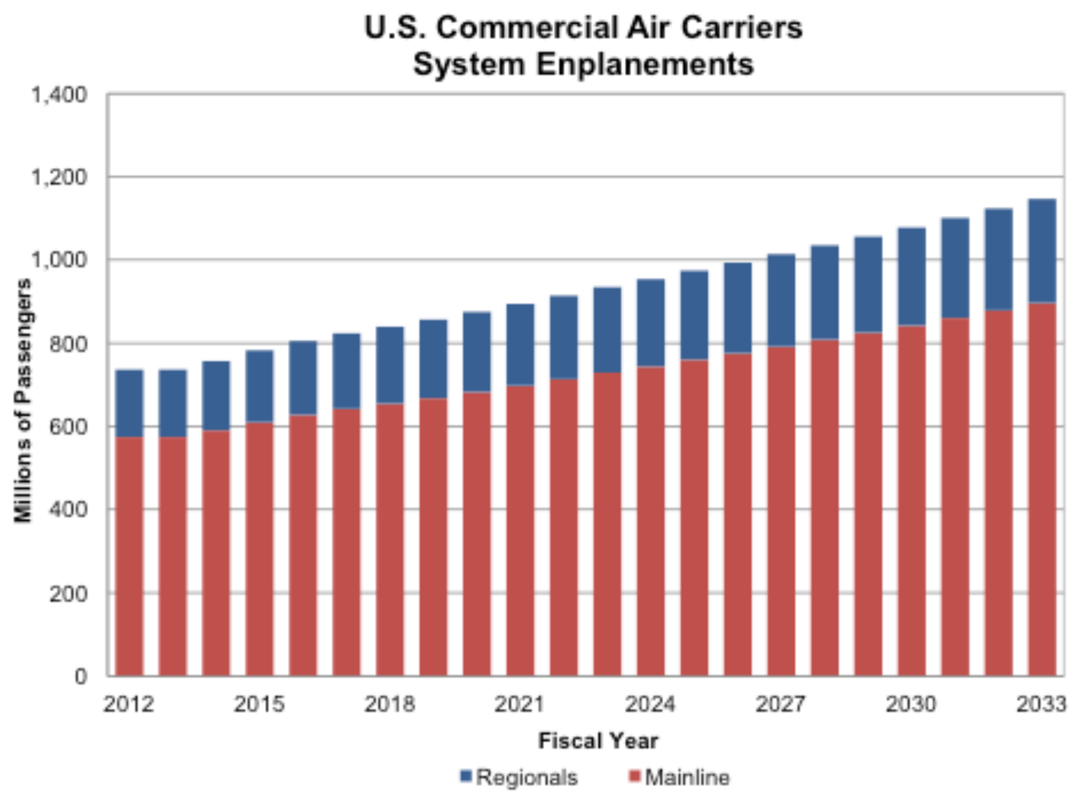


Figure 1.2 2013 FAA forecast for the system enplanements till the year 2033

(Source: 2013 FAA forecast)

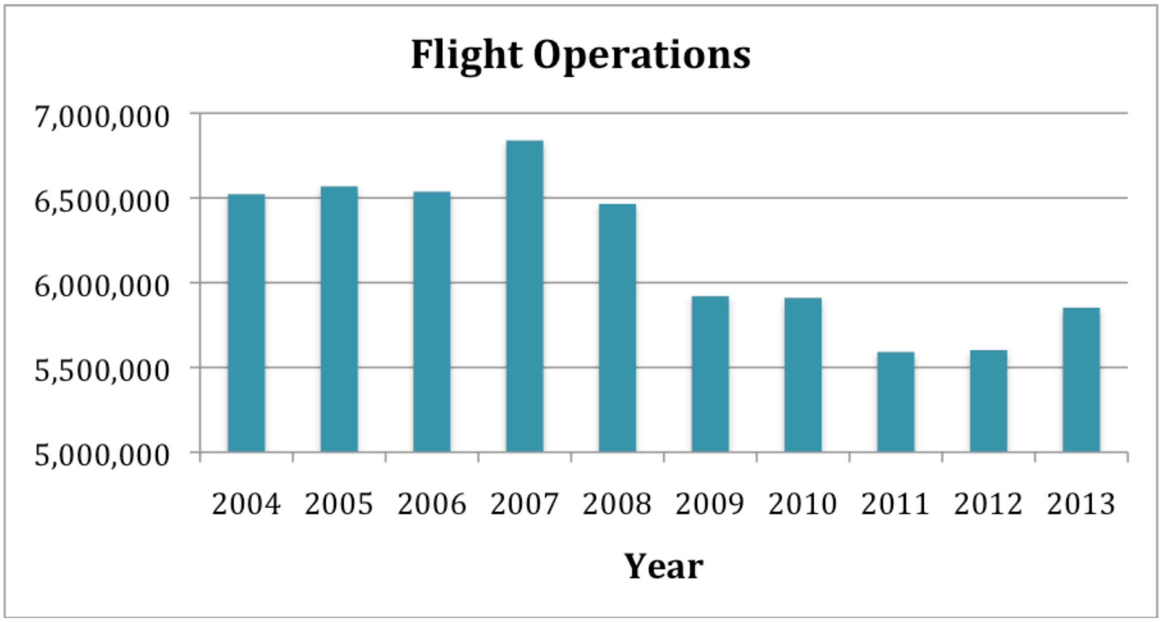


Figure 1.3 Flight operations for 2004-2013

For the same years if we look at the percentage of flights that were delayed, in Figure 1.4, we will see how these correlate to the traffic. Arrivals were delayed heavily in 2007, when traffic had reached a seasonal high. A commensurate pattern of slightly lower magnitude can be seen for the departures. When the system was at its worst, about 24 percent of arrivals and 21 percent of departures were delayed more than 15 minutes. The delays in 2007 were estimated to have cost the U.S. economy as much as $41 billion according to a report by the Joint Economy Committee (2008). More specifically, it was estimated that the traffic delays caused an increase of $19 billion of airlines' operating costs, a $12 billion cost of passengers' time and about $10 billion cost to other industries.

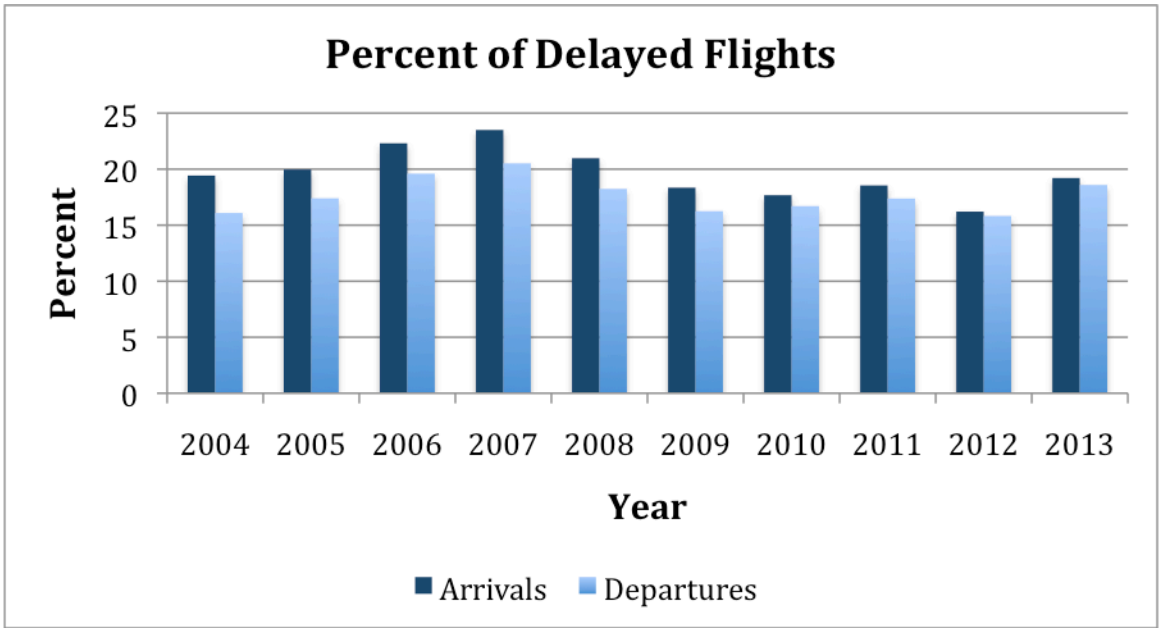


Figure 1.4 Percentage of arrivals and departures that were delayed in the years 2004 – 2013

Looking at Figure 1.4 we can see that once the economy started impacting air traffic, the delayed flights dropped to about 15 – 18 percent. For 2013, where air traffic showed some significant increase, the delayed flights also increased in turn, to reach about 20 percent.

It is evident that air traffic delays are always present and the more air traffic increases the more the delays will increase with very unwanted economic impacts. It is of great interest to study them further in order to be able to more effectively mitigate them. A first step would be to try to predict them under various circumstances. A second step would be to develop various mechanisms that will help in reducing delays in different settings.

Once the system is better handling unpredictable situations and delays are reduced, so will the need of airlines to pad their schedules against unforeseen circumstances. It is very well known that airlines, due to their need to adhere to their schedules, tend to add extra minutes to their scheduled flight times. Of course this allows them to be more resilient in cases where some flights suffer delays. In so doing, carriers can keep up with their schedules, have less missed connections and better on-time performance. At the same time this causes them to incur costs because of the extra time that planes and crews go unused. This is another aspect of traffic delays that is very interesting to look at further.

1.2 Air Traffic Management

In quite broad terms, the intended scope of this dissertation is to make inroads into three areas discussed so far: prediction of NAS system delays, system designs to allow carriers and the FAA to work together to reduce congestion and delays, and estimation of expected benefits from improving system efficiency. Before defining this scope of work in greater detail, a brief presentation of Air Traffic Management as it is contemporarily understood will follow in order to get a better understanding of the system, its components and actions.

Air Traffic Management (ATM) is essential for the safe and efficient operation of airports and the airspace system. Advanced ATM systems as defined by deNeufville and Odoni (2003) must:

- Accommodate an increasing number of users
- Achieve an exceptional level of safety

- Have a large number of skilled human operators to work seamlessly with a network of computers and communications, surveillance, and navigation equipment
- Take advantage of technological developments
- Keep the cost of all of this at a reasonable level

ATM is considered to consist of two major components: Air Traffic Control (ATC) and Air Traffic Flow Management (ATFM), as can be seen in Figure 1.5 below (Vossen, 2002). ATC refers to the processes that provide tactical separation services. The system users must maintain enough horizontal and vertical separation to avoid the risk of collision but at the same time capacity must be efficiently used (Ashford and Wright 1992).

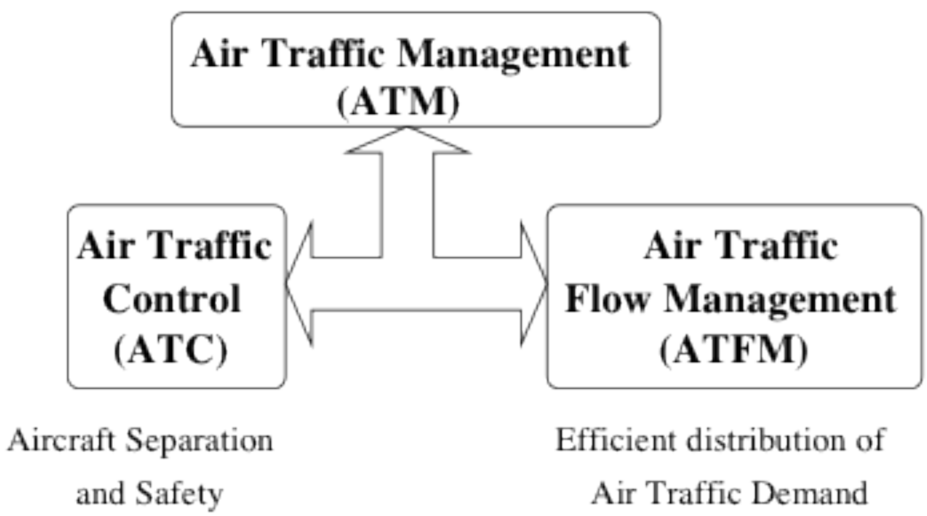


Figure 1.5 Air Traffic Management components

(Source: Vossen, 2002)

Air Traffic Flow Management aims to detect and resolve demand-capacity imbalances by adjusting the flow of aircraft so that demand matches as well as possible the available capacity.

The principal functions of ATFM, according to deNeufville and Odoni (2003), are:

- To predict the locations of potential overloads
- To develop strategies that will relieve these overloads
- To oversee the implementation of these strategies

Odoni (1987) has classified the ATFM initiatives that can resolve air traffic congestion in the following categories:

- Long-term approaches that are focused on increasing capacity. That is achievable by constructing new airports or adding more runways to existing facilities. This approach is characterized by very high costs and raises many environmental concerns. Thus it is more difficult to implement.
- Medium-term approaches that are more administrative and economic in nature and try to mitigate congestion by modifying temporal or spatial traffic patterns. For example at some airports, where the demand for airport infrastructure significantly exceeds the airport's capacity, slots are allocated to the airlines by a coordinator, according to International Air Transport Association (IATA) guidelines (IATA 2012). Other approaches considered are congestion pricing and slot auctions.
- Short-term approaches that consist mainly of adjusting the air traffic flows to match demand with available capacity. These approaches tend to mitigate

congestion caused by unpredictable disruptions such as bad weather and are performed a few hours in advance.

The research proposed for this dissertation is mostly motivated by some short-term initiatives.

1.2.1 Air Traffic Flow Management Initiatives

The Federal Aviation Administration (FAA) is responsible for the coordination of air traffic and for ensuring the proper separation requirements in the controlled airspace. In order to carry out these functions, the FAA has divided the airspace over the Continental United States into 20 areas, as shown in Figure 1.6 (FAA 2009). Each of these areas is controlled by the corresponding Air Route Traffic Control Center (ARTCC). Each en route center is then divided into smaller areas, called sectors, because the traffic load in any single center is too much to be handled by one controller.

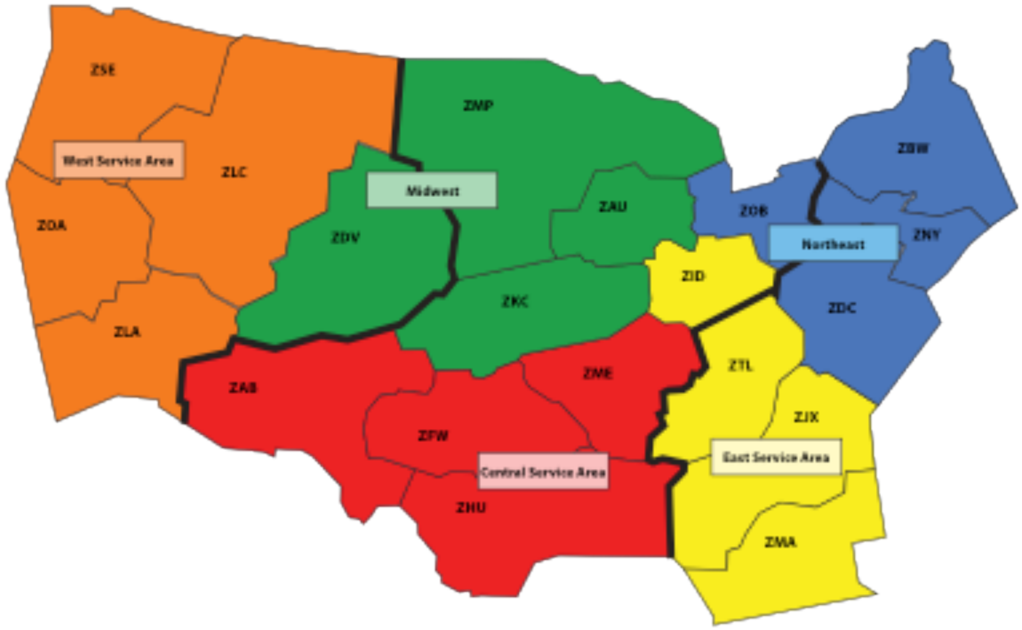


Figure 1.6 En-Route Traffic Control Centers, United States.

(Source: Federal Aviation Administration)

Traffic controllers guide the aircraft from one sector to another until they arrive approximately 50-150 nautical miles from the destination airport. Then the aircraft are handled by the Terminal Radar Approach Control facility (TRACON). Finally, for the last 5 nautical miles to the destination airport, the handling of traffic is performed by the airport traffic control towers. A detailed flow chart of a flight trajectory decomposed by phase of flight is depicted in Figure 1.7 below (FAA 2009).

Air traffic controllers in the above centers are responsible for the movement of aircraft within their area of responsibility and their decisions are based on real-time information about the flights entering their sectors. The strategic ATFM functions performed by the FAA are coordinated by the Air Traffic Control System Command Center (ATCSCC) located near Washington DC. The ATCSCC continuously monitors the current and forecast traffic demand throughout the National Airspace System (NAS) and identifies potential problems (like bad weather) that may constrain capacity. Whenever demand is predicted to exceed capacity the ATCSCC generates and implements strategies to mitigate the problem.

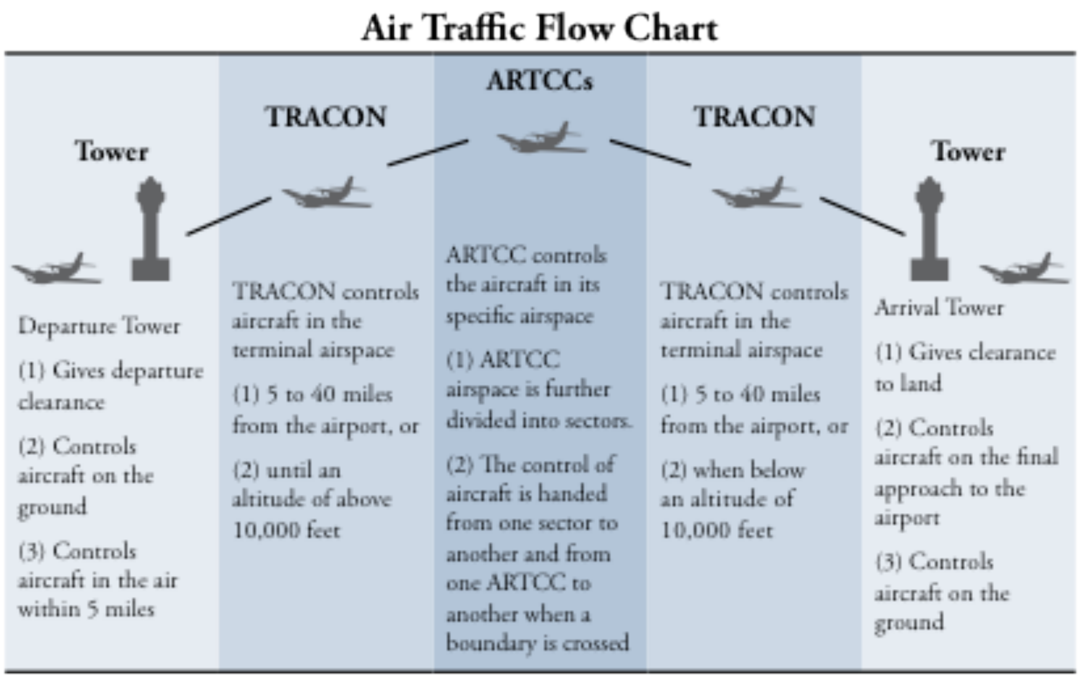


Figure 1.7 Air Traffic Flow Chart

(Source: Federal Aviation Administration)

At this point it is important to mention that there is a need for partial decentralization of decision making. The FAA does not necessarily have all the information needed to make decisions on behalf of the users. To this end, Collaborative Decision Making (CDM) is an effort to improve Air Traffic Management. It is essential for all stakeholders involved (FAA, airlines) to share information so that all will be aware of the current situation. All users of ATM will benefit from sharing information and collaborating. As described by deNeufville and Odoni (2003) the specific goals of CDM are:

- To provide the FAA and the airlines with a common picture of the current and forecast air traffic conditions
- To allow the person or organization in the best position to make each decision
- To make the decisions in open manner so all will know what is happening and can contribute as necessary

The ATFM initiatives that the FAA can implement under a CDM framework may be outlined as follows.

Ground Delay Programs

The Ground Delay Program (GDP) is a mechanism implemented when it is projected that the arrival demand at an airport would exceed capacity, usually because of adverse weather conditions around the airport area (Ball and Lulli 2004), although occasionally as a result of over-scheduling. The goal of a GDP is to decrease the arrival rate and this is achieved by intentionally delaying the take-off times for most flights intending to land at that airport (deNeufville and Odoni 2003). The motivation for doing so is that it is safer and cheaper for flights to absorb delays on the ground before take-off rather than while airborne.

Metering

Metering is used to control the rate that aircraft cross some specified spatial boundaries by adjusting the spacing between aircraft (deNeufville and Odoni 2003). Metering procedures can be divided in two categories. The first is time-based, which controls the minimum time headways with which aircraft are allowed to pass a specific geographical point. The second category is distance-based, which specifies a minimum separation (in miles) between aircraft moving at the same direction. This is also known as “Miles-In-Trail” (Vossen 2002).

Rerouting

When an airspace area is impacted by adverse weather conditions and its capacity is reduced, rerouting might occur. With rerouting some flight routes are changed or restructured to modify the distribution of traffic flows (deNeufville and Odoni 2003). Rerouting sometimes is part of Severe Weather Avoidance Programs (SWAPs), which are implemented when traffic flows are affected by widespread severe weather in the system (Vossen 2002).

Airspace Flow Program

Airspace Flow Program (AFP) is a relatively new traffic flow initiative that was first introduced in the summer of 2006 and it marked a new step in Air Traffic Management (FAA 2009). When there is a capacity reduction in an area of the airspace (not a specific airport) due to adverse weather conditions, and rerouting by itself is not enough to deal with the problem, the AFP is intended to solve the problem. The first step is to identify the problematic area by creating an FCA. A Flow Constrained Area (FCA) is an area of the airspace that the FAA has identified as potentially congested (Hoffman et al., 2004). NAS users are informed when FCAs are issued along with reroute advisories in order to reduce the number of planes passing through the impacted area to an acceptable level. Sometimes a Flow Evaluation Area

(FEA) precedes an FCA. The Traffic Situation Display (TSD) and the Common Constraint Situation Display (CCSD) provide traffic managers and flight dispatchers with the ability to define and display FEAs and/or FCAs (Vakili 2009). In the second step the Enhanced Traffic Management System (ETMS) takes the FCA description and produces a list of the affected flights and the times at which they are expected to pass through the FCA (Vakili 2009). This list is sent to the Flight Schedule Monitor (FSM) where flight operators, traffic controllers and service providers have access to the information and have a "common picture" of the situation. The traffic manager enters the expected capacity – expressed as the number of flights that can pass through the FCA per hour - in the FSM, which will then compute the best departure time for each flight scheduled to pass through the FCA, in order to lower the demand to meet the new capacity. During an AFP, resources are rationed by procedures such as Ration By Schedule (RBS), which is based on the principle of first-scheduled-first-served.

The FSM sends the controlled departure times to the flight operators and control towers. After they get the new flight-slot list, airlines have various options. They can have some of their flights depart according to the controlled times, they can reroute other flights around the FCA, and they can cancel some others. They can also swap departure times among their set of flights or they can even exchange slots with other operators, in a controlled transaction structure. All of these mechanisms are part of Collaborative Decision Making. Importantly, it is evident that airlines have at least some control of their flights and options *after* the initial allocation.

Collaborative Trajectory Options Program

One of the practical problems with the AFP, as described above, is that all of the mechanisms that carriers can employ to inject some of their individual preferences into the dispositions of their flights come *after* the initial capacity allocation done by the FAA. These machinations take time and are complicated to communicate quickly between the FAA and the carriers. Further, there is a sense that, if the carriers are to make significant revisions to each initial AFP allocation, then there is a waste of efficiency in the process to begin with. It might make more sense to try to capture at least some sense of their intentions *before* the initial allocation is conducted, thereby minimizing the effort involved in revising and coordinating with the FAA.

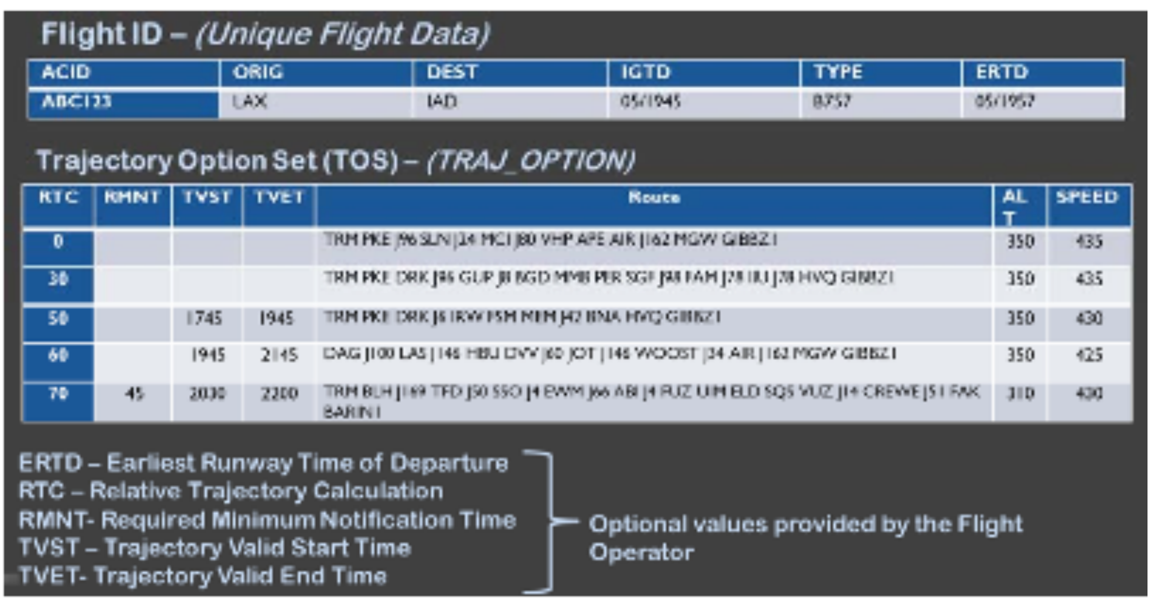
This is the primary driver behind the Collaborative Trajectory Options Program (CTOP), which is part of the Collaborative Airspace Constraint Resolution (CACR) concept. With this program, airlines can express which flights they want to be assigned to which slot in the AFP capacity allocation process.

There are three main points that distinguish CTOP from the current practices (FAA, 2011a).

- The communication between the FAA and the flight operators will be all electronic, which will make it faster and more precise than having voice calls or text advisories.
- The flight operators can express their preferences in great detail and FAA honors them to the extent that the problem permits.

- The decisions for the routes and assigned delays are made by automation and not humans, which allows a finely tuned solution to the congestion problem.

Here follows an example of message exchange in CTOP for a better understanding of it (FAA 2011a). First, the flight operators send Trajectory Options Sets (TOS's) to the Traffic Flow Management System (TFMS). A TOS is a set of trajectories for a given flight that are acceptable to the flight operator. Trajectories consist of route text, altitude, speed and departure time that specify the intended path of a flight through the NAS. An example of TOS for a flight from LAX to IAD is depicted in Figure 1.8 (Flow Evaluation Team, 2014). The flight operator specifies the relative cost of each trajectory options as the extra amount of delay compared with the best trajectory (this has a delay equal to zero).



Flight ID – *(Unique Flight Data)*

ACID	ORIG	DEST	IGTD	TYPE	ERTD
ABC123	LAX	IAD	05/1945	B757	05/1957

Trajectory Option Set (TOS) – *(TRAJ_OPTION)*

RTC	RMNT	TVST	TVET	Route	ALT	SPEED
0				TRM PKE J96 SLN J24 MCI J80 VHP APE AIR J162 MGW GIBBZ1	350	435
30				TRM PKE DRK J96 GUP J8 BGD MMB PER SGF J88 FAM J78 IIU J78 HVQ GIBBZ1	350	435
50		1745	1945	TRM PKE DRK J6 IRW FSM MEM J42 BNA HVQ GIBBZ1	350	430
60		1945	2145	DAG J100 LAS J146 HBU DVV J60 JOT J146 WOOST J34 AIR J162 MGW GIBBZ1	350	425
70	45	2030	2200	TRM BLH J169 TFD J50 SSO J4 EWM J66 ABI J4 FUZ UIM ELD SQS VUZ J14 CREWE J51 FAK BARIN1	310	400

Figure 1.8 Example of TOS for a flight from LAX to IAD

(Source: Flow Evaluation Team)

When the FAA defines one or more FCA's, the TFMS under CTOP sends a Traffic Management Initiative (TMI) with the airspace constraints to the flight operators, who reply with TOS's appropriate for the subject flights. Through TFMS the traffic manager determines the best solution from a system perspective, taking both overall FAA objectives and carrier-declared delays into account. The solution consists of assigned routes, arrival slots and departure times for the constrained flights. If a flight does not have multiple trajectory options, it will be controlled by the current route. Thus, clearly carriers are not bound to participate or to declare TOS's for all of their flights; there is a default posture for each flight when entering the optimization process.

This process is not static – flight operators monitoring their flights are allowed to send new TOS's (clearly only for those flights whose dispositions are not yet fixed), to update them according to evolving traffic conditions and other new information. In that case, TFMS will evaluate the updated TOS's and possibly find new solutions. If any are found, they will notify the flight operators of the changes.

Given any particular set of flight dispositions, the flight operators can determine the impact the solution has on their operations, and they might decide to cancel a flight, or ask for substitution. In the latter case, TFMS will check the feasibility of the proposed substitution and will give the associated flights their new assigned routes, departure times and slots if it is possible. Ultimately, the flight operator will file a flight plan where the route will be in accordance with the CTOP assigned route.

Collaborative Airspace Constraint Resolution

The Collaborative Airspace Constraint Resolution (CACR) concept extends the capabilities of the Collaborative Trajectory Options Program (CTOP) by managing flights within 45 minutes prior to departure and with adequate automation assistance provided to traffic managers for defining airspace constraints (Metron Aviation 2012). According to Stalnaker et al. (2009) CACR has four key components:

- Predicts sector demand and takes into account its uncertainty
- Predicts sector capacity and potential impact adverse weather will have
- Identifies the problem
- Generates congestion resolution initiatives

CACR will also be implemented under the CDM framework since it will collect and incorporate user preferences in terms of enhanced Trajectories Options Sets (TOSs).

The last two initiatives are of particular importance for this dissertation because they motivate the second proposed area of work, that of designing a simpler language of exchange between the FAA and carriers by which some carrier preferences can be incorporated into the capacity allocation process. The first and third aims of the dissertation were to develop new modeling methods for delay prediction and to assess the benefits associated with improved efficiency, particularly as it relates to the predictability of the system. As will be seen below, the first part is mainly concerned with predicting delays at a course scale where details of specific traffic management initiatives are not known, but rather the basic inputs of demand and capacity (either the natural capacity of the airport and runway configuration or the reduced capacity imposed by a Ground Delay Program) are enough to make broad predictions of delays at an airport. The third part of the dissertation is aimed at benefits assessment, in this case in response to technological innovations that might improve the predictability of the system. This is likely to be addressed at both strategic and tactical scales, so some consideration of the details of traffic management initiatives may be warranted.

1.3 Scope of Work

The scope of this dissertation is to look closer at a threefold approach to the problem of congestion in aviation mentioned above. In part, this is predicated on the fact that three separate sponsored research projects have already been conducted, each of which provides the fodder for a specific line of inquiry in the dissertation. The common thread between the three is airspace congestion and the management thereof.

The first effort is the prediction of delays and the development of a model that will make these predictions under a wide variety of distributional assumptions. This work was conducted as part of a NASA-sponsored project whose purpose was to develop queuing models as a means of assessing how precise adherence to 4D trajectories, that specify current and future aircraft position, will affect capacity and delay in the NAS. The focus of this effort was specifically on a continuum approximation using diffusion methods that enables efficient solutions under a wide variety of distributional assumptions.

The second part of the work effort proposed herein was initiated as part of an Aviation Cooperative Research Program (ACRP) fellowship that was awarded to Ms. Vlachou. Her proposed research topic for that fellowship was the design of a parsimonious language of exchange, with accompanying allocation mechanisms that allow carriers and the FAA to work together quickly, in a CDM environment, to allocate scarce capacity resources.

The third portion of the dissertation derives from a research project sponsored by the FAA that is only recently completed. The overall goal of that project was to develop metrics and assessment methods for technological innovations that might effect an improvement in predictability in various dimensions of the NAS. The scope of this work was specifically to assess the monetary benefits of improvements to predictability as manifested through carriers' scheduled block times.

1.3.1 Prediction of Delays

Studies of queuing delays in the National Airspace System (NAS), and other large networks, for that matter, are typically conducted either in a Monte Carlo simulation environment, where a considerable amount of fidelity is available at the expense of computational efficiency, or with closed-form equilibrium queuing models fraught with distributional assumptions that are typically not very representative of real situations. A common example of the latter is the use of the Poisson process to represent arrival processes to queues, motivated by its mathematical tractability, even in the face of fairly compelling evidence that the system is not Markovian.

1.3.1.1 Existing Queuing Models

One well-known aviation queuing model is LMINET (Lee et al., 1997, 1998), in which a network of airport queues is represented by means of interconnected single-server queues. Each queue has a time-dependent Poisson arrival process, and an Erlang-*k* service process. One serious problem with this approach is that because the input process to each downstream node is Poisson, one cannot have independent control of its mean and variance. Thus, while the outputs from upstream nodes may have variances different than what the Poisson process would be constrained to, the model cannot enforce these properly. More importantly, any technologies or policies that might be adopted to reduce variance in the system (such as improved trajectory accuracy) cannot be modeled accurately. The goal of this part of the dissertation is to provide a single-airport building block that might eventually be extended to a network environment, and that would allow for modeling of more complex and dependent interactions between aviation network nodes.

Another single-airport queuing model commonly used in aviation is the DELAYS model developed at MIT, the methodology behind which is captured in Kivestu (1976), Horanjic (1990), and Malone (1995). This model uses a time-dependent Poisson arrival process and an Erlang-*k* service process, much the same as LMINET (both models have a common heritage). One major difference is that the DELAYS model was later adopted for a network structure that does not suffer from the same independence problems as outlined for LMINET. The Approximate Network Delays (AND) model was originally proposed in Malone (1995), but was not assembled into a working model until more recently (Malone and Odoni, 2001). The idea driving the

AND model is that the DELAYS model, by itself, might produce excessively large estimates of delay, when fed purely scheduled arrival times. In reality, the network would not permit such large delays, as demands would be spread over time due to controller actions, metering by upstream queues, etc. Thus, the AND model iterates between the DELAYS model and a delay propagation algorithm, in an effort to find an estimate that more closely matches expectations. This is a heuristic approach, and it still suffers from the drawback this research is intended to address, which is the strong dependence between arrival process mean and variance.

Another aviation queuing model is the National Airspace System Performance Analysis Capability (NASPAC), which was developed beginning in the 1980s by the Federal Aviation Administration (FAA) and Mitre Corporation. A good description of the original model can be found in Millner (1993). The model is now housed at the FAA, and continues to be developed (see for example Post et al., 2008). The model includes a number of detailed components, such as realistic fleet information, fuel burn, etc., but its queuing engine is quite rudimentary, consisting of a simple deterministic queue with scalar capacity values for the airports. The claimed path forward to dealing with real stochastic queuing effects is to incorporate Monte Carlo simulation (Post et al., 2008), which will seriously impact the computational complexity of the model, as described above.

1.3.2 Delay Mitigation

As presented earlier, FAA implements various Traffic Management Initiatives (TMIs) in order to mitigate problems that arise due to the demand-capacity imbalance in the system. The most recently developed TMI is the Collaborative Trajectory Options

Program (CTOP), which is about to start being implemented and has as a great advantage the increased participation of the airlines in order to determine which flight is assigned to which slot. A piece of work that considered this feature (the extensive listing of airlines preferences), was conducted by Vakili (2009). In this research she considered a detailed way for airlines to express their preferences, where she provided for each flight in which priority they will be assigned to which slots. Then she presented various resource allocations mechanisms, other than Ration-by-Schedule, that take into account these preferences and allocate the slots to flights.

The allocation mechanism she proposed falls into a category of methods designed for fair treatment of claimants to, and allocation of, a scarce resource and this subject has received considerable attention in the applied economics literature. One of the more well-known problems is the apportionment problem, which exists when a set of indivisible objects must be distributed among numerous claimants in proportion to their claims (Young 1994).

1.3.3 Improved Predictability Will Ultimately Lead to Reduced Delays

The Federal Aviation Administration continues an effort to address customer requirements. One such requirement is to be accountable for the quality of service provided. The Office of Performance Analysis and Strategy contributes to the FAA's success by analyzing and monitoring performance through existing metrics and proposing new metrics that better evaluate National Aviation System efficiency and the FAA's customer service.

The FAA has identified 11 categories of system performance indicators: access and equity, capacity, cost-effectiveness, efficiency, environment, flexibility, global interoperability, safety, predictability, security, participation by the Air Traffic Management community (FAA 2011b). The one that has gained more attention lately is flight predictability and the work covered by this effort is contributing to research of the flight predictability concept in aviation.

By increasing flight predictability, airlines should experience significant benefits, mostly because this will allow them to reduce their scheduled block times. Scheduled block time is a major driver for crew costs and usage of equipment. Part of this research was to investigate how airlines set their scheduled block times. This will allow the examination of potential benefits for the airlines, at a strategic level, when the scheduled block times will be reduced because of the increased predictability in the system. Also by reducing the actual block time there will be benefits in the day-of-operations level, where the passenger delays will be reduced and regarding the crew perspective there will be less overtime and fewer crew time-outs.

1.3.3.1 Defining Predictability

There is some work done in the past related to flight predictability. The term flight predictability is defined in ICAO (2005) as "the ability of airspace users and ATM service providers to provide consistent and dependable levels of performance. Predictability is essential to airspace users as they develop and operate their schedules". In another report by Bolczak et al. (2007) predictability measures how the airspace user experiences the variation in the Air Traffic Management System.

Most of the studies in the literature for aviation and public transit that use predictability metrics are dealing with travel times. There are some that use delay and flights per day to quantify predictability, or the lack thereof. Following we summarize the metrics in each category.

Flight/Travel Times

In the work of Gulding et al. (2009) flight predictability is measured as the difference between the 80th and 20th percentiles of the distributions for taxi-out, en route, and taxi-in times. In this work they have broken the flight into different phases and considered one metric for each of them. In another report (Bolczak et al., 1997) predictability is defined as the deviation of ground movement times, the statistical spread of ground movement times and the statistical deviation of en route times. According to an ATSP Focus Group report (1999), a predictability metric considered is the ratio of the actual flight time to the scheduled flight time.

There are also studies related to predictability in public transit. In Taylor (1982), a metric for travel time variability is defined as the coefficient of variation (the ratio of the standard deviation to the mean) of travel time. In another work by Uniman et al. (2010), they consider the reliability from the passengers' perspective and they define it as the standard deviation of the travel time distribution. They also introduce the notion of the Reliability Buffer Time (RBT) metric, which is the buffer time that passengers must allow above their typical travel time to arrive on time with a specified level of certainty. This is analogous to air carriers building buffers into the scheduled block time.

Delay

A study that considers delay as a metric for flight predictability is by Gulding et al. (2009). They define that predictability is measured from the flight perspective as the difference between the 80th and 20th percentiles for pre-departure delay and arrival delay. In the report by Bolczak et al. (1997) predictability is measured by the arrival, departure and overflight delay. Predictability is assumed to be a measure of the variance between the planned and realized delay in a report by Ball et al. (2011).

Flight per Day

Finally there is a report by EUROCONTROL (2003) where they consider the seasonal and hourly variability as metrics for flight predictability. In more detail the metrics considered for the seasonal variability are the ratio of summer to winter traffic in flights per day and the ratio of traffic in the peak week to the average. For the hourly variation they define it as the ratio of the average hourly traffic to the average in the peak three hours.

1.4 Organization of the Document

In Chapter 2, the modeling effort for the delay prediction is presented. In Chapter 3 the allocation mechanism developed to reduce delays during severe enroute weather is presented, which takes into account the preferences of the airlines. In Chapter 4 the work done regarding identifying the benefits for the airlines as a result of increased predictability in the system is presented. Chapter 5 presents the conclusions, and suggestions for further extension of this work.

Chapter 2: Delay Prediction

Air traffic system undergoes a continuous transformation by shifting to smarter, satellite-based and more advanced technologies (FAA, 2013b). An important feature of the Next Generation Air Transportation System (NextGen) is the use of four-dimensional trajectories (4D Trajectories). Aircraft position will be known not only in space but also in time. This will increase the precision of the operations and lead to the reduction of the required spacing between the aircraft. Currently the system is stochastic and with 4D Trajectories in place will move to a more deterministic system. The queuing models that are more suitable to predict delays in the current state are stochastic in nature and based on solutions of differential equations. Delays in the future, when everything will be more predictable, will be best modeled with deterministic models.

As precision increases – and the system is moving from the stochastic state to the more deterministic-, the airspace capacity also will increase and this will lead to reduced delays. As is was shown in the work of Hansen et al. (2009), the delay reduction as it is derived by comparing a stochastic and a deterministic queuing model will be of the order of 35%. Poisson models are great to capture the stochastic nature of the system and set the lower boundary of the system's performance. Deterministic on the other hand are the upper boundary. The work conducted here presents the creation of a model that will be the intermediate part. The evolution of models capturing the increased precision can be seen in Figure 2.1. In the future with more precision in place, mean will remain the same but variance will reduce, so it is was desired to be able to handle mean and variance independently and diffusion models offer this great advantage.

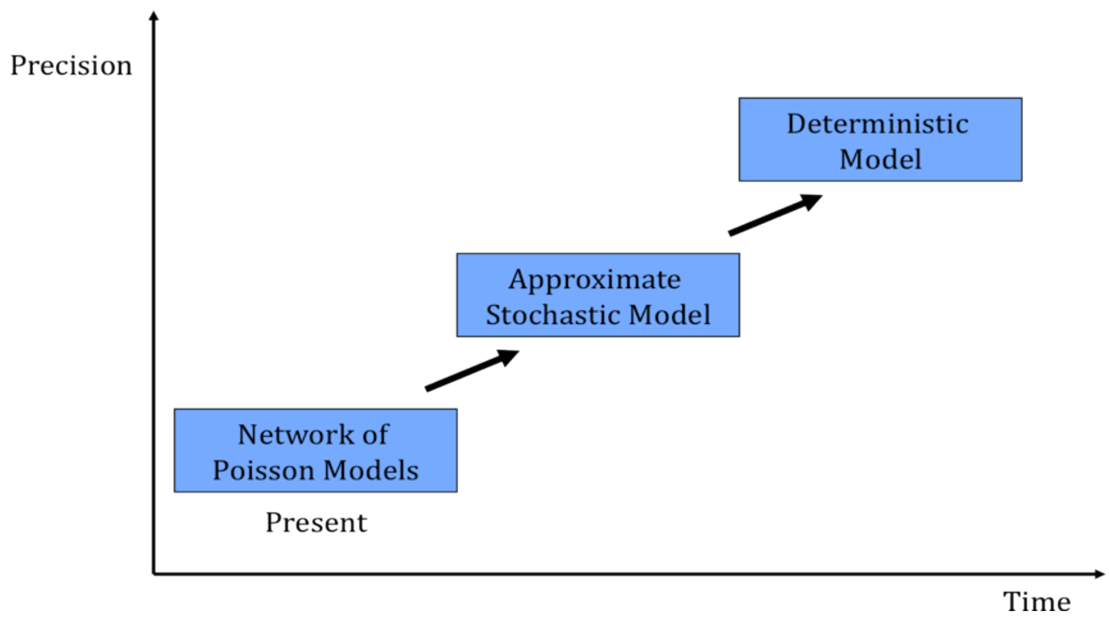


Figure 2.1 Queuing models under various degree of precision

With the aviation system in mind, the idea behind this research was to adapt a common continuous approximation technique known as the diffusion approximation to a queuing problem, with a specific interest in modeling arrival and departure delay statistics at an airport over the course of several hours or a day. The primary advantages of using the diffusion approximation for these purposes are that specific distributional assumptions can be relaxed in favor of an approximate description of the relevant stochastic processes by a small number of their time-dependent moments, that the full spectrum of probabilistic results can be obtained via a single run of the model, and that propagation of higher moments beyond the mean queue behavior can be captured. In general, is believed that it should be possible to represent a network of queues using methodology similar to the methods herein, although the results to date apply only to a single queue with a general arrival and general service process. A concise version of the results presented in this chapter is published in Lovell et al. (2013).

2.1 Model Development

In this section is introduced the modeling assumptions that lead to the particular continuum approximation for queuing systems known as the diffusion approximation. This consists of a governing differential equation, which is presented first, and which represents the primary dynamics of the system. This equation is valid for a closed subset of the real numbers representing all realistic values of the system state, but some boundary conditions must be imposed to prevent physically meaningless results outside of this interval. It is also described the set of initial conditions required to represent any particular queuing problem for which a solution is sought.

2.1.1 Governing Differential Equations

Diffusion methods have been applied to queuing problems in a variety of domains, including road transportation (Newell, 1971), computer networks (Kobayashi, 1974), and more general queuing systems (Gaver, 1968 and Kimura 1983). No significant use of them in an aviation setting is recorded in the literature. The development of the model shown in the following pages borrows very heavily from the exposition of Kimura, 1964, which develops the diffusion approximation in the context of a very different application, that of population genetics. The reason for following the template of that paper, however, is that the treatment is very thorough but also accessible to readers without prior experience in diffusion methods, and it can be adapted readily to the aviation context.

Suppose the arrival process to an airport is modeled as a single-server queue. This is admittedly an abstraction, because there are frequently multiple cornerpost entry points to an airport, often the possibility of multiple arrival runways, and incoming aircraft do not physically line up in queue in the same manner as customers at a grocery store, or even vehicles at a traffic signal. Nevertheless, it is common to model the competition amongst multiple arriving aircraft for the capacitated resource (the arrival runway system) as a queue, with the interpretation that the delays thereby imparted are assigned and incurred at en route locations farther away from the airport. Let $Q(t)$ represent the time-dependent random variable describing the length of the (virtual) queue for arrival aircraft at time *t*. While beyond the scope of this research, the ultimate goal of this endeavor is to model more complicated aviation networks. In that context, one could use the airport node being described here to model an arrival or departure resource like a runway, a gate, or an esoteric en route node intended to represent a capacity constraint in the airspace itself.

The first assumption necessary for consideration of continuum models is that of continuity; i.e., that the queue length measurement at any given time need not be an integer. Because aircraft only come in discrete units, this is obviously an artificial construct. However, it is mostly of interest in using queue length measurements as preliminaries to computing delay statistics, so they will be averaged over a large time domain. As a result, this assumption is probably no more malignant than assuming that there is such a thing as a "queue" at an arrival airport. This is a stochastic queuing system, and the probability density function for the queue length *x* at time *t* is denoted $f(x;t)$. A graphical example of f is shown in Figure 2.2. In this notional example, the queue density transitions over the time interval $[0,10]$, with a mean that increases and then decreases again, and a variance that changes similarly.

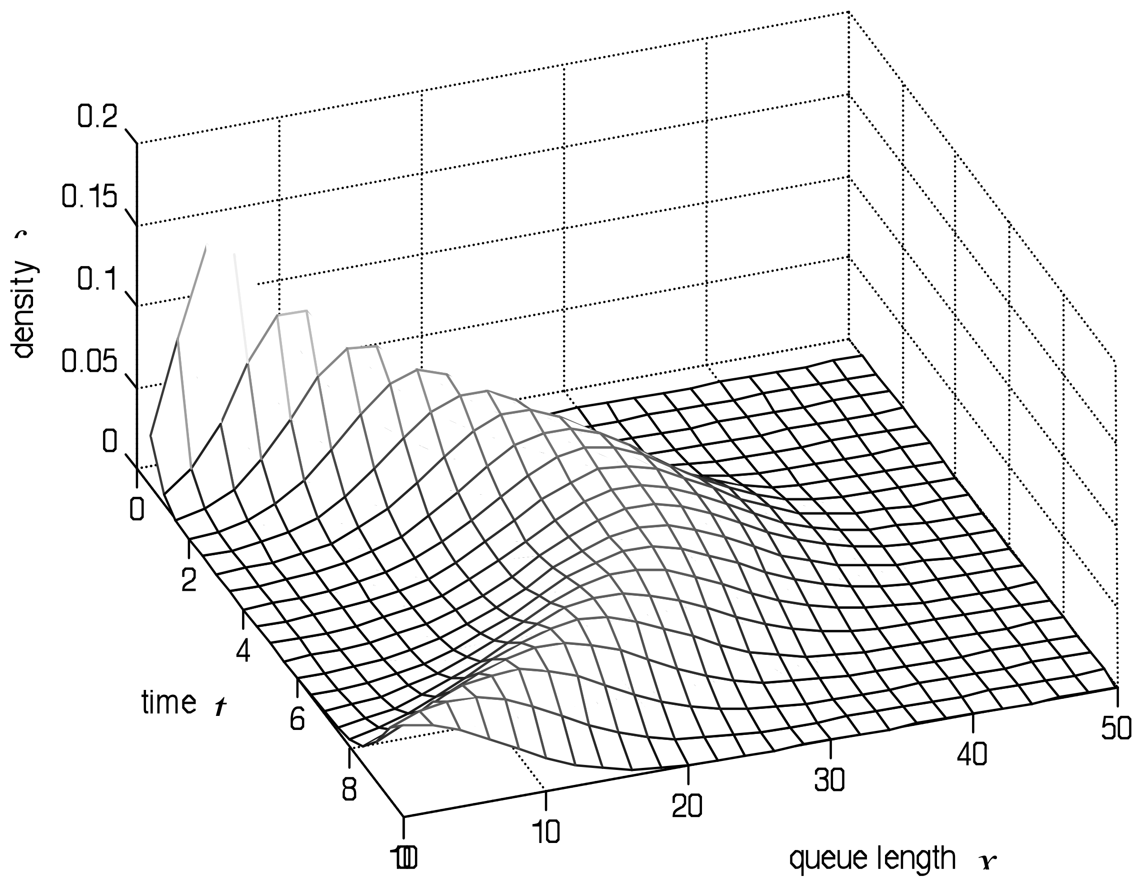


Figure 2.2 Queue length probability density function $f(x;t)$

The probability density transition function $g(\delta x, x; \delta t, t)$ is also defined as the probability density associated with a change in queue length from x to $x+\delta x$ in the time interval $[t, t+\delta t]$. An example of $g(\cdot)$ for a single choice of t and δt is shown in Figure 2.3.

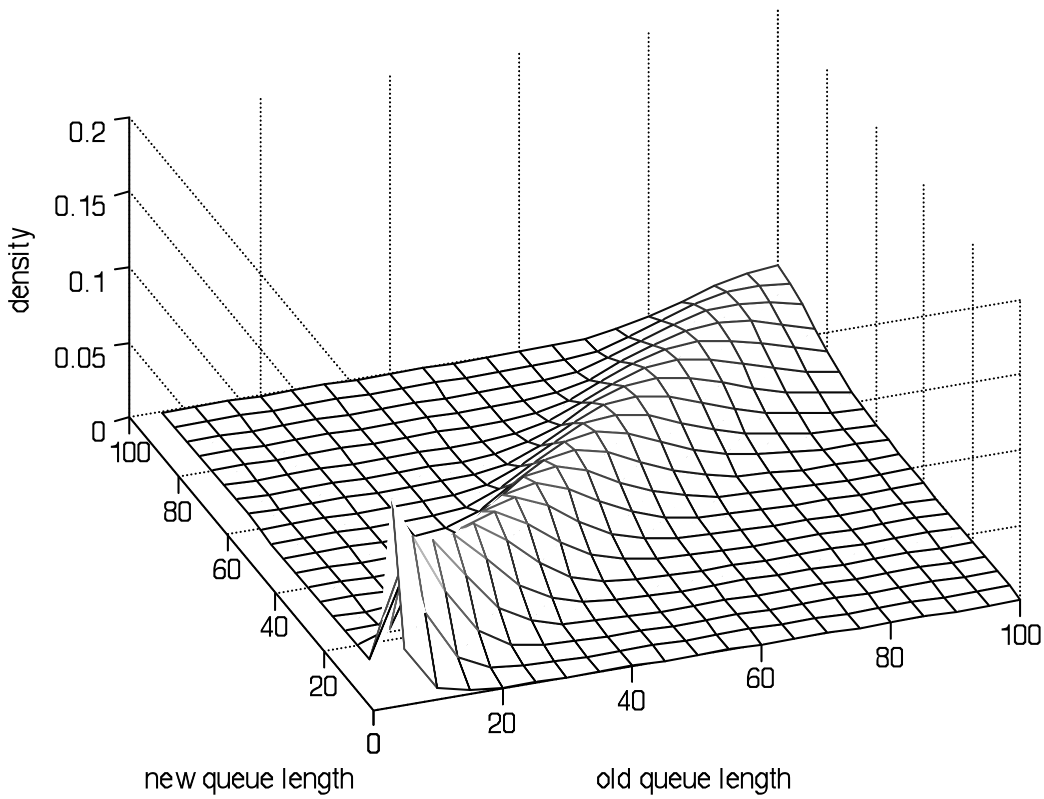


Figure 2.3 State transition probability function $g(\delta x, x; \delta t, t)$

The density function for the queue length at some future time $t + \delta t$ can be expressed using the continuous Kolmogorov-Chapman equation:

$$f(x; t + \delta t) = \int f(x - \delta x; t) g(\delta x, x - \delta x; \delta t, t) \, d(\delta x) \qquad (1)$$

This equation encapsulates conditioning over all of the possible queue states $x - \delta x$ at time t from which a transition to the state x at time $t + \delta t$ is possible. The necessary assumption to use this equation is that the transition probabilities of the state of the queue can be described entirely by the function g, regardless of the history of the prior queue states.

If the condensed notation $fg = f(x;t)\,g(\delta x, x; \delta t, t)$ is used, then the integrand of (1) can be expanded as a Taylor series around the point x as follows:

$$f(x-\delta x;t)\,g(\delta x, x-\delta x;\delta t,t) = fg - \delta x \frac{\partial}{\partial x}(fg) + \frac{(\delta x)^2}{2!}\frac{\partial^2}{\partial x^2}(fg) - \frac{(\delta x)^3}{3!}\frac{\partial^3}{\partial x^3}(fg) + \ldots \tag{2}$$

Then (2) is substituted back into (1), and integration and differentiation is interchanged. This presumes, of course that the functions are well-behaved (i.e., bounded).

$$f(x;t+\delta t) = f\int g\, d(\delta x) - \frac{\partial}{\partial x}\left\{f\int(\delta x)\,g\, d(\delta x)\right\} + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left\{f\int(\delta x)^2\, g\, d(\delta x)\right\} - \ldots \tag{3}$$

Since g is a proper density function, then for any choices x, t, and δt, it must be that $\int g\, d(\delta x) = 1$. Hence the first term on the RHS of (3) is simplified, and then f is subtracted from both sides and divided by δt:

$$\begin{aligned}\frac{f(x;t+\delta t) - f(x;t)}{\delta t} = &-\frac{\partial}{\partial x}\left\{f(x;t)\frac{1}{\delta t}\int(\delta x)\,g\, d(\delta x)\right\} + \\ &\frac{1}{2}\frac{\partial^2}{\partial x^2}\left\{f(x;t)\frac{1}{\delta t}\int(\delta x)^2\, g\, d(\delta x)\right\} - \ldots\end{aligned} \tag{4}$$

The limits of two of the elements contained in the RHS of (4) are frequently called the "infinitesimal" mean and variance, respectively:

$$\lim_{\delta t \to 0}\frac{1}{\delta t}\int(\delta x)\,g(\delta x, x;\delta t,t)\, d(\delta x) \equiv M(x;t) \quad \forall x,t \tag{5}$$

$$\lim_{\delta t \to 0}\frac{1}{\delta t}\int(\delta x)^2\, g(\delta x, x;\delta t,t)\, d(\delta x) \equiv V(x;t) \quad \forall x,t \tag{6}$$

A second assumption is made, which is that all of the important information about the transition density function g can be captured adequately in its first and second moments, as in (5) and (6), respectively. This is not a severe limitation; for situations where this is not the case, additional infinitesimal moments can be defined, and the analyst is then responsible for providing that information as well. In fact, in aviation applications, the best contemporary network models, such as LMINET (Lee et al., 1997, 1998) only deal with the propagation of average behavior, and usually with independent Poisson processes at each downstream node. Thus, including $V(x;t)$ is already a step forward. For the present case, assuming that the first two moments suffice, this is tantamount to the assumption:

$$\lim_{\delta t \to 0} \frac{1}{\delta t} \int (\delta x)^n \, g(\delta x, x; \delta t, t) \, d(\delta x) = 0 \quad n \geq 3, \forall x, t \tag{7}$$

Then, taking the limit of (4) as $\delta t \to 0$ and substituting (5) and (6) yields:

$$\frac{\partial f(x;t)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial x^2} V(x;t) f(x;t) - \frac{\partial}{\partial x} M(x;t) f(x;t) \tag{8}$$

Equation (8) is commonly called the *Kolmogorov forward equation* in the stochastic processes literature, or the *Fokker-Planck equation* in the physics literature. In the second case, the term $M(x;t)$ is referred to as *drift*, while the term $V(x;t)$ is called *diffusion*. Equation (8) is the governing differential equation (GDE) for our queuing system.

2.1.2 Boundary Conditions

In this section, the boundary conditions are developed to prevent the model from generating non-zero probabilities for states that are not physically possible, including negative values of the queue length. A similar constraint can be imposed to prevent the possibility of what might be considered unnaturally large queue lengths. The upper bound is more difficult to specify precisely, but it is necessary from a pragmatic standpoint in the numerical scheme because the solution space must be bounded, as will be seen in Section 2.2. It is also practically useful, since air carriers use cancellations and other initiatives to protect against unacceptably long delays.

Because the random variable $Q(t)$ represents a queue length, it makes no sense for it to be negative. Thus, an auxiliary condition is applied that can guarantee that

$$f(x;t)=0 \quad x<0, \quad \forall t \tag{9}$$

This cannot be accomplished by simply saying that (9) must be true; an additional differential equation must be specified that follows the same temporal evolution as (8) , and whose effect is to guarantee that (9) holds. Assuming that the initial conditions obey (9) (as they should, since they are controlled), a way to do this is to guarantee that the "net probability flux" (what would be thought of as the mass flux if this were a problem in physics) across the point $x=0$ is always zero.

A point x is fixed in one dimension and the probability flux across that point in both directions is considered. By integrating all possible increasing transitions that cross this barrier, and subsequently all possible decreasing transitions that cross the same barrier, and then adding them together, lead to the following requirement that the net probability flux be zero. This constraint is referred to in the physics or stochastic processes literature as a *reflecting barrier*.

$$f(0;t)M(0;t)-\frac{1}{2}\frac{\partial}{\partial x}f(x;t)V(x;t)\bigg|_{x=0}=0, \quad t>0 \tag{10}$$

At all times, f must also be a proper density function:

$$f(x;t) \geq 0 \quad \forall x,t \tag{11}$$

$$\int f(x;t)\,dx = 1 \quad \forall t \tag{12}$$

These last two conditions are notoriously difficult to enforce in a numerical solution scheme (Kumar et al., 2006). This is discussed further in Section 2.2.

2.1.3 Initial Conditions

The functions $M(x;t)$ and $V(x;t)$ represent the first and second moments, respectively, of the rate at which the length of the queue is changing at time *t*, given that its current state is *x*. In a queuing system where the arrival process is independent of the service process, then with the possible exception of $x = 0$ and an upper reflecting barrier, there is no reason to suspect that these functions should vary across the *x* dimension. In such situations, it is only necessary to specify how these functions change over time. For most aviation applications, for example, one would expect $M(x;t)$ to be positive at the beginning of the day, negative at the end of the day, and perhaps with some additional cycles in between. One would expect $V(x;t)$ to be small (approaching zero) at the beginning and end of the day and something larger in between, and of course never negative. If this construction was extended to a queuing network, these functions could be derived entirely from the outputs $\{f_i(x;t)\}$ of upstream queues *i*, with some time lags and with some rules for mixing them together.

Although negative queues are explicitly prevented, it also makes sense to preclude initial conditions that would seem in conflict with this goal. Thus, it is required that

$$M(0;t) \geq 0 \quad \forall t \tag{13}$$

At any node to which this method is applied, one can imagine that $M(x;t)$ will be computed as the differential of the difference between the arrival rate, which it might be got from the outputs of upstream processes, and the departure rate, which is related to the capacity of the airport or other resource. This being the case, (13) simply prevents an airport from serving traffic that does not exist.

At some airports, however, the rate of queue growth might depend on its current state. For example, at many airports, runway configurations prevail such that the total capacity of the airport is divided between arrivals and departures, and the airport has some control over that split. In such cases, when there is an excess of arrivals, the airport might choose to emphasize arrivals over departures to ameliorate this queue. This is tantamount to a temporary increase in the arrival capacity of the airport. If this were repeatable and quantifiable behavior, that could be captured in differences in $M(x;t)$ across different values of x.

An initial queue length distribution must be specified. For real airport problems, the queue is empty at the beginning of the day, so one might require:

$$f(x;0) = \delta(x), \tag{14}$$

where $\delta(\bullet)$ is the Dirac delta function. Alternatively, one might consider analyzing a problem starting at some other point in the middle of the day, in which case the restriction (13) is not required.

2.2 Numeric Solution Scheme

In order to solve a system including partial differential equations and their associated boundary and initial conditions, a numerical scheme is necessary to convert that continuum problem into some discrete form appropriate for solution by computer (Pepper and Heinrich, 1992). In this research is presented a discretization method based on the well-known finite element method (FEM) (with some elements of finite differences included as well) that is appropriate for this problem. The construction of numeric schemes for PDEs is very much an art, and certainly a host of other schemes could be attempted, including methods relying entirely on finite differences. The colloquial understanding of the competition between finite element (FE) methods and finite difference (FD) methods is that the former allows for an exact solution to an approximation of the problem, while the latter allows for an approximate solution to the exact problem. Neither is considered uniformly better than the other, and they both certainly have their proponents.

The FEM scheme developed for this problem consists of transforming the governing differential equation with its boundary and initial conditions into linear algebraic equations that can be solved at every time step. This transformation is possible by constructing a discrete approximation to the queue length density function $f(x;t)$ using the N Lagrange basis functions $\phi_1,\ldots,\phi_N$. Each basis function has a triangular shape; the collection of them is illustrated in Figure 2.4 for $N=4$. Mathematically, the basis functions can be represented as follows:

$$\phi_1(x) = \left(\frac{l_2 - x}{l_2 - l_1}\right)_+ ,$$

$$\phi_j(x) = \left(\min\left\{ \frac{x - l_{j-1}}{l_j - l_{j-1}}, \frac{l_{j+1} - x}{l_{j+1} - l_j} \right\} \right)_+, \quad j = 2, \ldots, N-1,$$

$$\phi_N(x) = \left(\frac{x - l_{N-1}}{l_N - l_{N-1}} \right)_+.$$

The approximation for *f* can then be expressed using these basis functions as:

$$f^{L,N}(x;t) = \sum_{j=1}^{N} a_j(L\Delta t)\phi_j(x), \tag{15}$$

where L is the number of time steps Δt, N is the number of Lagrange basis functions, and $\{a_j\}$ are the parameters of the approximation. Using the finite element method, the "solution" of the problem essentially amounts to determining the values $\{a_j\}$.

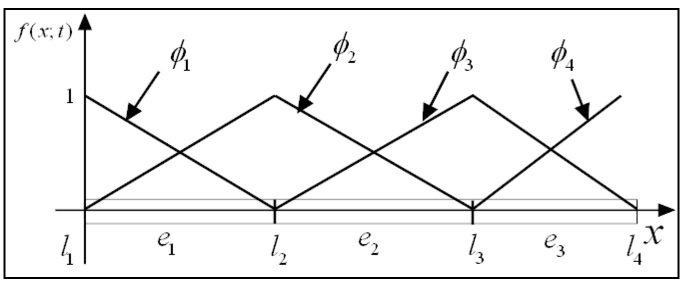


Figure 2.4 Lagrange basis functions for the finite element method

Using a finite difference approach, the left hand side of the PDE (8) can now be approximated by:

$$\frac{\partial f(x;t)}{\partial t} \simeq \frac{f^{L+1} - f^L}{\Delta t},$$

and the dynamics can be re-written as:

$$\frac{f^{L+1} - f^L}{\Delta t} = \frac{1}{2}\frac{d^2}{dx^2}\left(V^{L+1} f^{L+1}\right) - \frac{d}{dx}\left(M^{L+1} f^{L+1}\right) \tag{16}$$

The equation (16) is enforced by defining the residue r, which is essentially the difference between the LHS and RHS of (16),

$$r = \frac{1}{2}\frac{d^2}{dx^2}\left(V^{L+1} f^{L+1}\right) - \frac{d}{dx}\left(M^{L+1} f^{L+1}\right) - \frac{f^{L+1} - f^{L}}{\Delta t}$$

The residue is enforced to zero by using a test function $w(x)$. All of the projections of the residue on w are equate to be zero; i.e., $\int_{\Omega} r\, w\, dx = 0$, where Ω is the domain of interest in x and $\partial\Omega$ its boundary. Integrating by parts yields:

$$\begin{aligned}
&\frac{1}{2}\int_{\Omega}\frac{d}{dx}\left(V^{L+1} f^{L+1}\right)\frac{dw}{dx}dx - \int_{\Omega} M^{L+1} f^{L+1}\frac{dw}{dx}dx + \\
&\int_{\Omega}\frac{f^{L+1}}{\Delta t} w\, dx = \int_{\Omega}\frac{f^{L}}{\Delta t} w\, dx + \\
&\left[\frac{1}{2}\frac{d}{dx}\left(V^{L+1} f^{L+1}\right) - M^{L+1} f^{L+1}\right] w\Big|_{\partial\Omega}
\end{aligned} \tag{17}$$

where the last term on the RHS depends on the boundary conditions.

It is assumed that the interval is closed, and that at the right boundary $x = l$, and is desired the net probability flux to be 0. For some large l, the probability density function will approach 0 for all $x > l$. This will make the net probability flux approach zero at $x = l$, although it cannot be absolutely guaranteed. This is discussed more in the conclusions. Together with equation (10), we conclude:

$$\left[\frac{1}{2}\frac{d}{dx}\left(V^{L+1} f^{L+1}\right) - M^{L+1} f^{L+1}\right] w\Big|_{\partial\Omega} = 0 .$$

The test function w is parameterized with the Lagrange basis functions $\{\phi_i\}$ and parameters $\{b_i\}$:

$$w(x)=\sum_{i=1}^{N} b_i\phi_i(x) \tag{18}$$

The Lagrange approximations of w and f are used to obtain:

$$\sum_{i=1}^{N} b_i\left[\sum_{j=1}^{N} a_j^{L+1}K_{ij}-R_i\right]=0 \quad \forall\{b_i\} \tag{19}$$

where

$$K_{ij}=\frac{1}{2}\int_{\Omega} V^{L+1}\phi_j{}'\phi_i{}'\,dx-\int_{\Omega} M^{L+1}\phi_j\phi_i{}'\,dx+\frac{1}{\Delta t}\int_{\Omega}\phi_j\phi_i\,dx$$

$$R_i=\frac{1}{\Delta t}\int_{\Omega}\phi_i\left(\sum_{j=1}^{N} a_j^L\phi_j\right)dx$$

In the last two equations, is denoted $a_j^L=a_j(L\Delta t)$ and suppressed the dependence of the basis functions $\{\phi_i\}$ on x for the sake of clarity. As mentioned before, it is also assumed that the function $V(x;t)$ is constant in x.

Since the set $\{b_i\}$ is arbitrary, (19) is equivalent to solving the linear algebraic equations:

$$\sum_{j=1}^{N} a_j^{L+1}K_{ij}=R_i \quad \text{for } i=1,2,\dots,N \tag{20}$$

The solution of (20) is the set of parameters $\{a_j\}$ which define $f(x;t)$ according to (15). One of the advantages of the finite element method is the ability to solve these algebraic equations element by element. The N Lagrange basis function approximation defines $N-1$ elements, which makes it possible to solve $N-1$ independent algebraic equations.

The two remaining boundary conditions to enforce on the solution are (11) and (12). As described in Kumar et al. (2006), equation (12) is enforced by scaling the solution appropriately. The non-negativity constraint is harder to enforce. One possible solution is the partition of unity finite element method (PUFEM), described in Kumar et al. (2006). For the time being, however, the problems solved here always result in positively valued density functions, and they solve very quickly, so a more complex solution method is not justified unless that situation changes.

2.3 Model Validation and Results

In this section, some results are showed of applying the modeling with different input data sets. In order to validate this model, Monte Carlo simulation is used as ground truth. The simulation ran a number of iterations (1000 and 10,000) and averaged over this number in order to get the mean and the variance of the queue length. The first set of experiments involves comparisons against the steady state M/M/1 queue. This is not a very useful system for modeling airports, but because the stationary moments are known, it can demonstrate that the diffusion model converges to the proper equilibrium solution, so it is useful from a validation perspective.

Figure 2.5 shows how the results from the diffusion model compare to the results from the Monte Carlo model. The latter results are for an M/M/1 queue with arrival rate $\mu = 15$ aircraft/hour, and a service rate of $\lambda = 40$ aircraft/hour. The traffic intensity is thus $\rho = \mu/\lambda = 0.375$. The equilibrium queue length is then given by:

$$\bar{Q} = \frac{\rho}{1-\rho} = 0.6$$

and the equilibrium variance by:

$$Var(Q) = \frac{\rho}{(1-\rho)^2} = 0.96$$

Both the Monte Carlo and the diffusion results obviously converge to these values, although the diffusion model does so much more smoothly. That is because in this figure, only 1000 replications of the Monte Carlo simulation were conducted, hence a certain amount of noise around the equilibrium values. Figure 2.6 shows similar results for Monte Carlo runs with 10,000 replications instead.

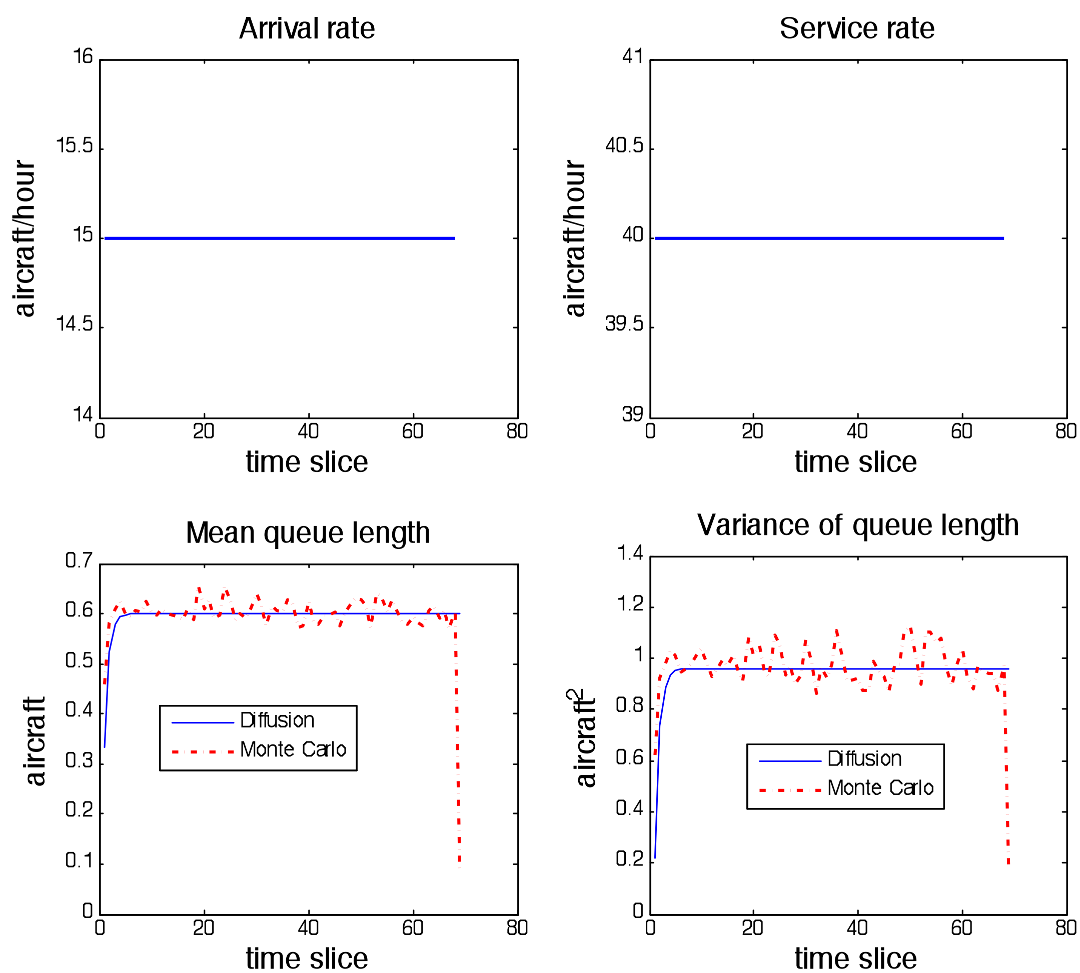


Figure 2.5 M/M/1 queue with 1000 replications, $\rho = \mu/\lambda = 0.375$

Observe that as the number of replications for the Monte Carlo simulation increases, it follows much better the diffusion solution and the equilibrium solution. One important advantage of the diffusion model is the solution time. The Monte Carlo simulation required 10.86 seconds for 1000 runs and 106.9 seconds for 10,000 runs. The diffusion model completes in one iteration, which takes about 8.2 seconds.

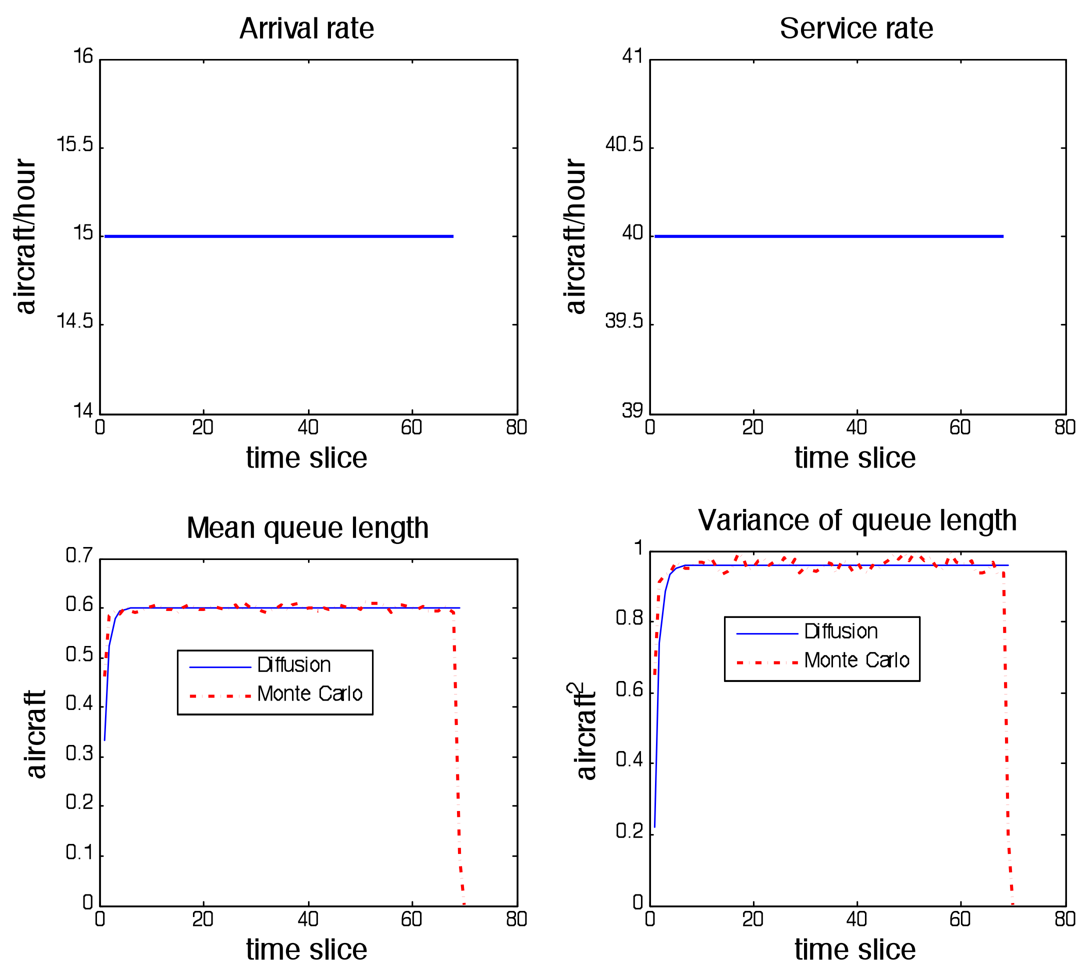


Figure 2.6 M/M/1 queue with 10,000 replications, $\rho = \mu/\lambda = 0.375$

In Figure 2.7, are shown results for some realistic airport demand and supply profiles. In this case, the demand profile is from the published (OAG) schedule for Miami

International Airport, from a peak day in 2007. The capacity profile is a single cluster from a *k*-means cluster analysis on airport arrival rates (AARs), generated using the methodology shown in Liu et al. (2008). These demand and capacity data have been used for previous studies on queuing see for example Hansen et al. (2009). The arrival data show considerable fluctuation over the course of the day, while the capacity profile is nearly flat. The arrival process was modeled as a non-stationary Poisson process, and the service process as a non-stationary Erlang-*k* process, with $k = 7$. This was done for two reasons, first to show that the diffusion model produces good results with different distribution assumptions, and second because this has been shown to be a reasonable model for a single airport server process in other literature (see for example Malone and Odoni, 2001). The Monte Carlo results include 1000 replications. The reason simulation was used to compare our results and not real data, is that real data include propagated delays (Churchill et al. 2008), and that would make the results not directly comparable.

From observation of the figure, one can tell that the diffusion model replicates the Monte Carlo ground truth quite well, in both the first and second moments. This is a very uncongested day, so the mean queue length remains quite low over the entire day.

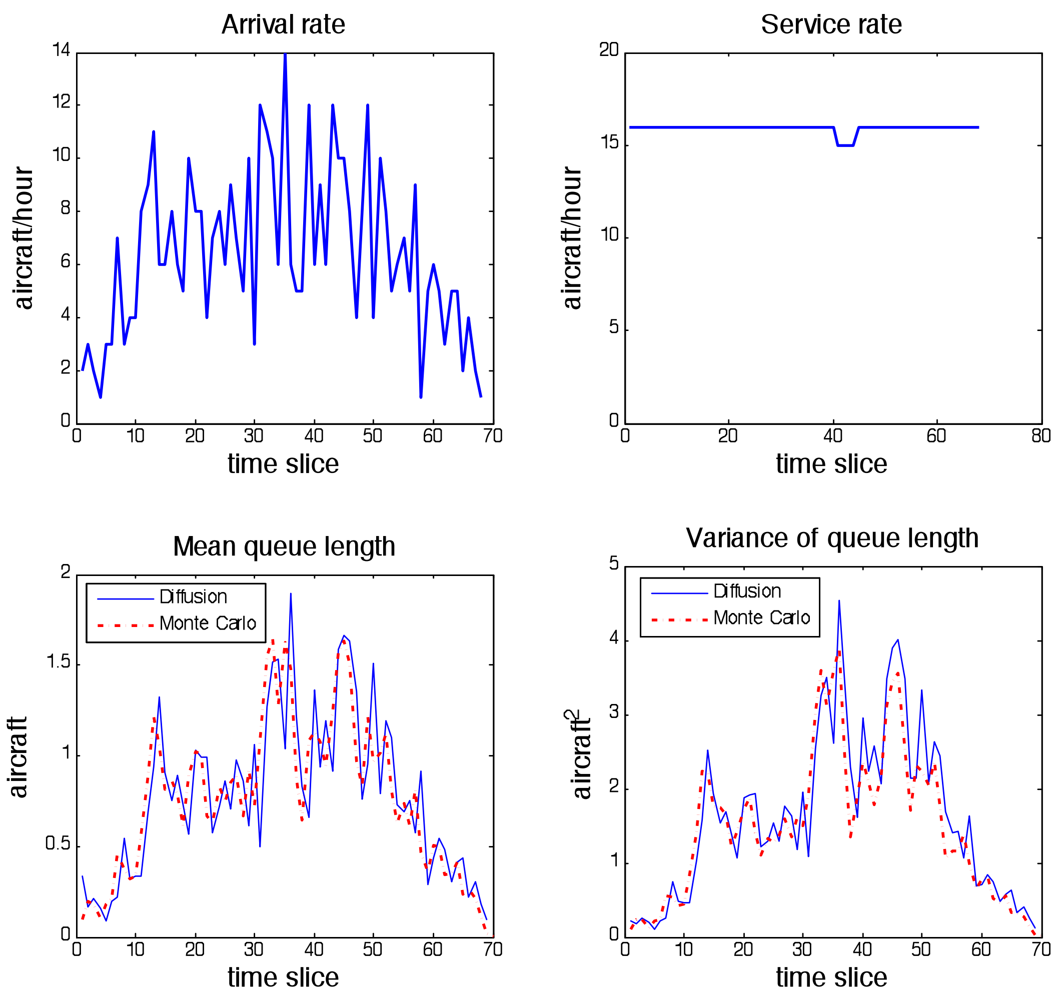


Figure 2.7 Diffusion and Monte Carlo queuing results for Miami International Airport

The final results come from a peak day at Chicago O'Hare International Airport. These results are shown in Figure 2.8. The demand profile is very oscillatory, and it frequently surpasses the capacity over the first three quarters of the day. Thus, larger mean queue lengths are to be expected. The demand subsides towards the end of the day. The smooth oscillations of the service rate between 22 and 23 aircraft per hour are an artifact of specifying integer 15-minute service rates derived from hourly rates that happen not to be multiples of 4. Again, the profiles of the first and second moments of the queue length over time match quite closely between the diffusion and the Monte Carlo models.

For both of these last two sets of results, the Monte Carlo runs complete in about 31 seconds, and the diffusion runs in about 8 seconds. The time required for the Monte Carlo runs is directly proportional to the number of replications, so if more precision were required, for example 10,000 runs, then the run time would be closer to 310 seconds. The diffusion model is immune to these considerations.

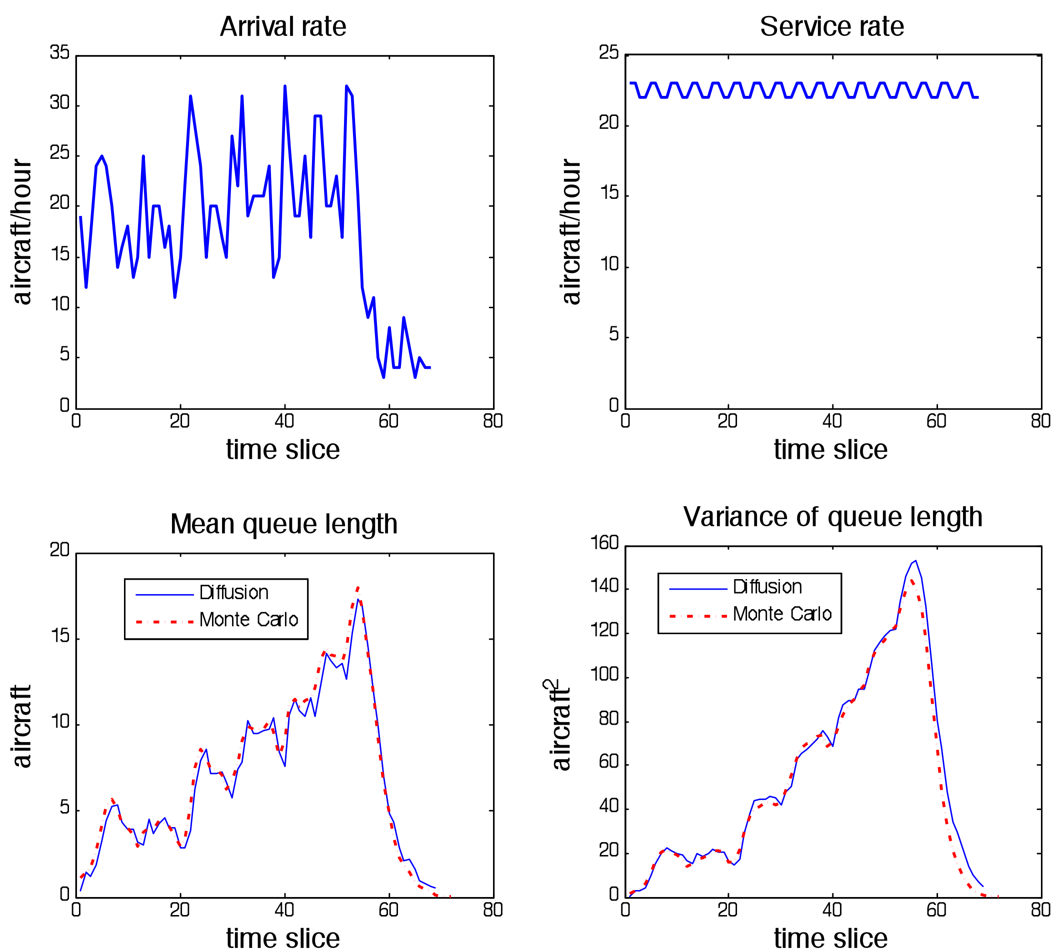


Figure 2.8 Diffusion and Monte Carlo results for Chicago O'Hare International Airport

2.4 Conclusions

The research has presented the mathematical construction of a continuum approximation to a queuing system that might represent a single congested resource in the National Airspace System, such as an airport, a runway, or some en route resource. The result is derived from the diffusion approximation. A numeric solution scheme based on the finite element method is also shown.

The use of this type of approximation requires one to be comfortable with some of the assumptions made in the research, such as the willingness to consider non-integer queue lengths. That notwithstanding, the method has seen considerable application in other areas of queuing theory that also deal with countable objects, so this assumption is not unique to the aviation context.

This result is a stepping-stone in what will hopefully be a larger system of inquiry into the use of such continuum approximations to study systems of aviation queues. In particular, the ability to model the propagation of both the mean and the variance of delay statistics through a connected network would mark a major leap forward in the performance analysis of the aviation system.

In this study a validation effort of the model is conducted. It was feasible to replicate the known steady-state results from that small set of queuing systems for which equilibrium results are known in closed form. The results in such cases showed that the diffusion approximation gives exactly the same results very quickly. Furthermore, a Monte Carlo exercise was also conducted for a number of other cases whose solutions cannot be found analytically. Again, the diffusion model seemed to perform very well, and it is much faster than running large numbers of Monte Carlo simulations.

Chapter 3: Incorporating Airlines' Preferences in Resource Allocation Mechanisms During Irregular Operations

3.1 Motivation for this Research

As mentioned in the first section of this dissertation a couple of new and promising Traffic Management Initiatives are about to start being implemented; i.e. the Collaborative Trajectories Options Program (CTOP) and the Collaborative Airspace Constraint Resolution (CACR). When these systems are fully implemented, there will be a capability at the FAA to allow carrier preferences to affect the allocation of constrained airspace resources. However it is not clear whether (FAA 2011a) carriers will be able to generate full and robust sets of trajectory alternatives on the fly, with associated costs, in response to suddenly changing capacity conditions, and (Ball and Lulli 2004) that such information, even if it could be generated, could be exploited in a systematic optimization of resource allocation. This is the motivation for the work presented in this research. A mechanism is proposed by which simpler, yet still useful, information could be submitted by carriers, and an algorithm is demonstrated that directly employs this information to influence the capacity allocation process. Also extensions are proposed for both the way that airlines can submit their preferences and the resource allocation algorithm. Finally the long run effect of using the proposed allocation mechanism with this preference structure is tested. A version of this work has also been published, in this case in Vlachou and Lovell (2013).

3.2 Feedback from Airline Experts

Part of this research was to find a meaningful way for carriers to express some preference structure - which flight-slot assignments are preferable for them - that would be different than the exchange language used in CTOP (TOS's). Since airlines play an important role in this program, input from experts from the aviation industry and airlines was sought in order to get their feedback and opinions regarding the usefulness of CTOP and more insight into what they would prefer, or what might not be very attractive to them with the existing structure. The latter will be a starting point in order to find a different way for carriers to express their priorities.

A number of people working for airlines were contacted via email and by phone and those who kindly provided their feedback are:

Jim Hamilton – Air Traffic System manager with UPS

George Kypreos – SOC sector manager at American Airlines

Frank Ketcham – Pilot with Delta Airlines and commercial aviation specialist

Don Wolford - Primary instructor for the United Airlines Flight Dispatch ATC coordinator desk and the flight operator lead on the CDM Future Concepts of TFM working group

Mark Hopkins - General Manager, ATM/CDM at Delta Airlines and chairman of the CDM Stakeholders Group (CSG)

There was an agreement in all opinions that the philosophy behind CTOP is something good since it gives some control to the airlines and airlines will benefit from it. Their operations will benefit because they will be able to recover better from delays and cancellations. It will also be more economical, safer and more comfortable for the passengers. Passenger satisfaction is one of the top priorities for all airlines. Another top priority was to keep the schedule intact. Some of the interviewees referred to their business models without revealing of course much information. Some of the things taken into consideration are the passenger connections, crew scheduling, and fuel burn.

From these discussions some concerns were expressed. The first one was about the amount of information that airlines have to share while expressing the various trajectory options sets. Also there was a concern for the extra workload that dispatchers will have to absorb in order to create the trajectory options, probably necessitating the hiring of additional people to perform this task.

Another concern that was brought up is that some airlines will be hesitant to provide information, invest, and participate in this program if there is no analysis that will show them that it is worth the investment, e.g., that it will be a money saver. Finally, another concern had to do with how equitable this will be and how for example international flights and pop-ups will be treated.

3.3 Expressing Airlines Priorities

It was ascertained from some of the interviewees that one of the issues surrounding incorporating carriers' preferences into a collaborative decision-making setting is that, while they would like the final allocation to be sensitive to their wishes, they would prefer not to articulate those wishes in such a clear manner that their internal business models might be discerned. Currently in CTOP airlines have to give great amount of detail about route text, altitude, speed, departure time, as depicted in the Figure 1.8 earlier. For this research, it is proposed to investigate simpler and more obscure ways of expressing preference information, with the hope that they can be just as powerful in injecting carrier preferences into the allocation process. By making the language simpler, this also addresses to an extent the concerns about the extra workload that dispatchers will have in order to create and submit the detailed options, or the updated TOS's. As mentioned above, this will not necessarily mean that the usefulness and the effectiveness of this tool is compromised or diminished. For example, one very simple mechanism that was investigated requires each airline to give to each of their flights a priority number ranging from 1 to 4. The greater the number assigned to a flight the more important this flight is. An extension requires that for each flight, carriers specify the maximum delay in minutes that they would allow it to be assigned on the ground. In Table 3.1 it can be seen an example of how an airline *A,* that had initially scheduled 6 flights to pass through the affected area within a two hour frame, can provide this information. The estimated time of arrival is the time that the flight would have reached the FCA boundary if there were no bad weather in the area.

Here must be mentioned that in addition to the preference language being different, the resource allocation mechanism will also be different (compared to the RBS paradigm that is currently used) in order to make use of the priorities. The details of the algorithm are not defined yet, which will be the next step for this research, but what it will do is to pick the flights with the highest priorities first and assign them to slots. Going back to the example, in this list, airline *A*, for its first flight (f_{A1}), will give a priority number 3, and for its second flight (f_{A2}), a priority number 4, which means that it considers the second flight as more important than the first.

Table 3.1 Example of list with Airline Priorities

Flight	**Estimated Time of Arrival**	**Priority Number**	**Max Delay Allowed (min)**
f_{A1}	3:04	3	35
f_{A2}	3:15	4	25
f_{A3}	3:40	4	23
f_{A4}	3:48	3	32
f_{A5}	4:12	2	50
f_{A6}	4:30	3	33

If the slot assigned to a flight is much later than the initial estimated time of arrival, so that the maximum delay allowed is exceeded then this flight will get rerouted and another flight from the same airline will get the slot. Again this flight will have a higher priority number than the others yet unassigned and its maximum delay allowed will not be violated.

Since the number of flights that are scheduled to pass through the FCA is reduced, each airline will have fewer flights passing through that area for the duration it is expected to last. Since some of their flights will be more important than others, airlines have no reason to claim that all their flights have the same high priority number. Also it is important to understand that airlines do not compete with the other airlines for a specific slot, so there is no point in submitting fallacious information. The priority numbers for flights are used to sort the flights of each airline separately

A flight of an airline with priority 4 does not compete with a flight with priority 4 of another airline. Importantly, such a scheme would not be useful in trying to game the system into ensuring that all their flights will go thought the FCA. Some of the flights will be pushed back in time and if they have intentionally given big priority numbers to flights that are not important, they essentially allow less important flights to pass the FCA before their more important ones. If all the flights are given the same priority number and high allowable delay, then the slots would end up being assigned following the Ration-By-Scheduled (RBS) method, as it is being currently used. This essentially would cancel the allocation mechanism proposed here and the potential benefits for the airlines to give priority to flights they are more important. Also, if the maximum allowed delay is set unreasonably small, then the flights – especially the ones further down the duration of the program - might not be able to be assigned to a slot. This will cause slots to go unused.

The times of the slots are not known in advance, so airlines cannot request specific flight-slot assignments. By giving a priority number and a maximum delay allowed, the airlines are given the chance to prioritize their flights without the need to reveal any information of why one flight is more important than the other. At the same time they are given some flexibility as to which slot they can get. For example from the above table it can be seen that the first flight was initially scheduled for 3:04 and is given a 35 minute allowance of delay, which means it will be considered in the system if it is given a slot before 3:39. For illustration purposes, assume that available slots might be at 3:10, 3:15, 3:20, 3:25, 3:30 and 3:35. This means that this flight can have a number of possible slot assignments that will be within the desired time window.

There is always the option of substitutions among the flights of the same airline after slots have been assigned to flights. One advantage of prioritizing the flights with the way suggested here is that it may reduce the need and the number for substitutions. Looking again at the example, flight f_{A2} has priority number 4 and flight f_{A1} has priority 3. With the allocation mechanism that is considered developing, f_{A2} will get picked first to get the first available slot for that airline S_{A1} and then flight f_{A1} would get the second one S_{A2}. Without this prioritization scheme the first flight f_{A1} would get S_{A1} and flight f_{A2} would get slot S_{A2}, and the airline later would have to ask for the swap.

3.4 Allocation Mechanism

Assuming that the FAA has a list of all flights determined to be affected by an FCA, and also has the accompanying preference information for those flights garnered from their respective carriers, the next proposed step is an allocation mechanism by which a subset of those flights would be allowed to use the FCA, and the provided information would play a role in that decision. It is not declared explicitly what happens to flights not captured in the AFP – carriers could choose to re-route them around the FCA, cancel them altogether, or re-schedule them to use the airspace in question at a later time. If enough flights elected to take extended ground delays, then presumably the FAA would have to extend the FCA or create a new one, as long as a demand-capacity imbalance continued to exist.

The allocation mechanism for airspace slots proposed by Vakili (2009) and is the basis for the allocation portion of this work. She had proposed an allocation scheme, called Preference Based Proportional Random Allocation (PBPRA), which was proven to be fair, equitable, and immune to gaming, and she used a different way for airlines to submit their preferences. The PBPRA is a two-step process:

Step 1: Determine the fair share of the constrained resource set for each carrier, using the original schedule as the basis of fairness

Step 2: Allocate flights to slots in a manner consistent with the fair share determined in Step 1.

The allocation mechanism tested in this research is consistent in motivation and basic construction with PBPRA. In the first step, the amount of claim is determined that each carrier has on each available slot – in other words, a number that should, in the long run, be proportional to the number of times that carrier is allocated that slot, under identical circumstances.

For example, consider the case where several carriers have flights, not previously assigned to slots, which can feasibly reach the FCA in time for a given slot. Each carrier can be thought of as having some "claim" on that slot. In trying to assign a numeric value to the claim of a particular carrier, one might consider allowing that number to depend on such things as the total number of flights that carrier has scheduled through the FCA that can feasibly use that slot, the number of flights owned by that carrier that were eligible for previous slots but were not assigned, and so on.

Here is an example of how fair share is computed. Let's assume that there are 6 ($f_{X,i}$) flights $\left[f_{A,1}, f_{B,1}, f_{A,2}, f_{B,2}, f_{C,1}, f_{C,2}\right]$ from 3 different (X) carriers $\left[A, B, C\right]$ that are scheduled to arrive at the boundary of the FCA in the following times $\left[222, 228, 234, 238, 242, 245\right]$ respectively. The time of available slots (S_j) due to reduced capacity are $\left[S_1, S_2, S_3, S_4\right] = \left[230, 235, 240, 245\right]$. The earliest slot that each flight can be assigned to is $\left[S_1, S_1, S_2, S_3, S_4, S_4\right]$. Then the total number of flights n_m that can be assigned to each slot is estimated by $n_m = \sum_i N_{i,m}$, where

$$N_{i,j} = \begin{cases} 1, \textit{if flight i can be assigned to slot j} \\ 0, \textit{otherwise} \end{cases}.$$

Then the share of each flight for each slot is computed by $Share_{sj}^{f_{X,i}} = N_{i,j} * \dfrac{1}{\prod_{m=k}^{j}(n_m - (m-1))}$, where k is the earliest slot the flight $f_{X,i}$ can be assigned to and n_m is the number of flights that can be assigned to the respective slot. For example the share of the first flight $f_{A,1}$ for the second slot S_2 is $Share_{s_2}^{f_{A,1}} = \dfrac{1}{(2-(1-1))\times(3-(2-1))} = \dfrac{1}{4}$. Then the next step is to find the total share of each flight to all slots and by adding these shares, the total fair share for each airline is computed. For this example, the final total fair share (FS_X) for each airline is $\left[FS_A, FS_B, FS_C\right] = \left[1.75, 1.58, 0.67\right]$.

A carrier can only have a claim for a slot if it has flights that can feasibly reach the edge of the FCA in time for that slot. Furthermore, is not allowed a flight to count towards a carrier's claim for a slot if the amount of delay required to fit that particular flight into that slot exceeds the carrier's declared maximum allowable delay for that flight. This is another departure from the PBPRA algorithm described in Vakili (2009). Thus, for each slot, is identified which flights are feasible, and therefore which carriers have some claim to that slot.

This mechanism ensures that all slots will be assigned to flights, as long as the maximum allowable delay is set reasonably. Additionally, since the number of available slots is less than the number of scheduled flights, the fair share for each airline will be a smaller number than their initial number of flights. The fair share is likely not to be integer-valued, as was shown in the example above.

When this initial step is completed, the fair share, or claim, that each carrier has on each slot has been computed. A carrier with many flights scheduled through the FCA might have a claim of something like 2.4 on a slot later in the program, while a carrier with less presence might have just a fractional claim of 0.25, for example. If this situation were to repeat itself over time, it would be expected the carrier with the 2.4 claim to be assigned the slot much more often than the one with the 0.25 claim (approximately 10 times as often). Because the smaller claim is not zero, however, the algorithm does not allow that carrier to be systematically denied that slot.

In the next step, slots are assigned to carriers, and then to specific flights. It begins by allocating only the fractional shares; the integer shares are allocated later. There is a specific reason for doing so. A carrier with a very small share to a slot, by virtue of having very few flights captured in the program, would be very unlikely to receive that slot by way of a deterministic allocation mechanism. As a result, there would be a systematic bias against that smaller carrier, particularly since FCAs tend to appear (or be declared) with some geographic regularity, due to recurrent congestion or recurrent weather patterns. Since there are strong geographic patterns to carriers' networks, this conspires to produce noticeable patterns in the extent to which carriers are represented in AFPs. This process of considering fractional shares first, coupled with the proportion-based assignment described below, obviates this bias.

A random process is used to assign a slot to a carrier, where each carrier's probability of being chosen is proportional to its fair share. Thus, a carrier's expected return is equal to its fair share, while its actual return in any particular AFP may differ from that. Again, in the first step is considered only the fractional shares. Once all carriers' shares have been reduced to integer values, the algorithm continues, considering their remaining integer shares until all slots have been allocated.

Once a carrier has been allocated a slot, the particular flight allocated to that slot is the one with the highest priority (as stated by the carrier itself) that is feasible for that slot. That flight is then removed from further consideration, as is that slot, and the process continues. This process ensures that carriers' preferences play a role, but not in such a way that there would be any advantage to gaming the priority numbers claimed by the carriers – they are only used to measure relative worth within a given carrier's own stable of flights. This is a further departure from the mechanism of Vakili (2009), which did not consider carrier priorities.

3.5 Trade-Off Between Slot Quantity and Minimizing Delay

For the purpose of this research, is also proposed an enhanced version of this allocation scheme, which is called Alternative Preference Based Proportional Random Allocation (A-PBPRA) and this is another contribution of this thesis. This allocation mechanism will also be a two-step process, and the first step is identical to PBPRA. In this scheme, however, each airline is allowed, when declaring its preferences, to also declare its intent to be considered as one of two different kinds of airlines: those that would prefer getting earlier slots (at a cost of depleting their fair share faster), and those that would prefer getting a larger number of slots overall, with the understanding that some of those will likely have large delays associated with them. For a particular AFP, this would allow carriers with a nearby hub and some higher priority connecting flights, for example, to choose to be treated differently than a regional carrier whose main concern is keeping the breadth of their schedule and moving their airframes to the next intended destination. So this consists of an additional airline preference input along with the flight priority number and the maximum allowable delay.

This modified scheme can be related to the first proposed scheme by imagining that in the first scheme, the "price" of each slot is one point of fair share, while in this second or modified scheme is allowed that "price" to be higher for carriers that prefer to be allocated a smaller number of premium slots. The allocation mechanism is modified so that when an airline of this type is chosen to be assigned a slot, it is also assigned to the slot that its next most valuable flight can use, albeit at a significant cost (for example two units of fair share).

As mentioned above, airlines always have the option of substitutions among their own flights after slots have been assigned to flights. It might be the case that, after looking at the resulting allocation, a carrier has a slightly different view of its relative priorities than it did when it first submitted its preferences, and it would be free to make swaps amongst its own assigned slots if it saw fit and if those swaps were feasible.

3.6 Results

The two allocation schemes (PBPRA and A-PBPRA) are compared to two schemes that can be thought of as representative of how AFPs are currently handled. One is a straightforward application of RBS, which is a presumption of what the FAA would do if it had no information on preferences whatsoever. The second is a proxy for RBS with intra-carrier substitutions allowed (RBS with substitutions). This process is mimicked by making first an RBS allocation, and then use the flight priorities to make an optimal assignment of flights to slots within each carrier's holdings.

The methodology is tested with a hypothetical AFP with a realistic sized FCA and a realistic capacity reduction. The AFP time is set to be 2 hours, and tested 3 different scenarios for capacity reduction: 25, 20, and 15 aircraft per hour, respectively. This is out of a nominal flow of approximately 30 aircraft per hour. Hypothetical carriers were invented, and in order to properly represent the variety of share sizes that each might have in an AFP, actual data were considered on flights into the Boston Logan airport, from the Bureau of Transportation Statistics (BTS 2012). From this the mean number of flights per airline was found, which is shown in Table 3.2.

Table 3.2 Average Number of Flights per Airline

Airline	Average Number of Flights Per Airline
1	3
2	2
3	6
4	5
5	10
6	16
7	6
8	11
9	4
total	63

Because each of the proposed schemes ensures that all AFP slots are utilized, there is no difference between these mechanisms and any version of RBS that differs in the efficiency with which the capacitated resource is used. It is not claimed that flights that are not given a slot in the 2-hour time period of the allocation mechanism are necessarily re-routed, delayed further, or cancelled, either for the proposed mechanisms or for the RBS mechanisms to which they are compared. The dispositions of these flights are unknown under all allocation schemes. The performance metrics, therefore, relate only to those flights captured in the 2-hour period of the AFP.

First the results from PBPRA are compared with the results from RBS and RBS with intra-carrier substitutions. The metric chosen for the comparison is the total weighted delay, which is computed by multiplying the delay for each flight by its priority number and summing across flights. For example, if a flight with priority number 3 was delayed for 9 minutes, its weighted delay would be 27. Since the PBPRA mechanism is a randomized procedure, each scenario is ran 100 times in a Monte Carlo simulation in Matlab and the average for each metric used is computed.

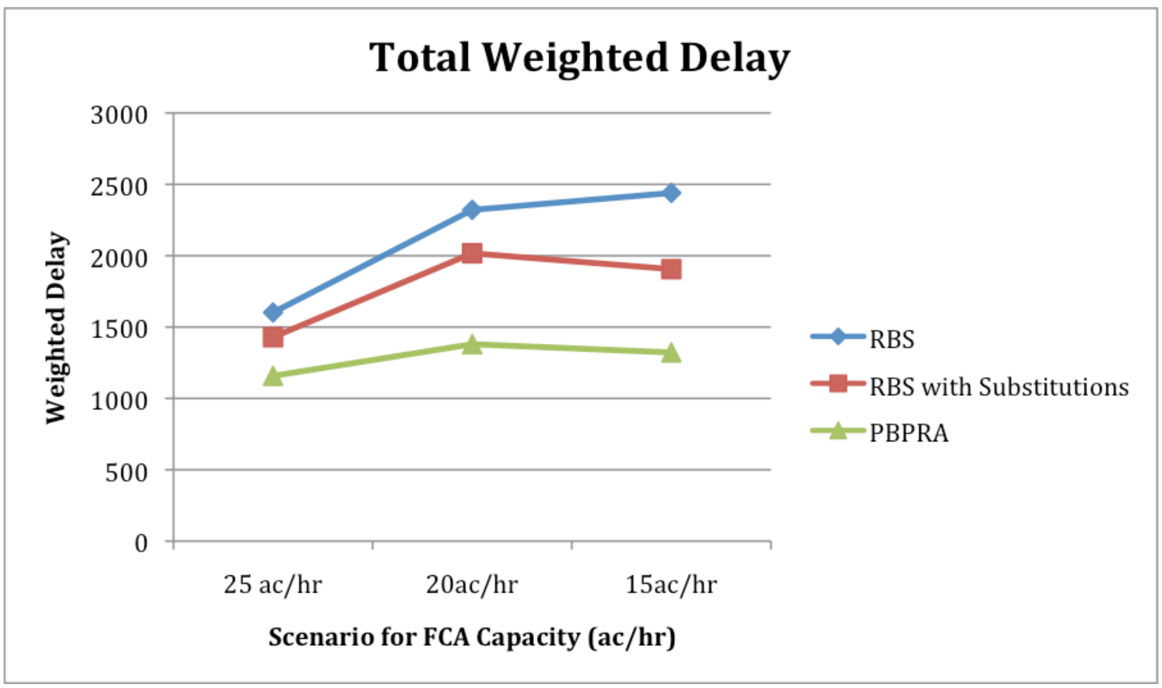


Figure 3.1 Total weighted delay for RBS, RBS with substitutions and PBPRA for each capacity scenario

In Figure 3.1 the total weighted delay accrued by the number of flights that were assigned to pass through the FCA during the two hour period is always less for

PBPRA compared to RBS and RBS with substitutions. RBS with substitutions performs better than RBS since flights with higher priority get earlier slots than initially assigned. With PBPRA not only are higher priority flights assigned earlier, but they are also assigned much closer to the desired slots in order to minimize the total delay.

To see how the allocation schemes compare, Figure 3.2 presents the weighted (by the priority number of each flight) average delay in minutes per flight for RBS, RBS with Substitutions and PBPRA. As capacity decreases the weighted average delay for RBS increases rapidly, for RBS with substitutions at a lesser rate, and for PBPRA at an even smaller rate. The weighted average delay for PBPRA is consistently lower than the other two schemes.

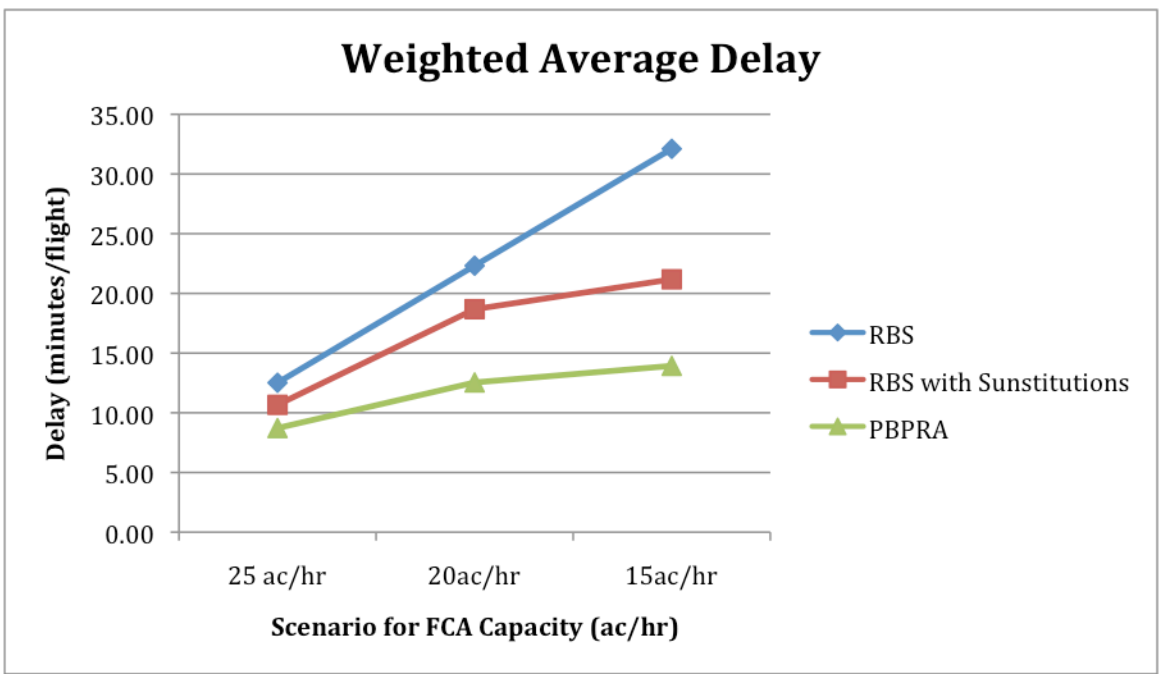


Figure 3.2 Weighted average delay in minutes/flight for RBS, RBS with Substitutions and PBPRA for each capacity scenario

For the second part of the analysis two different sets of airlines are considered. In the first set are included the airlines that are willing to pay more to get an extra early flight in the beginning of the allocation process and in this case is assigned Airline 6 and Airline 8 from Table 3.2 to belong to this set. The other set consists of the rest of the flights that do not want to pay extra and have a chance of getting more later slots. Also for this analysis since the A-PBPRA mechanism is a randomized procedure, each scenario ran 100 times in a Monte Carlo simulation in Matlab and the average for each metric used is computed.

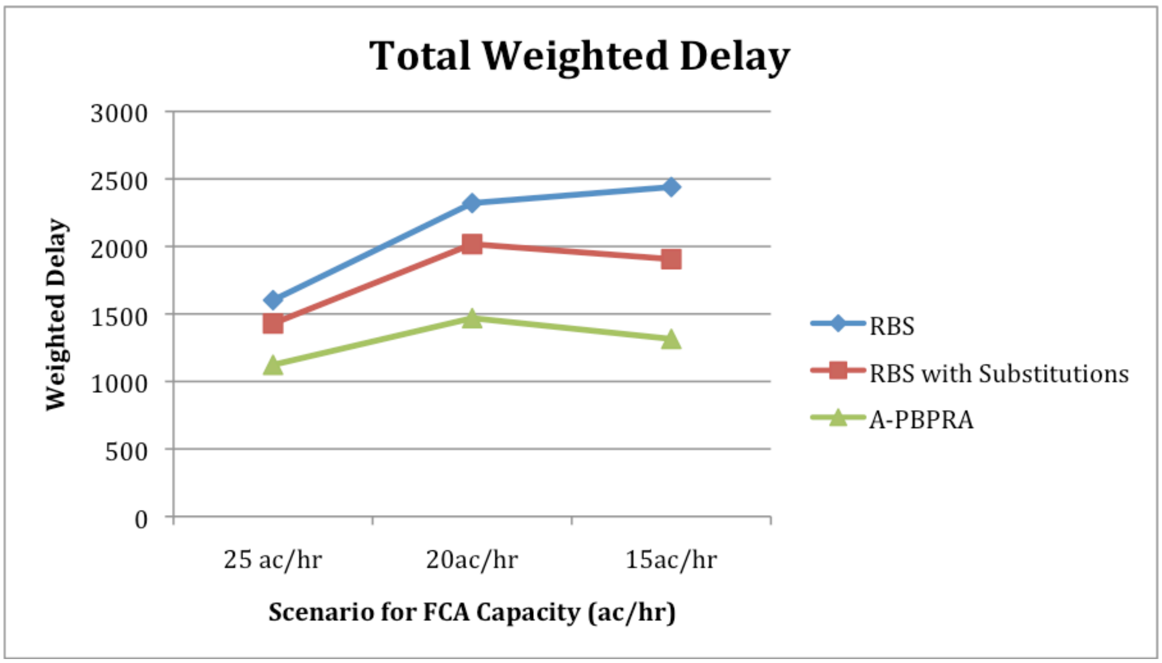


Figure 3.3 Total weighted delay in minutes for RBS, RBS with substitutions and A-PBPRA for each capacity scenario

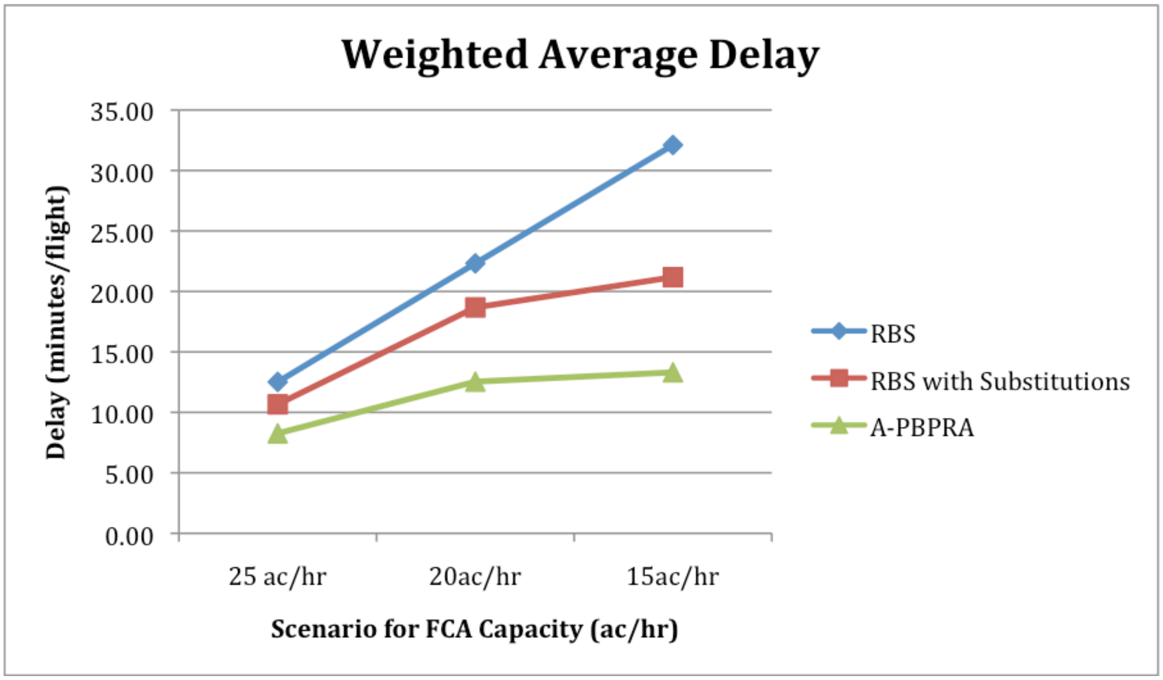


Figure 3.4 Weighted average delay in minutes/flight for RBS, RBS with substitutions and A-PBPRA for each capacity scenario

Figure 3.3 compares the total weighted delay accrued from all flights assigned to slots for the duration of the FCA, as determined by the Alternative PBPRA (A-PBPRA) mechanism, RBS, and RBS with substitutions. The total weighted delay for each capacity scenario with A-PBPRA is consistently lower than total weighted delay with RBS and RBS with substitutions. Figure 3.4 presents the weighted average delay for each flight that gets a slot. This is much lower for A-PBPRA compared to RBS and RBS with substitutions, especially for the cases for which the capacity is greatly reduced.

Overall, both proposed allocation mechanisms – PBPRA and A-PBPRA – perform better than RBS and RBS with substitutions, with regard to ensuring that carriers' higher priority flights are indeed treated better than lower priority flights. The total weighted delay for both mechanisms is consistently much lower than RBS and lower than RBS with substitutions. Also, each flight with the proposed allocation mechanisms has to suffer less delay than RBS and RBS with substitutions, since it is assigned to the closest possible slot in order to minimize its delay.

3.7 Estimating the Long Run Effect of Preference Based Proportional Random Allocation

During an AFP the number of available slots is less than the number of flights scheduled to pass through the FCA, which means that these slots must be divided in a fair way among the carriers. So it is important to have an equitable resource allocation mechanism to do so. But even then, on a given day the slots allocated to an airline will not match exactly its fair share. Some days they will get more and some others less, so it is important to see if in the long run they will get on average what they want. If the difference of the fair share from the actual allocation is considered as an error, another goal of this research is to measure this error.

Also another aspect of this problem is the variety in the sizes of carriers, or more precisely, the number of flights they have planned through the FCA. This does not stem only from the size of the carrier itself, but also takes into account the fact that FCAs are geographically specific, and carriers have definite geographic patterns with which they operate, regardless of their size. Nevertheless, in a given FCA, there will be "big" airlines, which will have many of their flights planned to pass through the affected area, and there will also be "smaller" ones with fewer flights. It would be interesting to see what would be the impact of allocation procedures on those two different categories of carriers. For example, there might be cases where an airline has one flight scheduled to pass through the FCA. This means that its share in the available slots will be less than one, which means the resulting slot allocation might omit this carrier altogether. It might be “fair” to do this on some fraction of days, but certainly not on every day, if the underlying pattern were to be repeated. So equity between carriers is a major concern, and perhaps more so for those who would expect to have a small presence in the schedule affected by a given FCA.

As mentioned before, one of the goals of this research is to examine if this way of expressing priorities is valid. It was desired to see if at the end most of the higher priority flights have been assigned to slots. Also another goal was to check if the delays occurred by the highest priority flights are less than the rest. It is requisite airlines to give truthful preferences for which they will be more willing to do if they actually see that it makes sense delays-wise for some of their flight to have higher priority. If airlines give a 4 in all their flights on purpose to game the system, and since they will not be able to assign all these flights to slots, they might miss the opportunity of actually a flight that is in reality more important than others to get a slot.

The delays accrued by flights that were given the highest priority number 4 are penalized more. The delay of these flights are multiplied by 4. Respectively the delays occurred by flights given priority number 3 were multiplied by 3, those with priority number 2 multiplied by 2 and those with priority 1 stayed as were. In this analysis was estimated the average delay per flight per priority number. The flights that essentially were accounted for their delays, were the ones that actually got allocated to a slot. The flights that did not assigned to a slot, for which each airline will decide whether to reroute, cancel, or delay them more, were not included in the calculations. Another thing also examined is the average delay per flight per airline and in this calculations the delays were not weighted according to the priority number assigned to them.

In this problem there are two levels of randomness that can be identified and examined. The first is due to the random selection of airlines in the allocation algorithm. Even when the number of flights and slots stay the same, each time the allocation scheme is implemented the selection of airlines can differ. The second level of randomness comes from the fact that each time an AFP is implemented the number of flights that each airline has will vary. For the purpose of this research and in order to be able to have multiple repetitions of the allocation mechanism with varying input data, simulation was used. Simulation is a very fast and reliable tool for analysis of this kind.

The first thing measured from the simulation output was, for each airline, how much variance there was in the number of assigned slots. Since each airline does not get the exact same number of slots each time, it would be desirable for these numbers to be quite close. For example, if an airline gets on average 10 slots but one time gets 5 and the next time 15, is this something that will not be easily acceptable from the airline's dispatchers.

Also it was of interest to see how much on average the slots assigned to the airlines deviated from their initially calculated fair share. In order to measure the deviation of the actual allocation to the fair share of each airline the following indicator was used:

$$Fair\,Share\,Deviation = \frac{Computed\,Fair\,Share - Actual\,Share}{Number\,of\,Slots}$$

The fair share deviation indicator, as it was similarly used by Carr et al. (1998), shows how much the actual average share (number of slots allocated to each airline) deviates from the fair share estimated before the allocation. If the numbers are zero then the actual share matches the fair share computed. The bigger this number becomes (positive) the actual share deviates and essentially the particular airline receives less slots than its fair share. If the number is negative the particular airline receives more slots than its fair share. What would be desirable is these numbers to be very close to zero.

For the purposes of this research simulation was divided into two different parts, in order to isolate the variance that will appear in the outputs into two different sources. In the first part, was tested the effect that the random allocation procedure has by itself. In order to do that, a deterministic set of input data was used, so the only randomness in the simulation is in the proportional allocation procedure itself. In the second part, recognizing that there are stochastic fluctuations in demand input data due to various causes (e.g., schedule changes, seasonality, unexpected cancellations and delays due to crew issues and maintenance, etc.) randomness to the input data was added.

3.7.1 Monte Carlo Simulation for Deterministic Set of Flights

In the first set of simulations was considered a deterministic set of input data to our Monte Carlo simulation. Rather than work with a particular geographic scenario and its associated FCA, a hypothetical AFP was used whose magnitude is commensurate with what tends to be observed in reality. The exact same data set used in the previous analysis was used here also, and the mean number of flights per airline is shown in Table 3.2. Each simulation ran for 1000 replications and the results are presented in the following figures and tables.

In Figure 3.5 can be seen the fair share for each airline, computed as described in section 3.4. After the allocation procedure ran, the average number of slots that each airline actually received was estimated, as can be seen in Figure 3.6. A first comparison of those two figures shows that, in the long run, airlines will be assigned numbers of slots that are close to their estimated fair share.

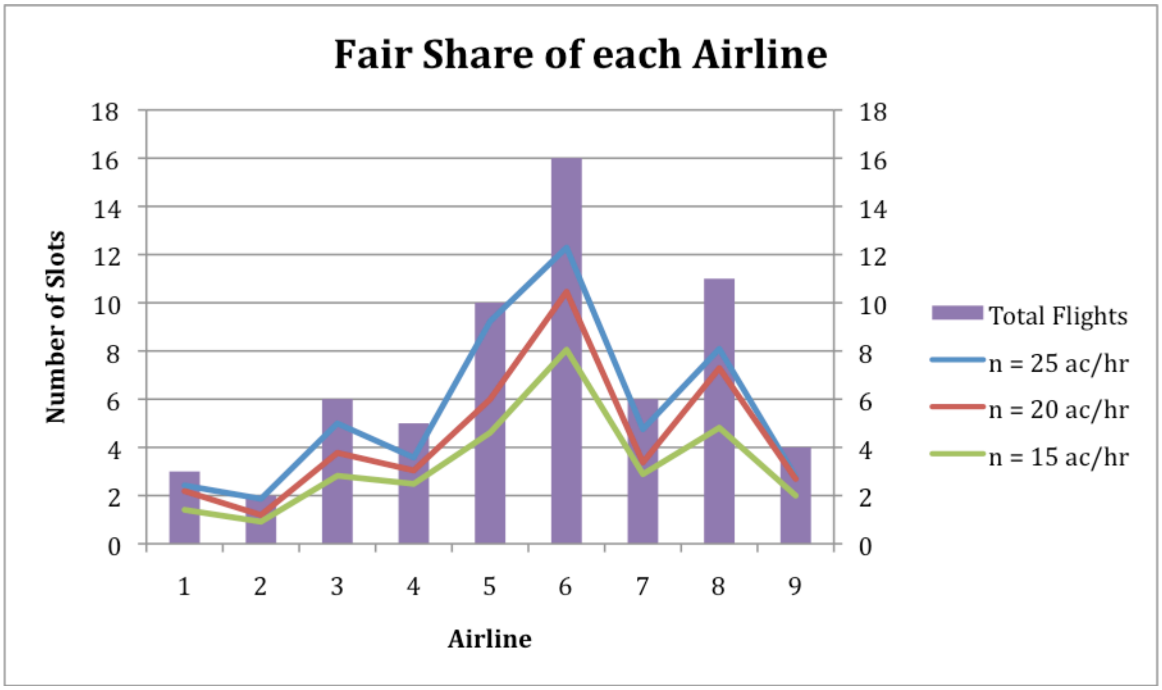


Figure 3.5 Computed fair share for each airline and for each capacity reduction scenario

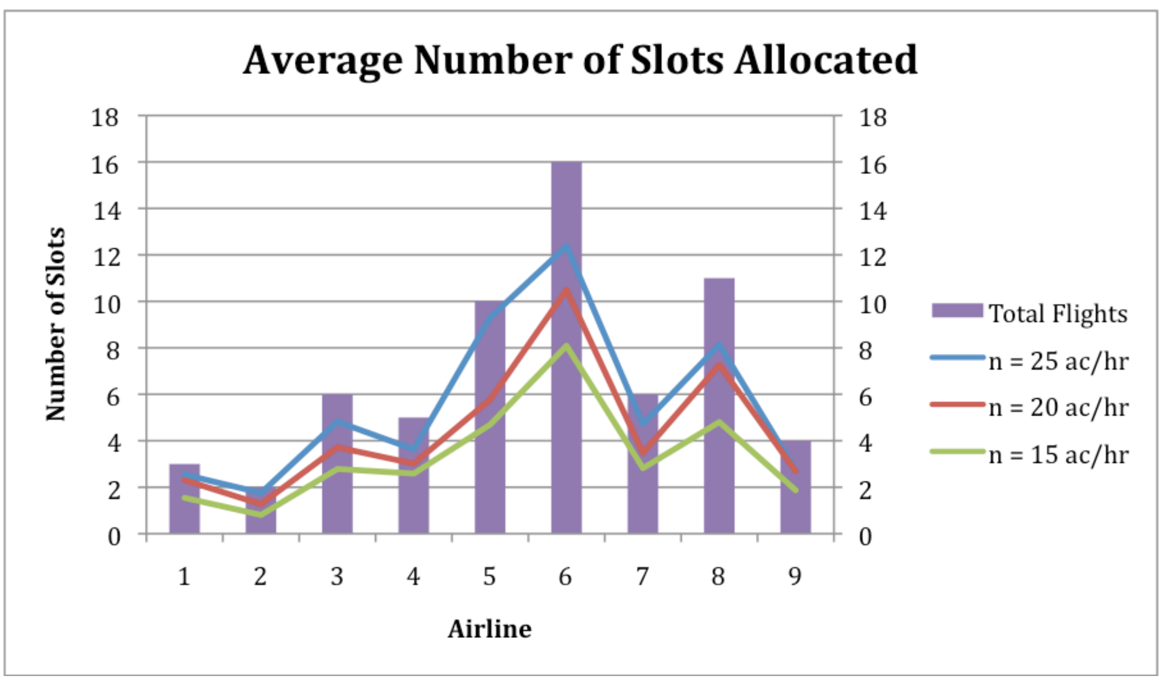


Figure 3.6 Average number of slots for each airline

The median number of slots for each airline is also presented in Figure 3.7, to have a better idea of how many slots airlines will usually get. Here it is clearer that smaller airlines have a good chance of getting slots. In Table 3.3 the coefficient of variation of slot allocation for each airline is presented. The coefficient tends to be smaller for bigger carriers. Probably this is caused by the fact that the resource allocation mechanism is specifically designed to be protective of small airlines' slot claims. In general the coefficients for all airlines are very close, which is a good indication of the fairness of the allocation process.

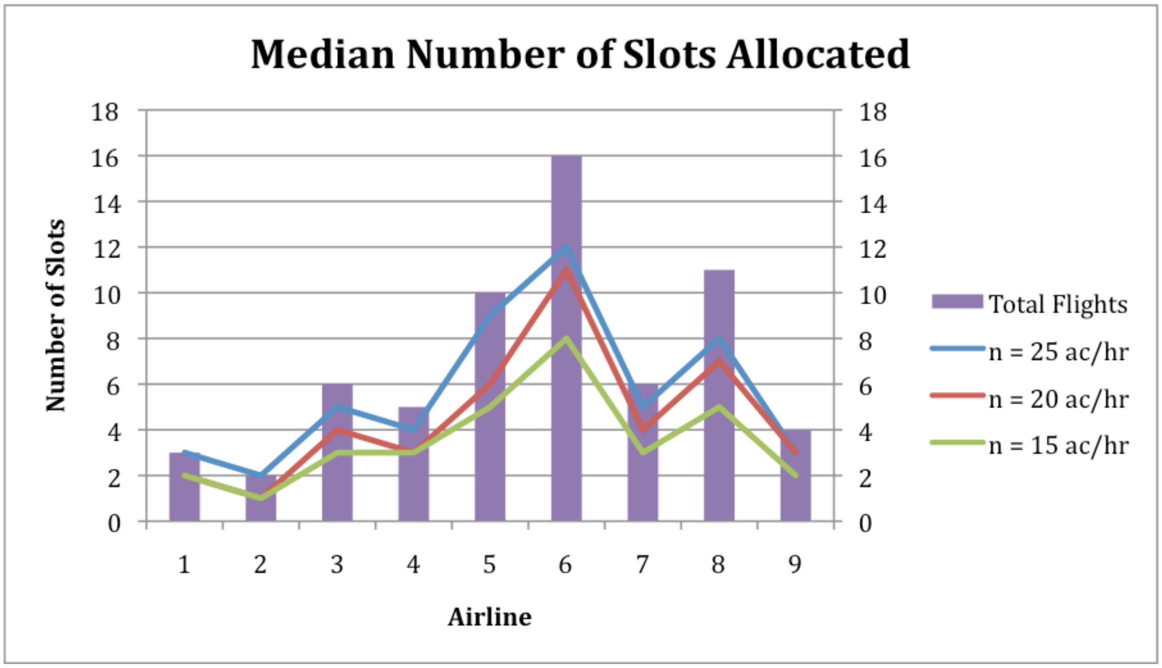


Figure 3.7 Median number of slots for each airline

Table 3.3 Coefficient of Variation of Slots Allocated to Each Airline

		Capacity Reduction to n ac/hr		
		25	**20**	**15**
Coefficient of Variation of Slots for Each Airline	**1**	0.1962	0.2229	0.3227
	2	0.2457	0.4366	0.4893
	3	0.0799	0.1301	0.1473
	4	0.1330	0.1153	0.1903
	5	0.0493	0.0768	0.0982
	6	0.0392	0.0554	0.0350
	7	0.0955	0.1607	0.1366
	8	0.0413	0.0867	0.0825
	9	0.1622	0.1848	0.1806

In Table 3.4 the results for the fair share deviation indicator can be seen. As explained before, the fair share deviation indicator shows how much the actual average share deviates from the fair share estimated before the allocation. As can be seen in this table, the indicator for each airline is quite close to zero, which means that in the long run airlines will receive numbers of slots that are very close to their fair share.

In Table 3.5 the average delay per flight per airline occurred by flights that actually got assigned to a slot is presented. Can't say a clear trend exists here. For some airlines as the number of slots available gets smaller, their average delay increases and for some other decreases. The smaller the airline the delays decrease, because they get fewer flights, so fewer flights are included in the calculations. The second airline has the least flights and when the capacity is reduced to 15ac/h, many times it wouldn't get a slot at all, which means the delay is accounted as zero. There isn't any extremely big difference in the delays among the airlines.

Table 3.4 Deviation of the Fair Share from the Actual Average Allocation

		Capacity Reduction to n ac/hr		
		25	**20**	**15**
Deviation of the Fair Share from the Actual Average Allocation	**1**	-0.0024	-0.0029	-0.0046
	2	0.0022	-0.0019	0.0036
	3	0.0035	0.0013	0.0012
	4	-0.0011	0.0007	-0.0034
	5	-0.0013	0.0055	-0.0028
	6	-0.0017	-0.0008	-0.0012
	7	0.0007	-0.0025	0.0023
	8	-0.0008	0.0006	0.0004
	9	0.0010	0.0000	0.0043

Table 3.5 Average Delay (in Minutes) per Flight per Airline

		Capacity Reduction to n ac/hr		
		25	**20**	**15**
Average Delay (in min) per Flight per Airline	**1**	17.5	25.9	20.4
	2	9.9	9.9	3.5
	3	6.1	15.5	22.4
	4	10.3	22.6	17.1
	5	4.0	13.1	17.8
	6	7.9	10.4	14.3
	7	11.5	18.2	17.9
	8	7.5	12.9	26.0
	9	18.0	12.9	14.2

Finally in Table 3.6 the results of the weighted average delay per priority number are presented. Although the delays of flights with priority given equal to 4 were weighted more, the average delay per flight is consistently less than the average delay of flights with priority 3. From the simulation was observed that most of the flights assigned to slots were of priority 4 and 3 and consequently most of the flights left unassigned had a priority of 2 and 1. This explains the fact that the delays for flights with priority 2 and 1 are less than the ones with higher priorities. It is not that they were assigned to slots that were closer to the initial scheduled times, but they weren't assigned to any slot at all. The fact that many flights with priorities 2 and 1 are excluded by the program and do not contribute to the metrics, makes it trickier to compare the numbers.

Table 3.6 Weighted Average Delay (in Minutes) per Flight per Priority Number

		Capacity Reduction to n ac/hr		
		25	**20**	**15**
Weighted Average Delay (in min) per Priority Number	**4**	16.4	35.3	67.2
	3	22.4	47.6	76.6
	2	21.5	28.8	13.6
	1	15.5	20.4	12.1

3.7.2 Monte Carlo Simulation for Random Set of Flights

For the next part of the simulation, variability to the input data (the number of flights for each airline) was added. In reality, the schedule of airlines fluctuates, and since it was a goal to comply with that, the traffic at Boston Logan for every Monday of two consecutive months was observed – February and March of 2011. For the same airlines as mentioned before, the number of flights was observed and incorporated similar levels of variation into the simulation. The mean number of flights matches the number of flights used in the previous set of experiments, but now the number of flights for each airline for each run varies following a uniform distribution whose extremes are identical to what was observed from the real data.

In Figure 3.8 is present the average fair share computed for each airline. Since the number of flights for each airline fluctuated from run to run, a new fair share was calculated in each run and in this table we have the average fair share from these runs.

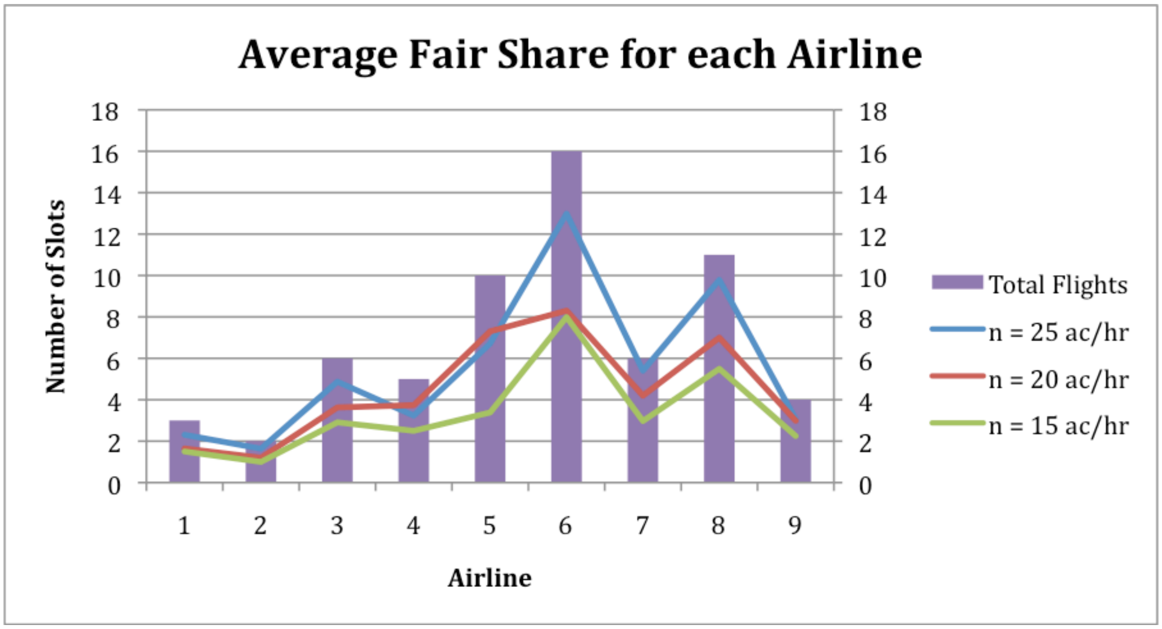


Figure 3.8 Computed average fair share for each airline and for each capacity reduction scenario

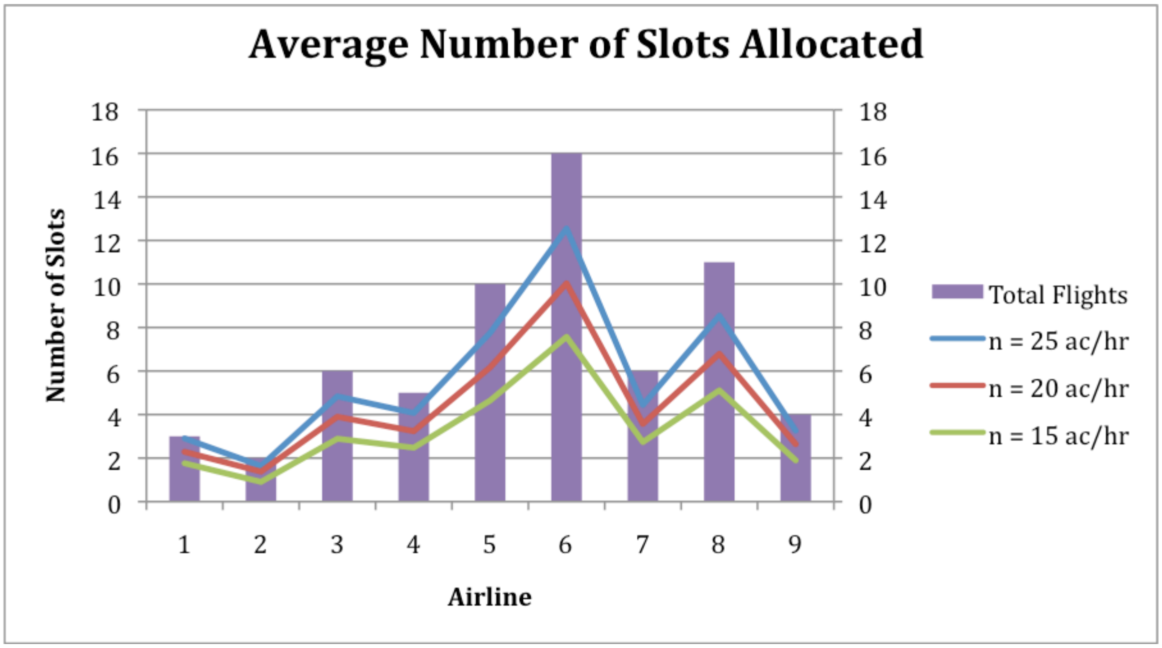


Figure 3.9 Average number of slots for each airline

In figures 3.9 and 3.10, are present the average and median number of slots that each airline actually got after the resource allocation mechanism was implemented. Again here it can be seen that the number of slots allocated to each airline matches very well its estimated fair share, and also smaller carriers have good chances to get slots.

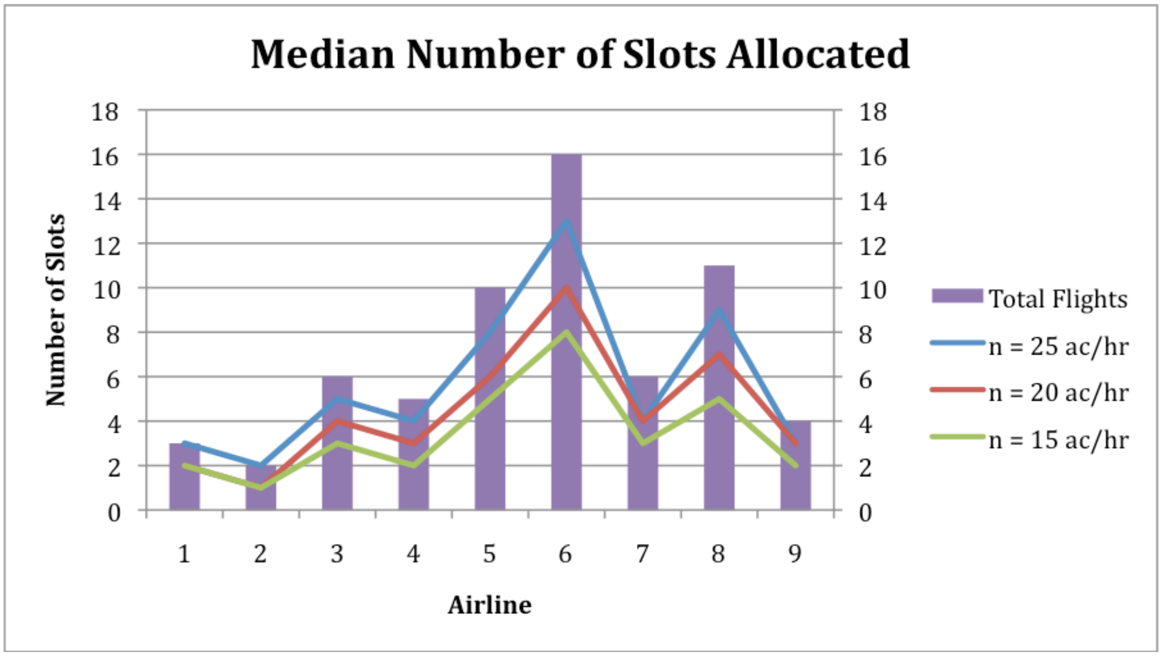


Figure 3.10 Median number of slots for each airline

In Table 3.7 can be seen the coefficient of variation of slots allocated to each airline. The coefficient for airlines with larger numbers of flights tends to be smaller than for those with fewer flights. This is partly because the bigger airlines have greater fluctuation in their schedules on a day-to-day basis. Overall the coefficients for all airlines are similar. Also the coefficients compared to the results with deterministic flights are a bit larger. This is expected, and the difference represents the marginal contribution of the noise in the schedule to the observed variation. The added effect of variability on number of flights has contributed to that. This can be better observed in the following Figures 3.11-3.13.

Table 3.7 Coefficient of Variation of Slots Allocated to Each Airline

		Capacity Reduction to n ac/hr		
		25	**20**	**15**
Coefficient of Variation of Slots for Each Airline	**1**	0.2336	0.2878	0.2743
	2	0.3007	0.3710	0.4073
	3	0.1405	0.1539	0.1517
	4	0.1930	0.2205	0.2353
	5	0.1313	0.1400	0.1440
	6	0.0853	0.0933	0.0914
	7	0.1600	0.1886	0.1922
	8	0.0960	0.1181	0.1145
	9	0.1816	0.2178	0.2190

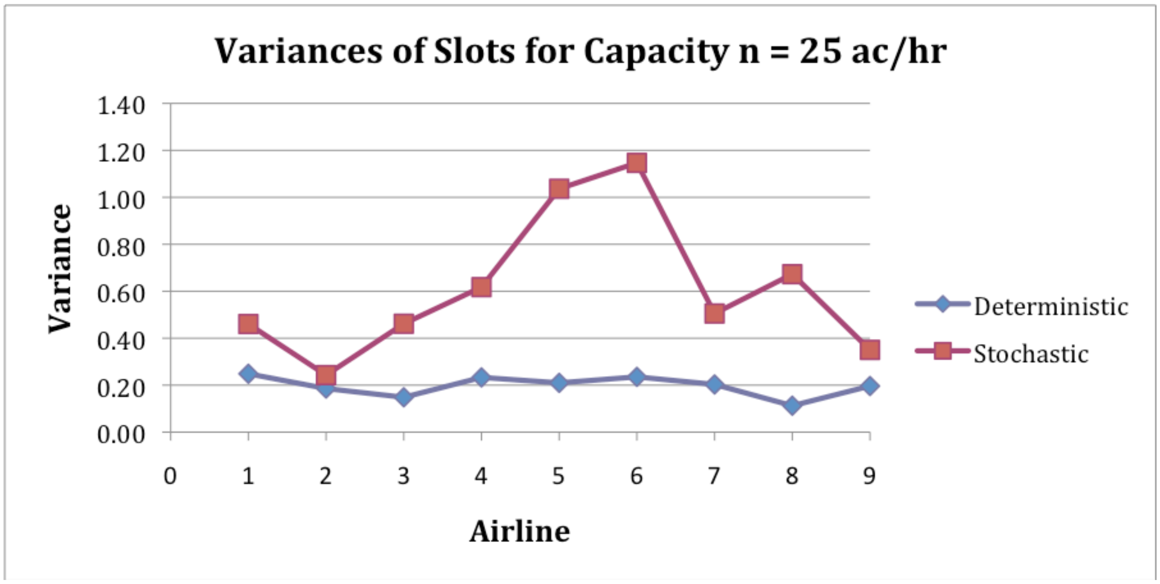


Figure 3.11 Variance of slots when capacity is reduced to 25 ac/hr

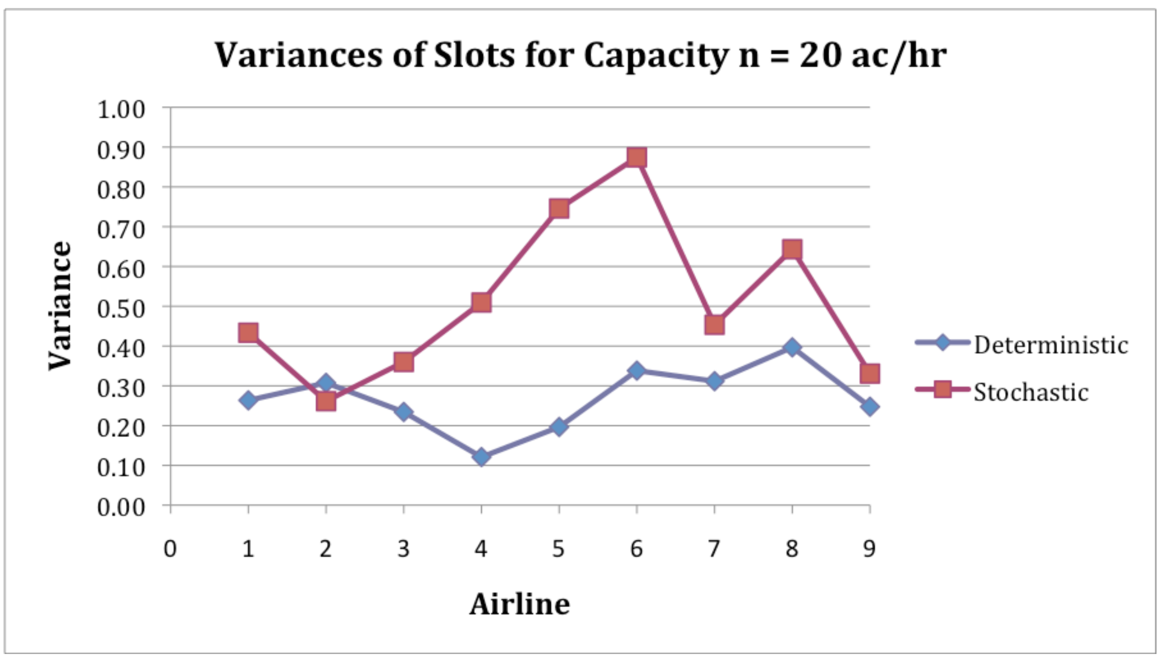


Figure 3.12 Variance of slots when capacity is reduced to 20 ac/hr

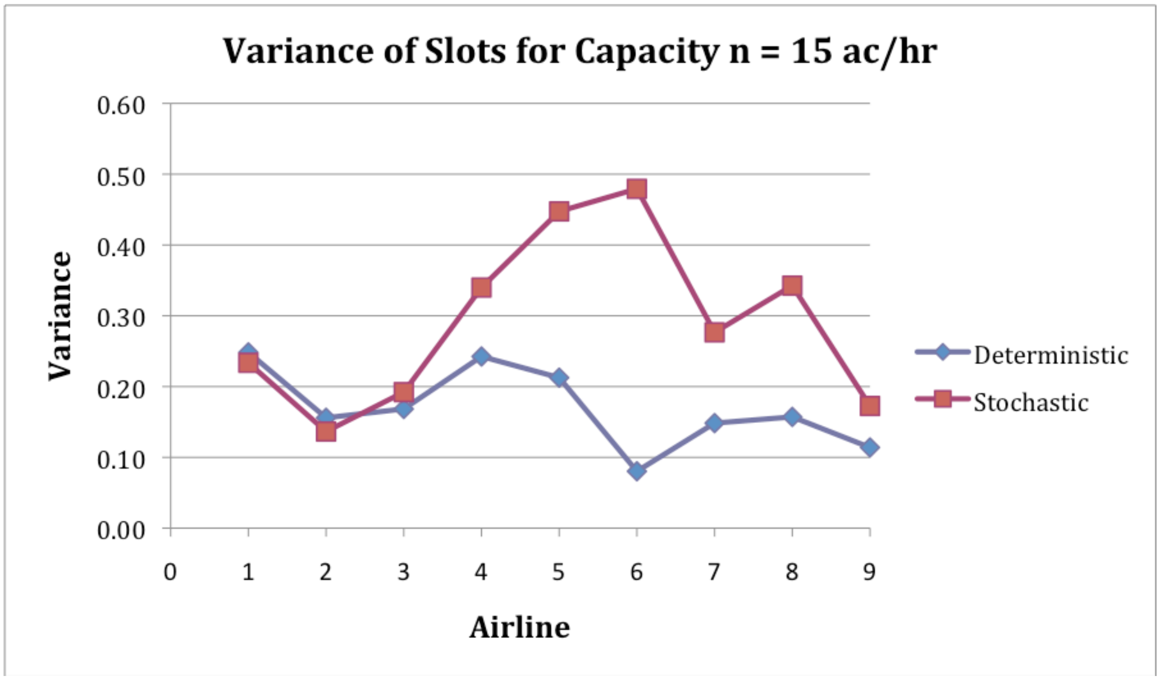


Figure 3.13 Variance of slots when capacity is reduced to 15 ac/hr

In Figures 3.11-3.13 can be seen that the airlines 5 and 6 have greater difference in their variances for each scenario, between the deterministic and the stochastic cases. This is due to the fact that they are bigger carriers and the range of flights they have is greater than the other airlines. For example airline 6 has on average 16 flights but it was observed that there were days that had 14 flights and other days up to 17. The airline 2, which represents the smaller carrier with only 2 flights, has small variance and the difference of it between the deterministic and stochastic cases is also small, because it was observed that the number of flights it has doesn't fluctuate with time.

In Table 3.8 are presented the results for the fair share deviation indicator. Here also, the indicator for each airline is very close to zero, which means that the airlines on average will be getting slots that are close to their fair share.

Table 3.8 Deviation of the Average Fair Share from the Actual Average Allocation

		Capacity Reduction to n ac/hr		
		25	**20**	**15**
Deviation of the Fair Share from the Actual Average Allocation	**1**	-0.00093	-0.00008	-0.00181
	2	-0.00043	-0.00173	0.00245
	3	0.00009	0.00074	0.00152
	4	-0.00007	0.00009	-0.00116
	5	-0.00024	0.00018	-0.00111
	6	0.00043	0.00085	-0.00041
	7	0.00077	-0.00011	-0.00136
	8	-0.00011	0.00095	-0.00021
	9	0.00050	-0.00089	0.00209

Table 3.9 Average Delay (in Minutes) per Flight per Airline

		Capacity Reduction to n ac/hr		
		25	**20**	**15**
Average Delay (in min) per Flight per Airline	**1**	12.2	17.1	21.2
	2	13.0	19.9	3.8
	3	10.0	15.6	19.9
	4	10.8	17.4	21.2
	5	8.2	13.3	18.9
	6	6.5	11.4	16.1
	7	10.6	16.1	20.9
	8	7.7	13.0	18.5
	9	11.6	18.2	19.6

Table 3.10 Weighted Average Delay (in Minutes) per Flight per Priority Number

		Capacity Reduction to n ac/hr		
		25	**20**	**15**
Weighted Average Delay (in min) per Priority Number	**4**	23.2	36.8	48.1
	3	24.8	44.0	64.4
	2	20.1	35.6	47.3
	1	11.6	17.1	18.0

Finally at tables 3.9 and 3.10 are presented the average delays per flight per airline and per priority number assigned respectively. In the first of these two tables, it is clear that when capacity reduces the amount of delay accrued by each airline increases. With the exception of airline 2, which has the least slots and when capacity reduces much it doesn't often being assigned to any slot, the differences in delays among the airlines are reasonable. In the other table can be seen again the same trend as before. The weighted average delay for flights with priority 4 is consistently less than the delay of flights with priority 3. As the number of priority reduces it was observed that the number of flights left without being assigned increased, which caused the reduced weighted delay compared to the higher priority flights.

3.8 Conclusions

In this research was proposed a meaningful way for carriers to express some preference structure during AFP. Also two resource allocation mechanisms were proposed that will improve the system efficiency and at the same time will take into account the preferences of the airlines. First, was examined how the results from using the proposed preference structure of airlines in the first proposed allocation scheme –PBPRA- is compared to RBS and RBS with substitutions. The results showed that the total weighted delay and weighted average delay for PBPRA is consistently lower than RBS and RBS with substitutions. Then was examined how the second allocation mechanism proposed, A-PBPRA, works compared to RBS and RBS with substitutions. The results showed that the total weighted and weighted average delay accrued by the flights with the proposed mechanism – A-PBPRA- is also much lower than the ones with RBS and RBS with substitutions. This work can be extended to look at other allocation mechanisms that can for example consider some airlines getting a number of early flights (more than one) for a higher price and then the rest of their share of flights to be of lesser value. Also the airlines preference structure can be extended to have some additional components, apart from the priority number and the maximum delay allowed, to make it even more rich and at the same time not containing much proprietary information.

During AFPs the airlines in the long-run will be getting on average what they want. As was estimated the smaller carriers have good chances of actually getting slots in the constraint areas. For smaller airlines the variance of slots allocated tends to be smaller than the variances for the bigger carriers. It was also shown that for the flights with priorities 4 and 3 were most of them assigned to slots and most of the ones with priorities 2 and 1 were not. The weighted delays for the flights with priority 4 were less than the ones with priority 3.

Chapter 4: Impact of Improved Predictability

As mentioned in the beginning of this dissertation, airlines will benefit from increased flight predictability. Airlines tend to add extra time to their scheduled block times in order to absorb delays and maintain their schedules intact as much as possible. It is a way to deal with unexpected delays that occur frequently and cause many problems to the airlines due to missed connections, crews being overtime, unhappy passengers etc. The anticipated mechanisms by which benefits could be realized as a result of improvements in strategic flight predictability can be articulated as follows:

- A reduction in the variability of actual flight times should lead to a reduction in scheduled block times and fuel buffers.
- The reduction in scheduled block times should lead to shorter actual block times.
- The reduction in fuel buffer will lead to a reduction in contingency fuel loaded, which will also lead to a reduction in actual fuel usage.
- With improvements in scheduled and actual block times, carriers could hypothetically achieve the same levels of scheduled operations with fewer aircraft and less total crew duty time.

While the number and duration of operations is not expected to change under this hypothesis, the fuel burned on every segment of each itinerary would be reduced.

4.1 Analyzing Scheduled Block Times

Since scheduled block time is the key component of the benefits analysis and in order to set a baseline of how airlines set their scheduled block times, historical data for specific flights were analyzed. Before any analysis, it is important to clarify the definitions of some of the phases of flight. A graphical display of them is presented in the following Figure 4.1.

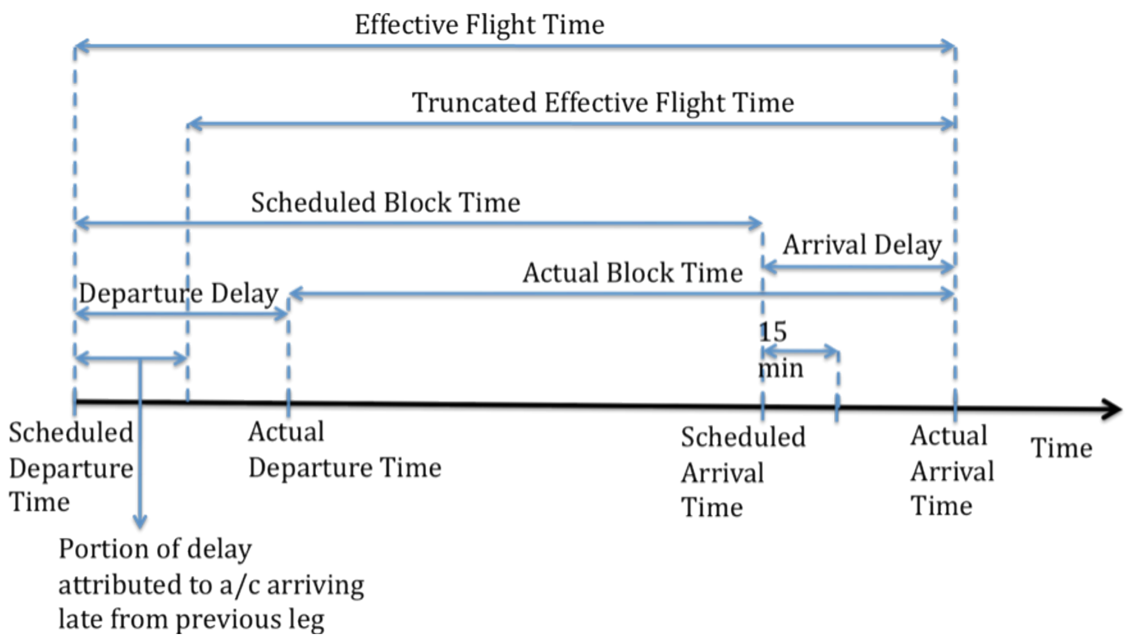


Figure 4.1 Definitions of phases of flight

The actual block time is the time between the actual departure and actual arrival time. As the effective flight time is considered the time elapsed between the scheduled departure time and the actual arrival time. In other words it includes the actual block time and the departure delay. Some portion of the departure delay is caused by the late arrival of the aircraft from its previous leg (late aircraft delay) and since this delay is already counted in the previous leg, the idea is to remove this portion from the current flight. So the truncated effective flight time can be considered, which is the effective flight time minus the late aircraft delay.

4.1.1 Scheduled Block Time for a Single Flight

First data for a single flight from January 2009 to December 2011 was collected (BTS 2013) and broken it down by quarter. Weekdays only were taken into account. The flight chosen was a United Airlines from Boston to San Francisco that leaves around 6am. In the data the scheduled time of departure was ranging is between 6.00am to 6.20am. The flight number for this was UA171 until October 2010 and changed to UA893 afterwards. After the analysis of the data the results are presented in the following figures.

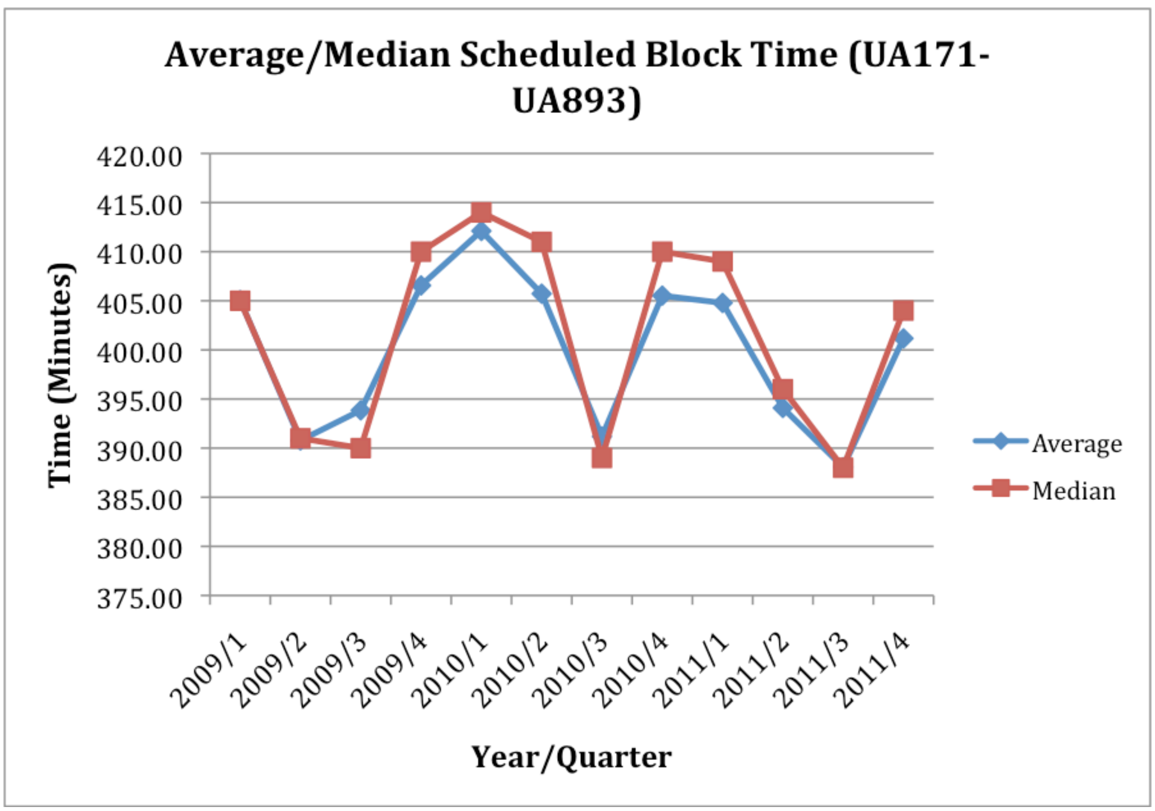


Figure 4.2 Average and median scheduled block time for a single flight between BOS and SFO

From the figures 4.2 and 4.3, it is evident that airlines do not have a fixed schedule block time for each flight within the same quarter. As it can be seen in Figure 4.2, the scheduled block times oscillate quarterly and during the 1st and 4th quarter these times are higher. In Figure 4.3 can be seen that during some quarters, especially the 4th there is increased standard deviation.

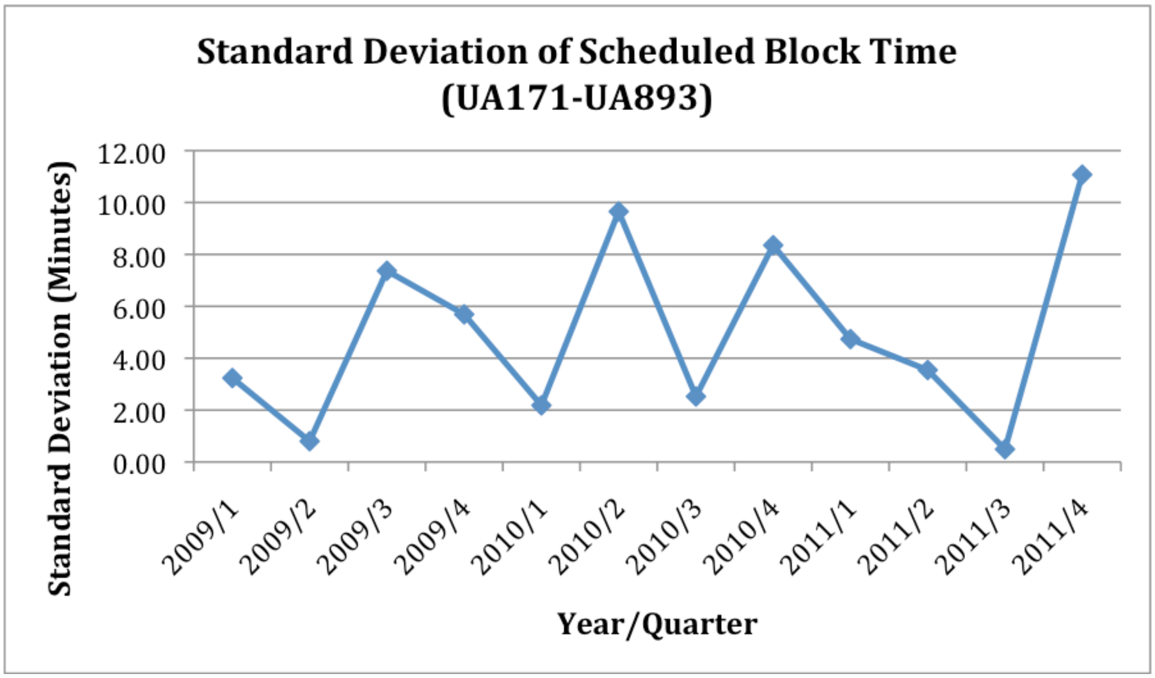


Figure 4.3 Standard deviation of the scheduled block time for a single flight between BOS-SFO

4.1.2 Scheduled Block Time for Multiple Flights in a Day

The next step in the data analysis was to examine the effect of different departure times to the scheduled block times for the same Origin-Destination (O-D) pair for the same airline. For this part, other flights scheduled between BOS and SFO from United Airlines were included in the analysis. Again flights from January 2009 to December 2011 were considered, weekdays only, and broken down by quarter. The flights were grouped to 5 different sets of departure times from BOS (6AM, 8AM, 11AM, 3PM, 6PM, local times), because the departure times were changing a bit throughout the years. In the following figures, Figures 4.4-4.6, the results are presented.

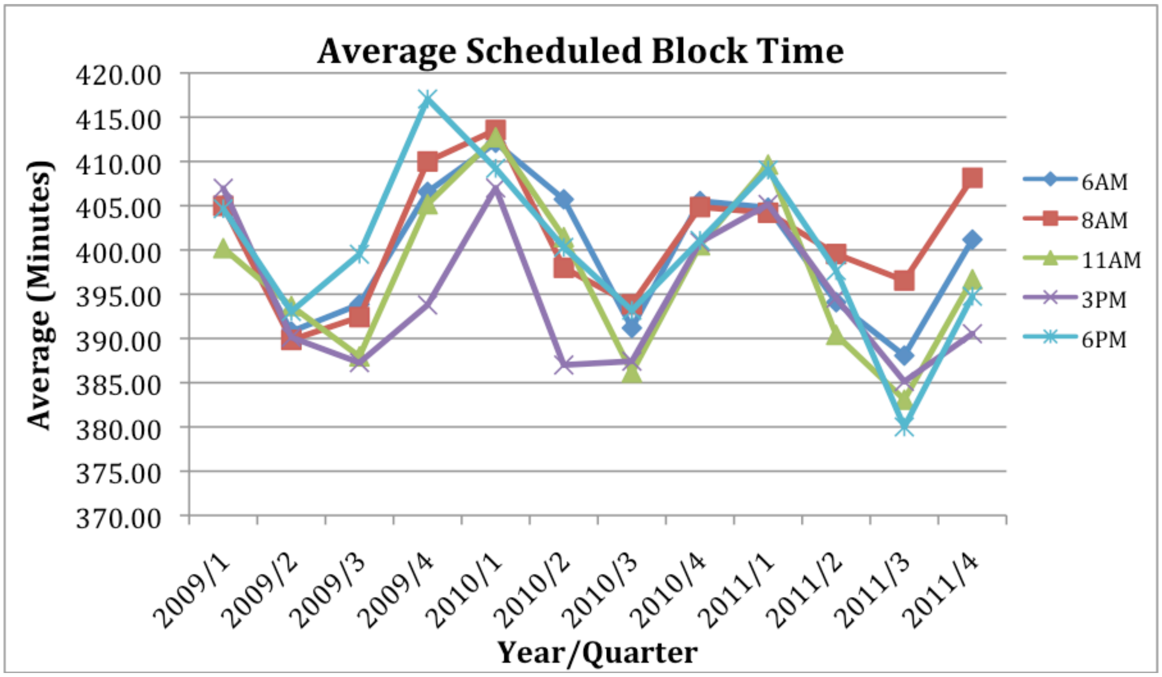


Figure 4.4 Average scheduled block times for all United Airlines flights between BOS-SFO

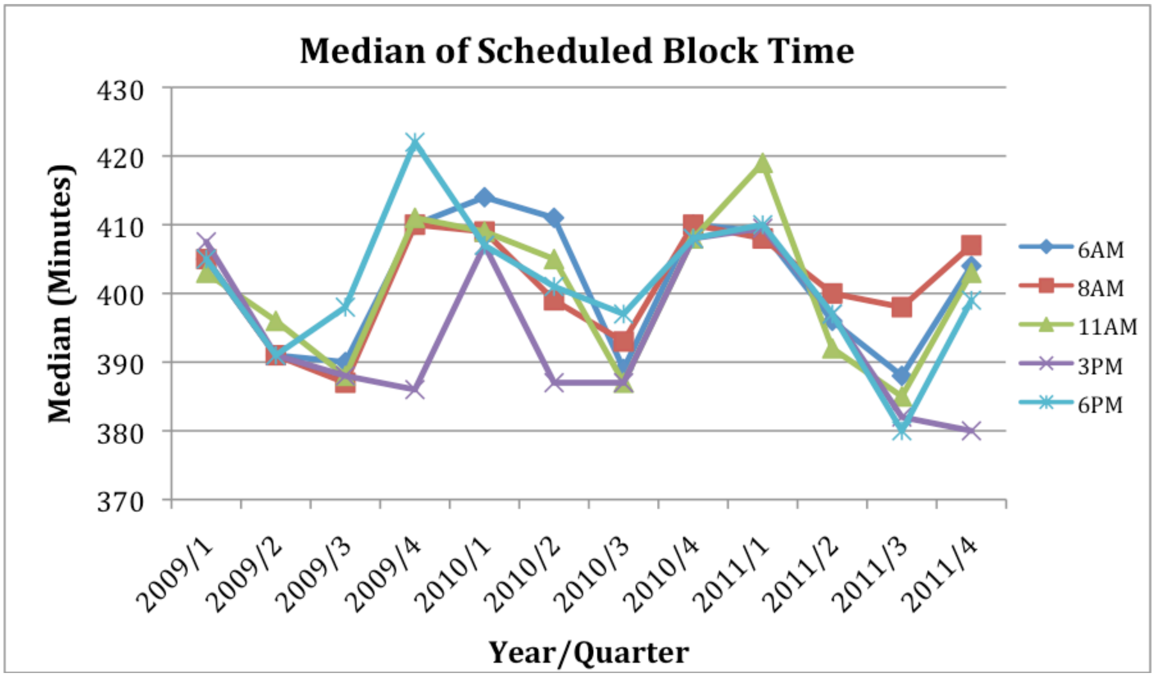


Figure 4.5 Median scheduled block times for all United Airlines flights between BOS-SFO

As it can be seen in figures 4.4 and 4.5, flights with different departure times in the day for the same O-D pair have different scheduled block times. The difference in scheduled block times can exceed 20 minutes, for example the flight departing at 3PM and flight departing at 6PM the 4th quarter in 2009. Finally the scheduled block times oscillate similarly by quarter and during the 1st and 4th quarter these times appear to be higher.

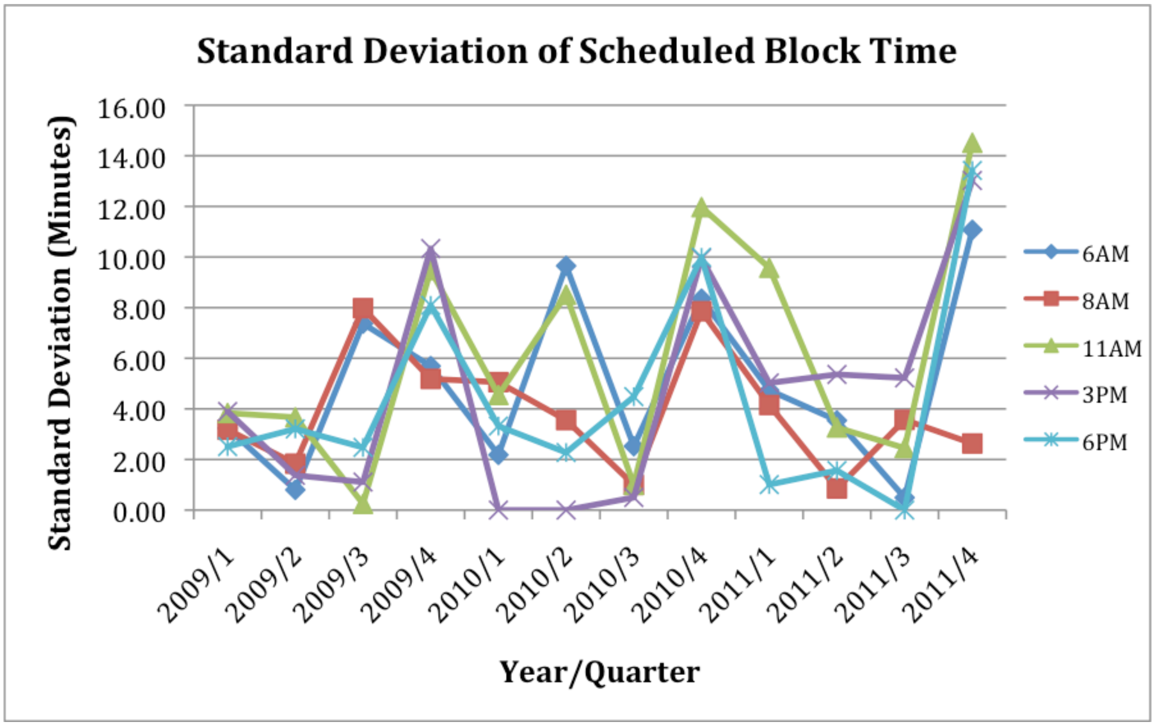


Figure 4.6 Standard deviations of the scheduled block times for all United Airlines flights between BOS-SFO

4.1.3 Estimating Distributions

It is known that airlines look at historical data, up to 5 years, for each Origin-Destination (OD) pair and look at the distribution of the effective flight times. It is also common for them to use historical data of competitor airlines for the same OD pair. From these distributions they set an on-time performance goal, which is flight specific and it may vary from 60% to 75%. Some airlines might ignore the Late Aircraft Delay (LAD), which is defined as the portion of the departure delay attributed to the aircraft arriving late from its previous flight leg, and construct the distributions of the truncated effective flight time (Deshpande and Arikan 2012). In this section these distributions are analyzed to find which ones fit better the actual data.

In Figure 4.7 and Figure 4.8 are presented the distribution of the effective flight time and the truncated effective flight time respectively for the third quarter of the years 2009 and 2010. It appears that both distributions are skewed to the right and possible distributions that could fit this shape are the log-normal, log-Laplace, gamma.

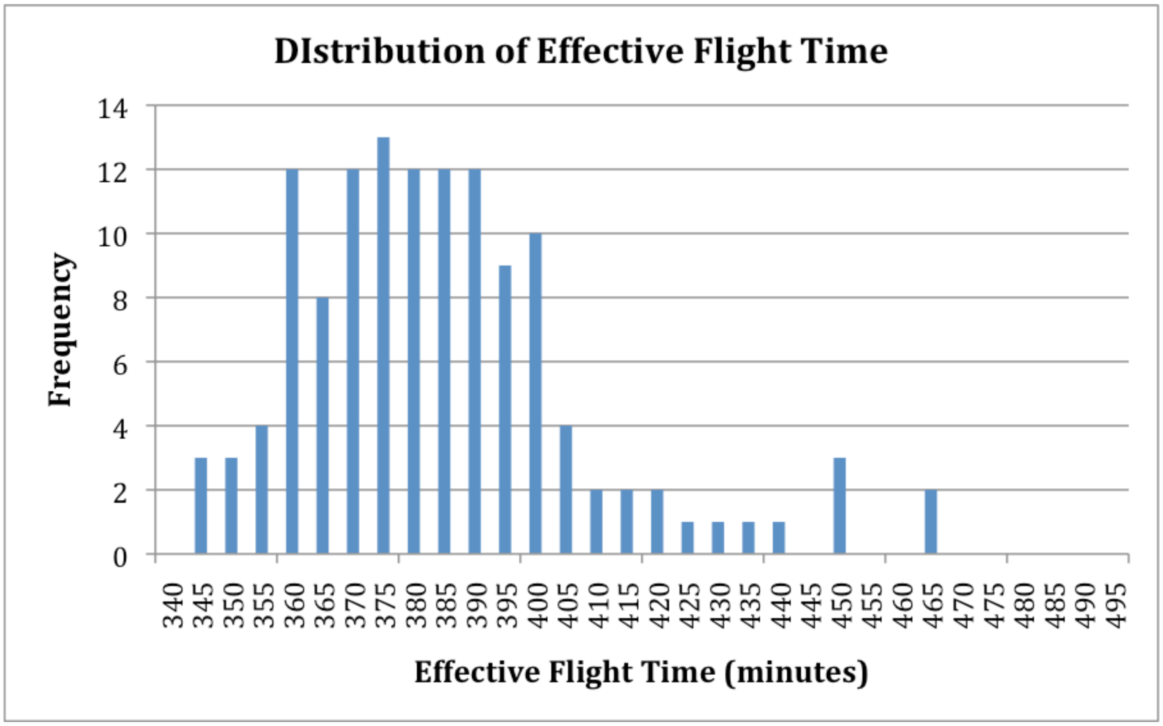


Figure 4.7 Distribution of the effective flight time of flight with departure time 3PM

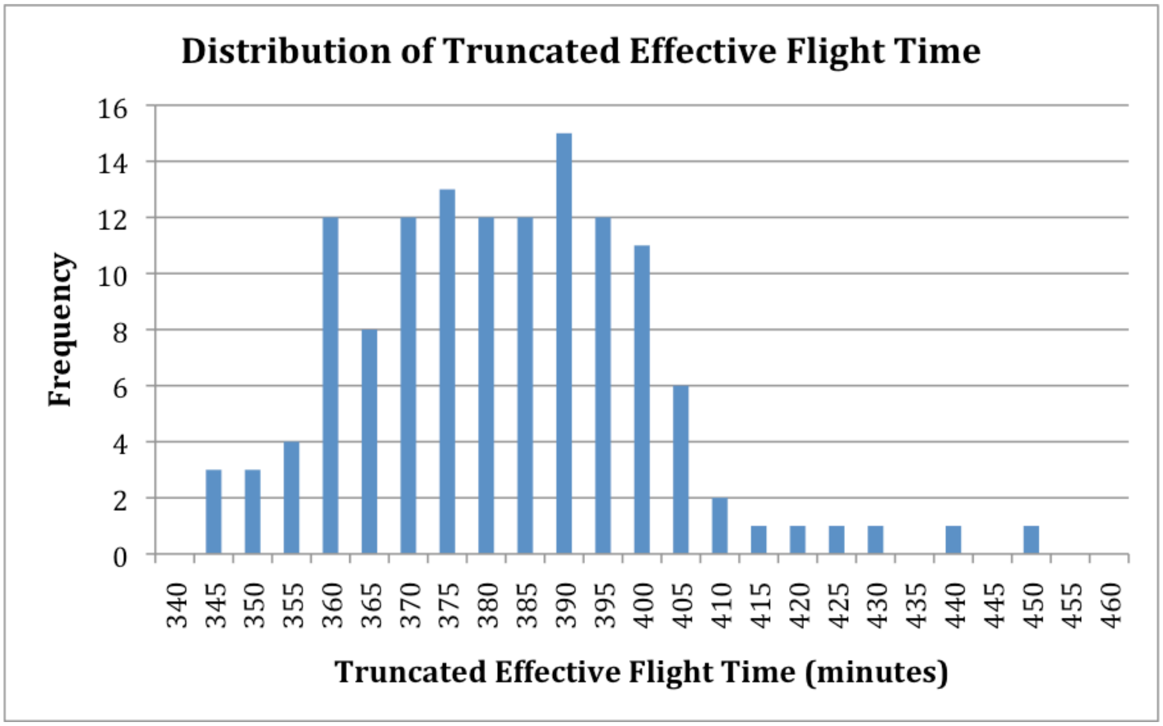


Figure 4.8 Distribution of the truncated effective flight time of flight with departure time 3PM

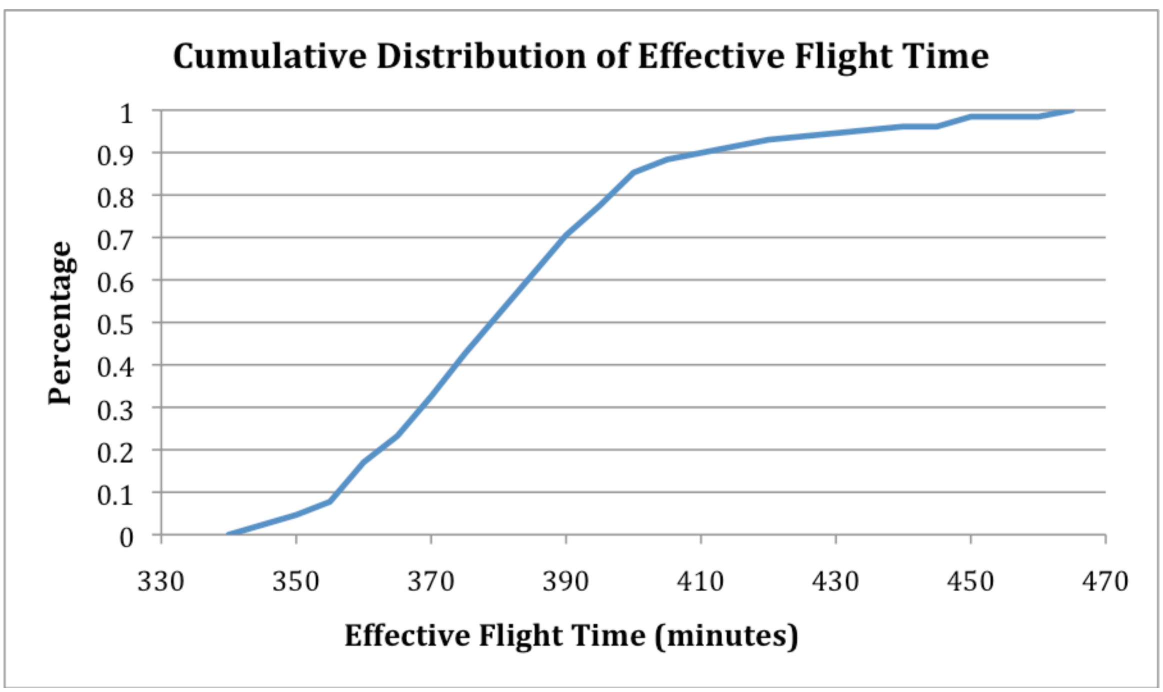


Figure 4.9 Cumulative distribution of the effective flight time of flight with departure time 3PM

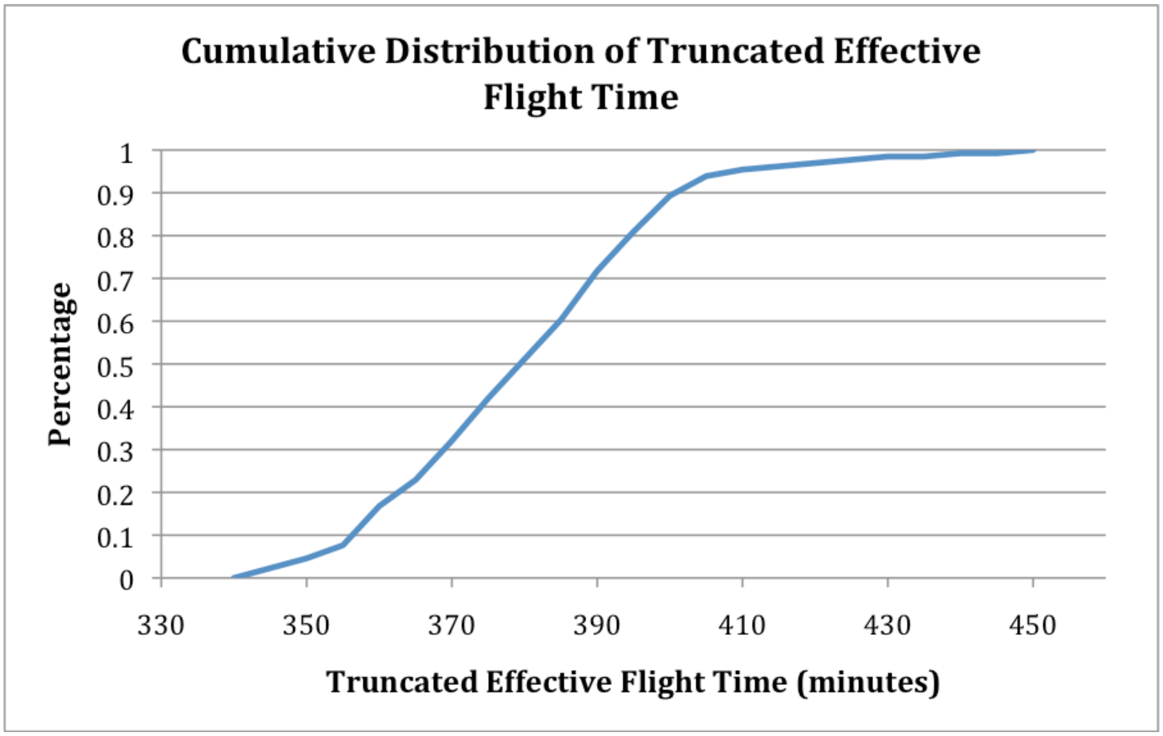


Figure 4.10 Cumulative distribution of the truncated effective flight time of flight with departure time 3PM

In Figure 4.9 and 4.10 are presented the cumulative distribution of the effective flight time and truncated effective flight time respectively. As mentioned above, it is believed that airlines look at the distributions of the past few years and try to set a goal for the next year, to reach a certain level of service. While looking at the average scheduled block time set for the third quarter of 2011, which was 385.1 minutes, and at Figure 4.10 it can be deduced that United had set it's goal to be approximately 60% of flights to be on time. The actual percent of flights that arrived with delay less than 15 minutes in 2011 was 74%. For the rest quarters of 2011, by looking at the actual average scheduled block times set for each of them, and looking at the distributions constructed from 2009-2010 data, can be concluded that United had set it's goal to be 60% of flights to be on time for this specific OD pair.

In Figure 4.11 the distribution of the actual block time can be seen, which appears to be slightly skewed to the right.

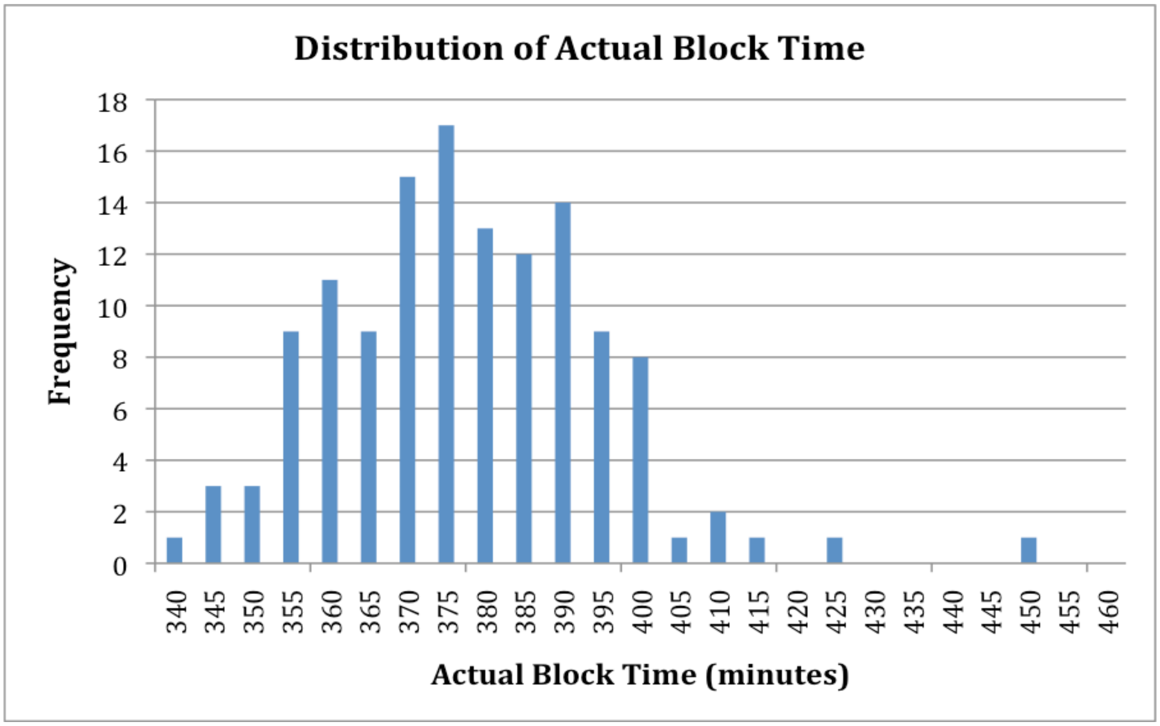


Figure 4.11 Distribution of the actual block time of flight with departure time 3PM

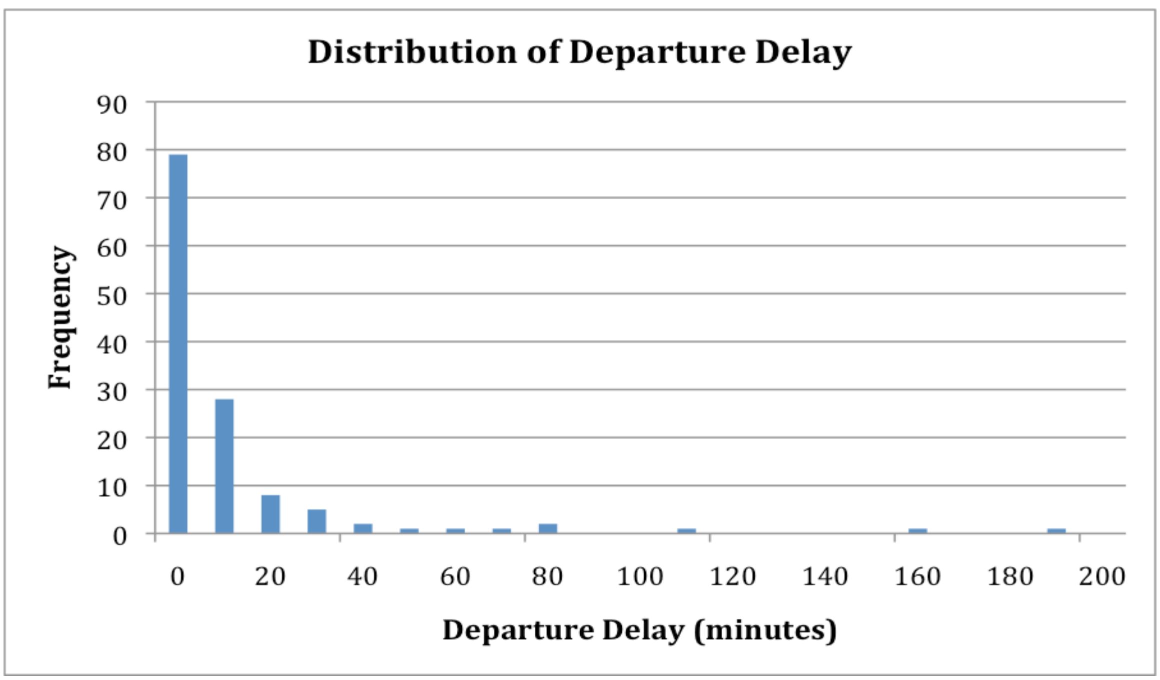


Figure 4.12 Distribution of the departure delay of flight with departure time 3PM

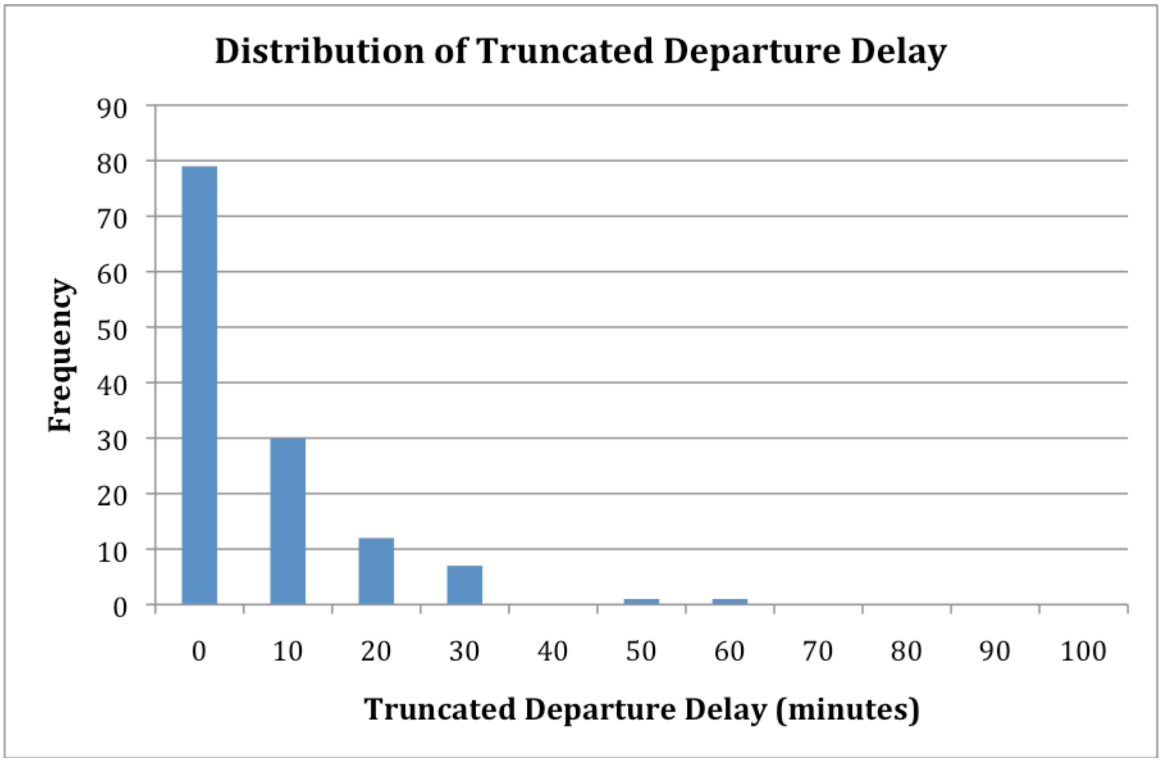


Figure 4.13 Distribution of the truncated departure delay of flight with departure time 3PM

Finally in figures 4.12 and 4.13 are presented the distribution of the departure delay and truncated departure delay respectively. Possible distributions that could fit their shape is the exponential and gamma.

4.1.4 Fitting Gamma Distribution to Data

The next step presented in this section is to fit the actual data to a distribution. The distribution chosen is gamma, because of its flexible shape, depending on the parameters chosen – shape and scale -. First the gamma distribution was fitted to the effective flight times for the 3rd quarter of 2009 and 2010. In Figure 4.14 is depicted the fitting for which the shape parameter is 275.977 and scale parameter is 1.387.

To have 385.1 average scheduled block time for the 3rd quarter in 2011, the percent given from the gamma distribution is 54.9%. In other words the airline must had chosen 54.9% of their flights to be on time.

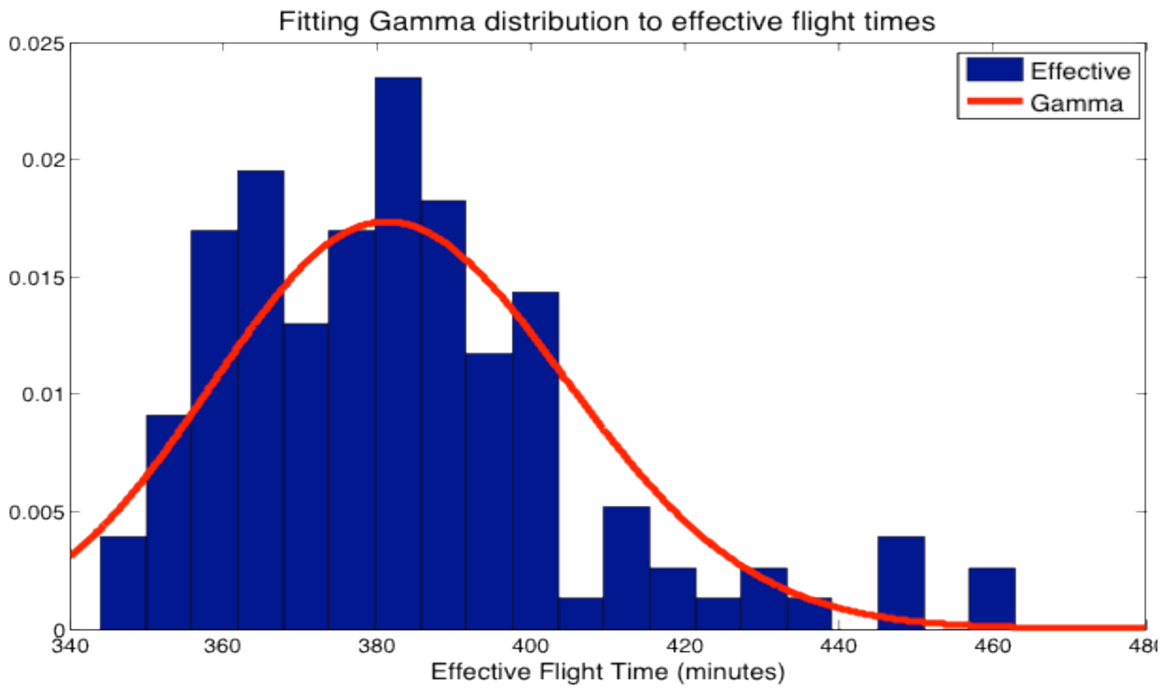


Figure 4.14 Fitting gamma distribution to the actual data of effective flight time

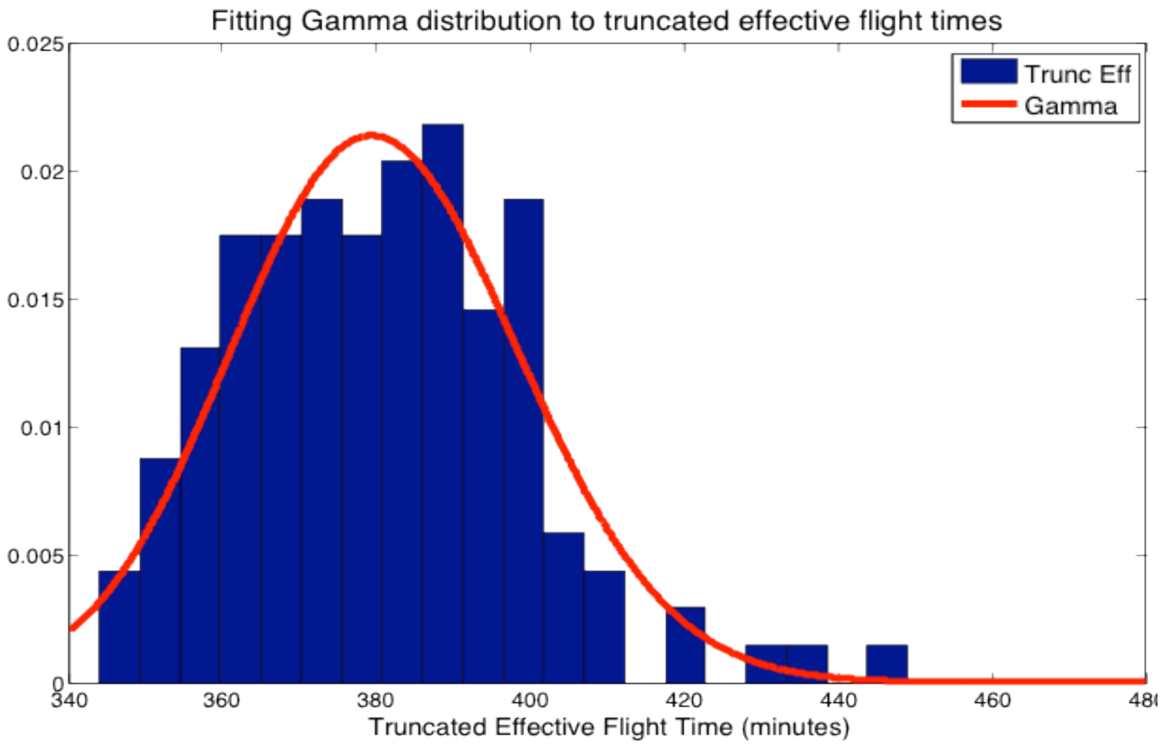


Figure 4.15 Fitting gamma distribution to the actual data of truncated effective flight time

Then the gamma distribution was fitted to the truncated effective flight times for the 3rd quarter of 2009 and 2010. The shape parameter is 413.889 and scale parameter is 0.919 for this gamma distribution, which is depicted in Figure 4.15. To have 385.1 average scheduled block time for 3rd quarter in 2011 the percent given by this distribution is 60.6%. If the findings from the two fittings are compared with what we get from the cumulative distribution functions, which is approximately 60% of flights to be considered on time, can be concluded that the fitting of the truncated effective flight times data gives better estimate than the effective flight data.

4.2 Estimating Scheduled Block Times

4.2.1 Source of Information

In order to solidify how airlines set their scheduled block times, input from industry experts was asked for. Individual meetings and phone interviews with the following people were held:

- Meeting with the Delta team responsible for the estimation of SBT
- Teleconference with Jim Hamilton, from UPS
- Teleconference with George Kypreos from American Airlines
- Teleconference with Michael Clarke from Sabre and Tuell Green from American Airlines

4.2.2 General Process

From the discussions came out that all airlines follow the same general approach in order to estimate the SBT. They all break the block time in 3 components: taxi-in, airtime, taxi-out, as we can see in Figure 3.16. They look at historical data for each component and plot the distribution of block time. From this distribution they set an on-time performance goal, which is flight specific. As it can be seen in Figure 3.17, for the distribution of historical block times an airline sets its goal to α%, which means that they want α percent of the time their flight to be on-time.

Figure 3.16 Block time breakdown

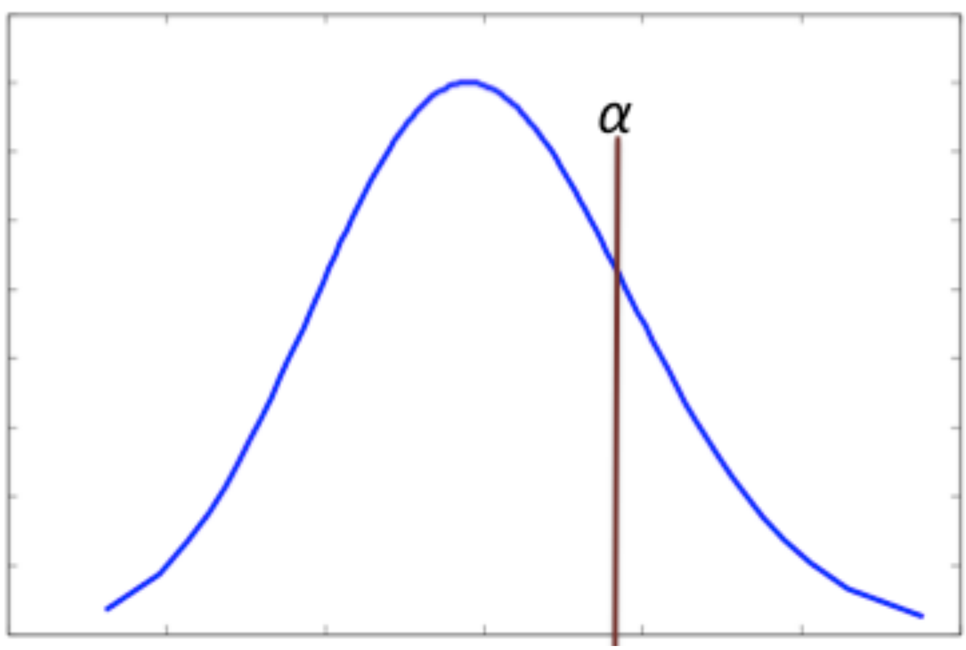


Figure 3.17 Block time distribution and on-time performance goal set to α%

The above three time components depend on the Origin-Destination (O-D) pair, the aircraft used, the time of the day, day of week and season.

The airtime depends on what aircraft is used for the specific route. Airlines tend to estimate their en-route times by considering aircraft flying at their optimal speed. So for the same O-D pair, if various aircraft types are used, then for each of them a different SBT is calculated. Airtime also depends on the O-D pair because of some traffic issues in certain areas, like the Northeast where there are many metroplexes, which cause a greater variance in the flight times. Finally airtimes depend on the season since different wind patterns exist.

Taxi times depend on the origin and destination airports. For each airport a different taxi-in and taxi-out time is calculated in order to take into account the specific characteristics of the airport (like runway usage, traffic, runway configuration etc). Taxi times also depend on the time of the day (peak and non-peak traffic) and the day of the week.

All airlines exclude the outliers from their data and also constantly monitor the on-time performance of flights and tweak the schedule as needed. Also, they do not take into account directly the propagated delay.

4.2.3 Airlines' Differentiations

The general process is common for each airline. Of course depending on the airline and its business needs and models some parameters differentiate. The ones that vary are summarized as follows.

Percentiles

The range of percentiles that each carrier considers varies. UPS for example, a freight carrier, gives a great emphasis on the on-time arrival of parcels in critical markets, especially for the early morning deliveries. So for these kinds of markets they look at the 80^{th} to 90^{th} percentile instead of the 60^{th} percentile that they usually do. The range of percentiles for Delta Airlines is 65%-75% and for American Airlines 70%-75%.

Taxi-in and taxi-out estimation

Different approaches are used to estimate the taxi-out and taxi-in times. One airline takes the average taxi-in and taxi-out times observed. Another airline takes the average taxi-in and taxi-out times observed the few previous months. In the case of contingencies such as runway construction they use surface simulations to estimate the new taxi times. In another airline, they estimate the taxi-out times with the assistance of a simulation software after they input all the parameters i.e. other airline traffic, runway usage during different time periods, taxiway traffic, separation, runway configuration etc. For the taxi-in times they usually set to 5-8 minutes, unless there is a big issue with terminal location versus runway usage.

Data used

The data that each airline uses for their analysis is also different. One of the airlines looks at historical data for each O-D pair since 1988. Another airline looks at the last 3-4 last years if nothing has changed. The third airline for the airtime looks at the last 5 years and for the taxi times only the most recent – few months-.

Seasons considered

Two of the airlines consider only two seasons (Summer-Winter). At the other airline they consider eight seasons for the domestic flights and two seasons for their international flights.

Types of days

For this parameter there is a great variation among the airlines. One airline considers all days together. Another airline considers each day separately. And for the third airline there are two types of days: Saturdays and all other days.

4.3 Estimating the Strategic Benefits of Increased Flight Predictability

The scope of this work is to estimate the benefits for airlines due to the improvement in strategic flight predictability. The work flow for the benefits assessment related to scheduled block time begins with the work conducted by U.C. Berkeley, which developed airline specific scenarios for the impact of changes in empirical block time distributions on scheduled and actual block times. These scenarios were used as the entry points for an economic benefits assessment. In all scenarios the median block time remains the same. Scenario1 depicts a condition where flight time variability is reduced due to the increase of the shortest flight times. The second scenario depicts the exact opposite case, where the longest block times are reduced the most. Finally Scenario 3 considers a consistent change throughout the flight time distribution.

The benefits assessment process will use the results shown in Table 4.1, for Low Cost Carriers (LCC) and legacy carriers, Delta Airlines, American Airlines, and United Airlines.

Table 4.1 Evaluation results under different scenarios

	LCC			***Delta Airlines***			***American Airlines***			***United Airlines***		
Scenario	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
Mean Block Time Difference	0.96	-3.27	-1.16	1.19	-3.44	-1.12	1.23	-3.58	-1.17	1.31	-3.35	-1.02
SBT Difference	-1.82	-3.03	-2.38	-2.16	-3.88	-3.02	-1.93	-4.51	-3.22	-0.19	-0.27	-0.23

These changes in mean actual and scheduled block time can be related to changes in average daily flights per aircraft and average pilot salary per available seat mile (ASM) using regression coefficients from a study by Moreno-Hines and Kirkman (2013). In their analysis was included the estimation of coefficients that can be directly applied to the task of converting between average scheduled block time and certain dependent variables that, by themselves, are not monetized, but that can be monetized in a subsequent step using some reasonable assumptions. Table 4.2 below shows the coefficients used in this benefits assessment, as taken from that reference.

Table 4.2 Block time benefits regression coefficients

Δ in Dependent Variable	**Δ in Explanatory Variable**	***AA***	***American Eagle***	***Delta***	***JetBlue***	***Northwest Airlines***	***Southwest Airlines***	***UA***	***US***
Average Scheduled Block Time	Average Actual Block Time	0.75	0.75	0.91	0.9	0.8	0.78	0.62	0.89
Average Daily Flights per Aircraft	Average Scheduled Block Time	-0.02		-0.02		-0.05		-0.02	
Average Pilot Salary per Available Seat-mile	Average Scheduled Block Time	6.4E-5				8.2E-5	6.7E-5		

There are two limitations of that study that add some complication to the prospect of using the results in this benefits analysis. First, the empty cells in Table 4.2 represent coefficients that would be necessary for completeness, but were not included in the paper. In the paper, each model was developed in a parsimonious form, so independent variables that did not significantly improve the fit of the model were removed from the model specification. There is no way of knowing, at this point, what the un-estimated coefficients might have been, or what their p-values might have been.

Secondly, the set of carriers reported in that study does not match exactly the set used for the scenario generation as part of the scheduled block time impact modeling performed for this work. As a result, a mapping between the two sets was constructed, which then necessitates the assumption that the effects to certain carriers are expected to be the same as those of other "similar" carriers. In particular, both studies used the major carriers American, Delta, and United, so those results were directly transferable. The Moreno-Hines and Kirkman study included Northwest Airlines, which was absorbed into Delta Airlines in 2010. In this study, the results for Northwest are ignored, although one could argue that their expected behavior might in some way manifest itself as part of Delta, but there is no way of quantifying this. Among low-cost carriers, the Moreno-Hines and Kirkman paper included American Eagle, JetBlue, and Southwest. In the current study, low-cost carriers were consolidated into one entity with respect to scenario generation.

Table 4.3 below shows the data used for the benefits assessment. The Available Seat Miles (ASM), numbers of pilots/copilots, fleet size, and departures per year are from BTS. The wage data is also from BTS, specifically US DOT Form 41, Schedule P6 & P10. The yearly ownership cost per aircraft is computed as a weighted average. For each carrier, their fleet is stratified into different airframe types, each with different ownership costs. The average ownership cost per aircraft is then a weighted average of these values, weighted by the fraction of the total fleet represented by that particular airframe type. The data for this analysis are from Aviation Daily (2013).

The three following tables 4.4-4.6 show the results of the analysis, one for each of the scenarios. In Scenario 1, American and Delta can save over 1.5 aircraft apiece, resulting in significant savings in ownership costs. The savings to United are more modest. As mentioned above, the necessary data to compute these savings for the low cost carriers are missing. For all of the carriers, however, it was possible to compute expected savings in pilot and flight attendant salaries that would be realized by conducting the same operational tempo with fewer total aircraft. Again, the numbers for American and Delta are high, and for United quite low. The savings for the LCCs are somewhere in between.

Table 4.3 Data for benefits assessment

				Low Cost Carriers (LCC)		
	American Airlines	***Delta Airlines***	***United Airlines***	***Airtran***	***JetBlue***	***Southwest***
Available Seat Miles (2011)	9.00E+10	1.09E+11	6.27E+10	2.34E+10	3.09E+10	1.03E+11
Pilots/Co-pilots (2011)	4898	6980	3731	1570	1730	5676
Departures per year (2011)	531,000	729,000	319,000	246,000	209,000	1,142,000
Mean Annual Wage for Pilots (2011)	$139,963	$150,099	$125,690	$128,225	$139,744	$203,196
Mean Annual Wage for Flight Attendants (2011)	$51,197	$40,475	$37,888	$32,088	$37,987	$54,120
Fleet Count	608	722	697	129	183	582
Yearly Ownership Cost per Aircraft	$1,766,492	$1,748,785	$2,366,392	$1,949,242	$1,735,440	$1,457,649

Table 4.4 Results for benefit Scenario 1

				Low Cost Carriers (LCC)		
	American Airlines	***Delta Airlines***	***United Airlines***	***Airtran***	***JetBlue***	***Southwest***
% of Saved Aircraft	1.61	1.56	0.30	0.00	0.00	0.00
Savings from Aircraft Saved	$17,326,246	$19,717,863	$4,998,482	$-	$-	$-
Reduction in Pilot salaries for current scenario (salary savings)	$11,116,924	$15,040,097	$762,955	$2,857,420	$3,769,531	$12,578,477
Reduction in Flight Attendant Salaries (salary savings)	$4,066,454	$4,055,643	$229,985	$715,063	$1,024,682	$3,350,200

Scenario 2 exhibited the most pronounced reduction in scheduled block times, and hence should produce the greatest expected savings. The results in Table 4.5 below can verify this. The relative standings amongst the airlines are the same as before, which is to be expected, because the cost coefficients are the same. Finally, Table 4.6 below shows the results for Scenario 3, which was representative of an intermediate level of reduction of scheduled block time.

Table 4.5 Results for benefit Scenario 2

				Low Cost Carriers (LCC)		
	American Airlines	***Delta Airlines***	***United Airlines***	***Airtran***	***JetBlue***	***Southwest***
% of Saved Aircraft	3.77	2.81	0.43	0.00	0.00	0.00
Savings from Aircraft Saved	$40,487,757	$35,419,125	$7,103,105	$-	$-	$-
Reduction in Pilot salaries for current scenario (salary savings)	$25,977,889	$27,016,471	$1,084,199	$4,757,133	$6,275,648	$20,941,091
Reduction in Flight Attendant Salaries (salary savings)	$9,502,440	$7,285,136	$326,821	$1,190,461	$1,705,927	$5,577,530

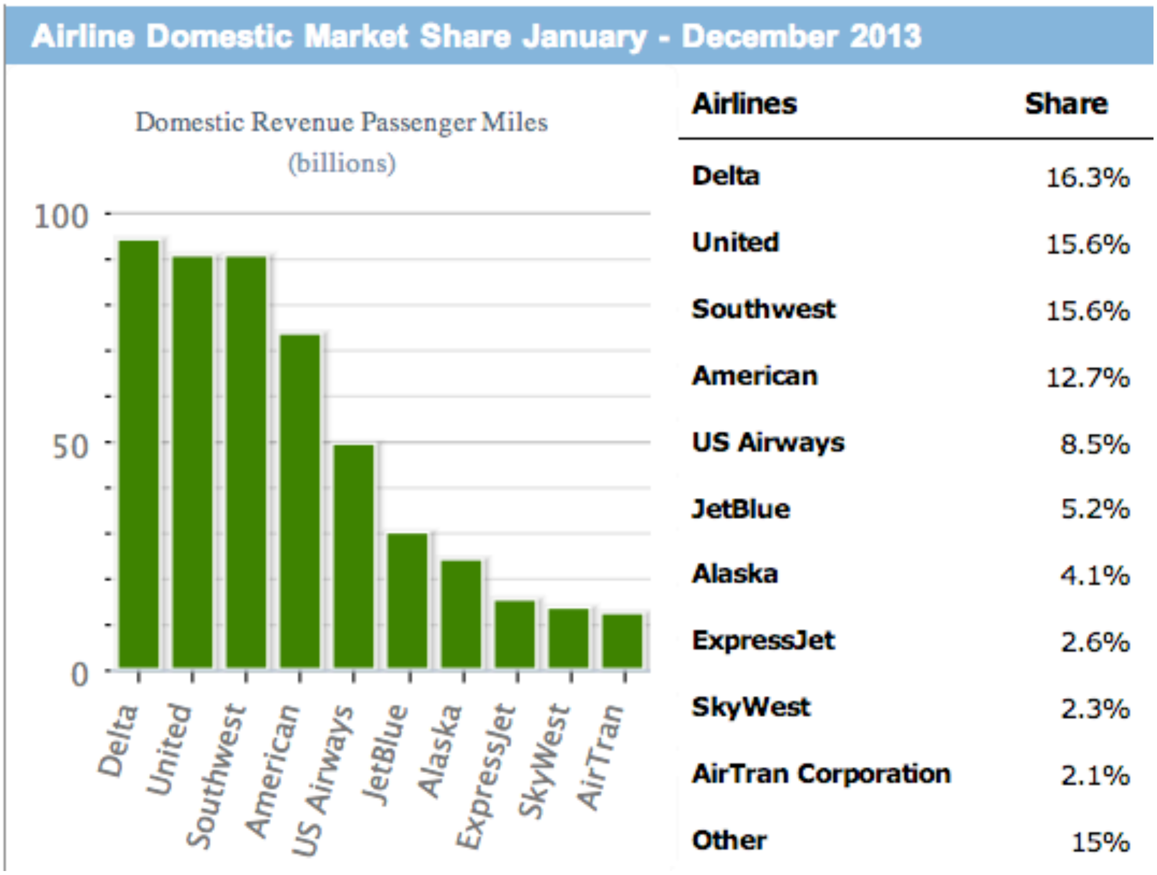
Table 4.6 Results for benefit Scenario 3

				Low Cost Carriers (LCC)		
	American Airlines	***Delta Airlines***	***United Airlines***	***Airtran***	***JetBlue***	***Southwest***
% of Saved Aircraft	2.69	2.18	0.37	0.00	0.00	0.00
Savings from Aircraft Saved	$28,907,002	$27,568,494	$6,050,794	$-	$-	$-
Reduction in Pilot salaries for current scenario (salary savings)	$18,547,406	$21,028,284	$923,577	$3,736,626	$4,929,387	$16,448,777
Reduction in Flight Attendant Salaries (salary savings)	$6,784,447	$5,670,390	$278,403	$935,082	$1,339,969	$4,381,030

4.3.1 Estimating the Benefits for the NAS

An extension of this work is to estimate what would be the benefits across the National Airspace System (NAS). In order to do so, it was need to include a few more carriers in this analysis. From Figure 3.18 can be seen which carriers have the biggest domestic market share. Most of the top airlines have been included already in the analysis, but two major carriers are missing and these are US Airways and Alaska Airlines. Since the coefficients are not known for these airlines, some assumption need to be made, that these carriers are similar with some of the carriers used previously and have the necessary data and assume that the benefits will be similar. For US Airways and Alaska Airlines is assumed that they have the same coefficients as American Airlines. From Figure 3.18 can be seen that their share is closer to American Airline's than the other two major carriers, Delta and United.

In the figure is also depicted that two regional carriers are included ExpressJet and Skywest. Unfortunately in the initial analysis there was no regional carrier included, so there are no data available to work with. However, since some Low Cost Carriers are included, this work could be extended by considering an additional one. Frontier is also a LCC and the assumption has to be made is that it will get similar benefits as Airtran, so this input can be used. Another type of carriers that are not included in this work is freight carriers. Freight carriers operate under different business model than major carriers, so their benefits cannot be directly associated. Also, most of their critical operations take place overnight where traffic levels are low and queuing delays are present. This means that the need to add contingency to their scheduled block times is not as big as is for the passenger carriers, so the benefit of increased predictability will not be as profound. Additionally freight carriers do not report to BTS and that would make very difficult to find the necessary input data.



Airlines	Share
Delta	16.3%
United	15.6%
Southwest	15.6%
American	12.7%
US Airways	8.5%
JetBlue	5.2%
Alaska	4.1%
ExpressJet	2.6%
SkyWest	2.3%
AirTran Corporation	2.1%
Other	15%

Figure 3.18 Airline Domestic Market Share (source: BTS)

The following Table 4.7 shows the data we used for this analysis. The sources of the data are exactly the same as with the previous analysis.

Table 4.7 Additional data for benefit assessment

	US Airways	***Alaska Airlines***	***Frontier***
Available Seat Miles (2011)	5.25E+10	2.38E+10	1.11E+10
Pilots/Co-pilots (2011)	4003	1286	676
Departures per year (2011)	403000	145000	85000
Mean Annual Wage for Pilots (2011)	$109,535	$155,024	$113,250
Mean Annual Wage for Flight Attendants (2011)	$40,442	$35,433	$19,105
Fleet Count	327	131	60
Yearly Ownership Cost per Aircraft	$2,469,821	$2,138,191	$2,496,360

In the following tables 4.8-4.10 are presented the results for each of the scenarios used in the previous analysis. Again, the scenario that shows the most cost savings is Scenario 2.

Table 4.8 Additional results for Scenario 1

	US Airways	***Alaska Airlines***	***Frontier***
% of Saved Aircraft	1.14	1.27	0.00
Savings from Aircraft Saved	$9,232,857	$3,565,342	$-
Reduction in Pilot salaries for current scenario (salary savings)	$6,481,836	$2,940,270	$1,352,193
Reduction in Flight Attendant Salaries (salary savings)	$2,393,193	$672,042	$228,112

Table 4.9 Additional results for Scenario 2

	US Airways	***Alaska Airlines***	***Frontier***
% of Saved Aircraft	2.67	2.97	0
Savings from Aircraft Saved	$21,575,227	$8,331,446	$-
Reduction in Pilot salaries for current scenario (salary savings)	$15,146,673	$6,870,787	$2,251,178
Reduction in Flight Attendant Salaries (salary savings)	$5,592,384	$1,570,419	$379,768

Table 4.10 Additional results for Scenario 3

	US Airways	***Alaska Airlines***	***Frontier***
% of Saved Aircraft	1.91	2.12	0
Savings from Aircraft Saved	$15,404,042	$5,948,394	$-
Reduction in Pilot salaries for current scenario (salary savings)	$10,814,254	$4,905,528	$1,768,252
Reduction in Flight Attendant Salaries (salary savings)	$3,992,788	$1,121,230	$298,300

In order to find the overall benefits across the NAS we add the benefits of each airline for each scenario, and the results are shown in Table 4.11

Table 4.11 Total benefits

	Total Savings from Aircraft Saved	Total Pilot Salary Reduction	Total Flight Attendant Salary Reduction
Scenario 1	$54,840,791	$56,899,702	$16,735,373
Scenario 2	$112,916,660	$110,321,068	$33,130,886
Scenario 3	$83,878,725	$83,102,092	$24,801,639

The benefits for Scenario 2 will be the greatest, since for this scenario the reduction in scheduled block time was the highest. In general the savings from the pilot salary reduction will be approximately equal with the savings from aircraft saved. The savings from the flight attendant salary reduction will be approximately a third of the other two. Here we have to note that these benefits will take time to get realized.

Once the system gets more predictable, the actual block times will be reduced. When this reduction will be profound enough will lead airlines to reduce their scheduled blocks times.

Chapter 5: Conclusion

The air transportation system in the United States is one of the most complex systems in the world. Projections of increasing air traffic demand in conjunction with limited capacity, that is volatile and affected by exogenous random events, represent a major problem in aviation system management. Air traffic delays are always present and the more air traffic increases the more the delays will increase with serious economic impacts. The scope of this dissertation was to look closer at a threefold approach to the problem of congestion in aviation.

The first part of this thesis was related to the prediction of delays and the development of a model that will make these predictions under a wide variety of distributional assumptions. In this work the mathematical construction of a continuum approximation to a queuing system was presented, that might represent a single congested resource in the National Airspace System, such as an airport, a runway, or some en route resource. This was the first time ever to consider diffusion approximation in the aviation setting. While the model formulation was based on past work done in other areas like biology, the numeric solution scheme – Finite Element Method (FEM) - was part of this work. A discrete approximation to the queue length density function was constructed by using triangular basis functions, instead of Gauss-Legendre quadrature, that have known integrals and can be easily solved. The Monte Carlo simulation was set up to serve as the ground truth to compare with the results from the diffusion approximation. It was achieved the replication of the known steady-state results from that small set of queuing systems for which equilibrium results are known in closed form. The results in such cases showed that the diffusion approximation gives exactly the same results very quickly. Furthermore, a Monte Carlo exercise was also conducted for a number of other cases whose solutions cannot be found analytically. Again, the diffusion model seemed to perform very well, and it is much faster than running large numbers of Monte Carlo simulations. This is one of the advantages of using this model. While the simulation running time increased significantly with the number of iterations, about 310 seconds for 100,000, the diffusion model was giving results in less than 10 seconds.

In the second part of this work a parsimonious language of exchange was designed, with accompanying allocation mechanisms that allow carriers and the FAA to work together quickly, in a CDM environment, to allocate scarce capacity resources. A simple mechanism was proposed that requires each airline to give to each of their flights a priority number ranging from 1 to 4. The greater the number assigned to a flight the more important this flight is. It is also required that for each flight, carriers specify the maximum delay in minutes that they would allow it to be assigned on the ground. An extension of this was also proposed, where airlines can give an additional input regarding their preferences. Airlines can declare their intent to be considered as one of two different kinds of airlines: those that would prefer getting earlier slots (at a cost of depleting their fair share faster), and those that would prefer getting a larger number of slots overall, with the understanding that some of those will likely have large delays associated with them. While the allocation mechanism - Preference Based Proportional Random Allocation (PBPRA) - used in the first part of the work was proposed in past work, in this work it was modified accordingly in order to use the suggested preference structure. Then an enhanced version of this allocation scheme was proposed, which was called Alternative Preference Based Proportional Random Allocation (A-PBPRA). This modified scheme can be related by imagining that in the first scheme, the "price" of each slot is one point of fair share, while in this second or modified scheme we allow that "price" to be higher for carriers that prefer to be allocated a smaller number of premium slots.

The results showed that the total weighted delay and weighted average delay for PBPRA is consistently lower than RBS and RBS with substitutions. Then it was examined how the second allocation mechanism proposed, A-PBPRA, works compared to RBS and RBS with substitutions. The results showed that the total weighted and weighted average delay accrued by the flights with the proposed mechanism – A-PBPRA- is also much lower than the ones with RBS and RBS with substitutions.

On a given day the slots allocated to an airline will not match exactly its fair share. Some days they will get more and some others less, so it is important to see if in the long run they will get on average what they want. As mentioned before, one of the goals of this research is to examine if this way of expressing priorities is valid. It was interesting to check if the delays occurred by the highest priority flights are less than the rest. In this problem there are two levels of randomness that were identified and were examined. The first is due to the random selection of airlines in the allocation algorithm. Even when the number of flights and slots stay the same, each time the allocation scheme is implemented the selection of airlines can differ. The second level of randomness comes from the fact that each time an AFP is implemented the number of flights that each airline has will vary. For the purpose of this research and in order to be able to have multiple repetitions of the allocation mechanism (PBPRA) with varying input data, simulation was used.

During AFPs the airlines in the long-run will be getting on average what they want. As it was observed the smaller carriers have good chances of actually getting slots in the constraint areas. For smaller airlines the variance of slots allocated tends to be smaller than the variances for the bigger carriers. It was also shown that for the flights with priorities 4 and 3 were most of them assigned to slots and most of the ones with priorities 2 and 1 were not. The weighted delays for the flights with priority 4 were less than the ones with priority 3.

In the final part of this work the monetary benefits of improvements in strategic flight predictability as manifested through carriers' scheduled block times were assessed. Airlines tend to add extra time to their scheduled block times in order to absorb delays and maintain their schedules intact as much as possible. It is a way to deal with unexpected delays that occur frequently and cause many problems to the airlines due to missed connections, crews being overtime, unhappy passengers etc. Once the variability of the actual block times reduces, the scheduled block times will reduce and airlines will be able to achieve the same levels of scheduled operations with fewer aircraft and less total crew duty time.

This work contributed in establishing the process that airlines follow to set their scheduled block times, since scheduled block time is the key component of the benefits analysis. Then the benefits of reduced SBT for some major carriers and Low Cost Carriers were estimated. Finally the number of carriers included in the analysis was extended in order to estimate the benefits across the NAS. The results show that the savings from aircraft saved and reduced pilot salaries can be very significant and reaching more than 100 million dollars per year for each of them. Also much savings can be achieved due to reduced flight attendants salaries, and they will be approximately a third of the other two categories.

All this work could be extended in various ways. For the first part, a single queue model was proposed, so one of the next steps would be to extend it for multiple queues. Then a network of queues would be appropriate to depict the NAS. Similar work has been done in the past for waterway delays, which share similar traits, such as two-way traffic between nodes, interdependence between arrivals etc. (Dai and Schonfeld, 1998). For the second part a single FCA was considered with a predetermined duration and fixed area. It would be interesting to develop an algorithm that would model the presence of 2 FCAs in certain proximity so there will be many flights that are scheduled to pass from both areas. Also another extension would be to study a moving FCA or an FCA that is terminated early or have to be extended further. This work did not examine what happens to the flights that are left without an assigned slot, so further research on this can be conducted. Finally in the last part of this work the benefits of improvement in strategic flight predictability were estimated. The work can be extended for benefits of improvement in operational flight predictability. Carriers will be able to better respond to scheduled disruptions on a particular day of operations and this will allow them to better re-accommodate passengers, reduce crew overtime and wasted crew sources.

Bibliography

Air Traffic Services Performance Focus Group (ATSP FG), CNS/ATM Focused Team (C/AFT). Airline Metric Concepts for Evaluating Air Traffic Service Performance. Technical Report, August 1999.

Ashford N., and P.H. Wright. *Airport Engineering*. Wiley-Interscience, New York, 3rd Edition, 1992.

Aviation Daily. Aviation week intelligence network market briefing. Aviation Daily, McGraw-Hill Financial, Inc., 2013

Ball M.O., and G. Lulli. Ground Delay Programs: Optimizing over the Included Flight Set Based on Distance. *Air Traffic Control Quarterly*, Vol. 12, pp 1-25, 2004.

Ball Michael, Cynthia Barnhart, Antony Evans, Mark Hansen, Yi Liu, Prem Swaroop and Vikrant Vaze. Distributed Mechanisms for Determining NAS-Wide Service Level Expectations: Year 1 Report. May 2011

Bolczak Catherine N., Jonathan H. Hoffman, Anne J. Jensen, William W. Trigeiro. National Airspace System Performance Measurement: Overview. MITRE Technical Report, MTR 97W0000035, August 1997.

BTS, Bureau of Transportation Statistics official website, www.bts.gov. Last accessed February 2013.

Carr G. C., H. Erzberger and F. Neuman. Airline arrival prioritazation in sequencing and scheduling. *2nd USA/Europe Air Traffic Management Research and Development Seminar*, Orlando, FL, 1998.

Churchill A., K. Vlachou and D.J. Lovell. Filtering and aggregation schemes for delay model calibration. *3rd International Conference on Research on Air Transportation, ICRAT*, Washington DC, 2008.

Dai M. D. M. and P. Schonfeld. Metamodels for Estimating Waterway Delays through Series of Queues. *Transportation Reseasch Part B*, Vol. 32, No. 1, pp 1-19, 1998

Deshpande V., and M. Arikan. The Impact of Airline Flight Schedules on Flight Delays. *Manufacturing & Service Operations Management, Articles in Advance*, pp. 1–18, 2012.

deNeufville R., and A. Odoni. *Airport Systems: Planning, Design, and Management*. The McGraw Hill, New York, 2003.

Dinar A., M.W. Rosengrant, and R. Meinzen-Dick. Water Allocation Mechanisms: Principles and Examples. World Bank paper, World Bank, Washington, DC, 1997.

EUROCONTROL Performance Review Commission. A Comparison of Performance in Selected US and European En-Route Centers. May 2003.

Federal Aviation Administration. FAA Aerospace Forecast, Fiscal Years 2013-2033. U.S. Department of Transportation, Federal Aviation Administration Policy and Plans, 2013a.

Federal Aviation Administration, NextGen Implementation Plan, 2013b.

Federal Aviation Administration. Traffic Flow Management in the National Airspace System. Federal Aviation Administration, Air Traffic Organization, October 2009.

Federal Aviation Administration. Collaborative Trajectory Options Program (CTOP) Interface Control Document (ICD) for the Traffic Flow Management System (TFMS). FAA Air Traffic Organization, Contract Number: DTFAWA-04-C-00045, 2011a.

Federal Aviation Administration (FAA). Performance (Service) Based NAS ATM. Air Traffic Organization (ATO) White Paper, June 2011b.

Flow Evaluation Team (FET). Operating in a CTOP (Collaborative Trajectory Options Program) Environment. FAA Collaborative Decision Making Program, FAA Human Factors Division, February 2014.

Gaver, D.P., Jr,. Diffusion approximations and models for certain congestion problems. *Journal of Applied Probability*, 5(3), pp.607-623, 1968.

Gibson S. Allocation of Capacity in the Rail Industry. *Utilities Policy*, Vol. 11, pp. 39-42, 2003.

Gulding John, David Knorr, Marc Rose, James Bonn, Philippe Enaud, and Holger Hegendoerfer. US/ Europe comparison of ATM-related Operational Performance. *Eighth USA/Europe Air Traffic Management Research and Development*, 2009.

Hansen, M., T. Nikoleris, D. Lovell, K. Vlachou, and A. Odoni. Use of queuing models to estimate delay savings from 4D trajectory precision, *Proceedings of the 8th USA/Europe Air Traffic Management Research and Development Seminar*. Napa Valley, CA, 2009.

Hoffman R., T. Lewis, and R. Jakobovits. Flow Constraint Area Rerouting Decision Support Tool Phase I SBIR. Metron Aviation, Technical Report, 32N0704-001-R0, July 2004.

Horanjic, B. R.,. *Some queuing models of airport delays.* S.M. Thesis. Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, 1990.

International Air Transport Association. Worldwide Slot Guidelines. International Air Transport Association, 2nd Edition, January 2012.

International Civil Aviation Organization (ICAO). Global Air Traffic Management Operational Concept. First Edition – 2005.

Joint Economic Committee Majority Stuff. Your flight has been delayed again: Flight delays cost passengers, airlines, and the U.S. economy billions. May 2008.

Kiker C., and G.D. Lynne. Water Allocation under Administrative Regulation: Some Economic Considerations. *Southern Journal of Agricultural Economics*, pp. 57-63, 1976.

Kimura, M. Diffusion models in population genetics. *Journal of Applied Probability,* 1(2), pp.177-232, 1964.

Kimura, T. Diffusion approximation for an M/G/m queue. *Operations Research,* 31(2), pp.304-321, 1983.

Kivestu, P. Alternative Methods of Investigating the Time-dependent M/G/K Queue. S.M. Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, 1976.

Kobayashi, H. Application of the diffusion approximation to queuing networks II: Nonequilibrium distributions and applications to computer modeling. *Journal of the ACM,* 21(3), pp.459-469, 1974.

Kumar, M., P. Singla, S. Chakravorty, and J. L. Junkins. The partition of unity finite element approach to the stationary Fokker-Planck equation. *Proceedings of the AIAA/AAS Astrodynamics Specialist Conference.* Keystone, CO, 2006.

Lee, D. A., P. F. Kostiuk, B. J. Kaplan, M. Escobar, A. Odoni, and B. Malone. *Technical and economic analysis of air transportation management issues related to free flight*. Logistics Management Institute , NS501T1, 1997.

Lee, D. A., D. Long, M. R. Etheridge, J. R. Plugge, J. P. Johnson, and P. F. Kostiuk. A method for making cross-comparable estimates of the benefits of decision support technologies for air traffic management. Logistics Management Institute, NS710S1, 1998.

Liu, P.-c. B., M. Hansen, and A. Mukherjee. Scenario-based air traffic flow management: From theory to practice. *Transportation Research Part B*, 42, pp.685-702, 2008.

Long, D., D. Lee, J. Johnson, E. Gaier, and P. Kostiuk. Modeling Air Traffic Management Technologies with a Queuing Network Model of the National Airspace System. Technical Report CR-1999-208988, National Aeronautics and Space Administration, Hampton, VA, January 1999.

Lovell, D. J., K. Vlachou, T. Rabbani, and A. Bayen (2013). A diffusion approximation to a single airport queue. *Transportation Research Part C: Emerging Technologies*, vol. 33, pp. 227-237.

Malone, K. *Dynamic Queuing Systems: Behavior and Approximations for Individual Queues and Networks*. Ph.D. Dissertation, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, 1995.

Malone, K. and A. Odoni. The Approximate Network Delays Model,Working Paper, Operations Research Center, MIT, 2001.

Metron Aviation official website. http://www.metronaviation.com/ . Last access January 2012.

Millner, D. C. Design of the NASPAC simulation modeling system. The MITRE Corporation, MTR 92W0000135, FAA Contract No. DTFEA01-93-C-00001, 1993.

Moreno-Hines, F., and D. Kirkman. Assessing the Nextgen avionics business case from the airline perspective: The implications of airline responses to changes in operational performance. 10th USA/Europe Air Traffic Management Research and Development Seminar (ATM2013), Chicago, IL, USA, 2013.

Mueller, E. R. and G. B. Chatterji. Analysis of Aircraft Arrival and Departure Delay Characteristics. *AIAA's Aircraft Technology, Integration, and Operations Forum*, Los Angeles, California, 2002.

Nash C., S. Coulthard, and B. Matthews. Rail Track Charges in Great Britain – The Issue of Charging for Capacity. *Transport Policy*, Vol. 11, pp. 315-327, 2004.

Newell, G.F. *Applications of Queueing Theory*. Chapman-Hall, 1971.

Odoni A. The Flow Management Problem in Air Traffic Control. In A.R. Odoni, L. Bianco, and G. Szego, editors, *Flow Control of Congested Networks*, 269-288. Springer-Verlag, Berlin, 1987.

Pepper, D. W., and J. C. Heinrich. *The finite element method: Basic concepts and applications*. Taylor & Francis, 1992.

Post, J., J. Gulding, K. Noonan, D. Murphy, J. Bonn, and M. Graham. The modernized National Airspace System Performance Analysis Capability (NASPAC), *Proceedings of the 26th Congress of the International Council of the Aeronautical Sciences (ICAS)*. Anchorage, AK, 2008.

Stalnaker S.E., J.S. DeArmon, and R.D. Katkin. Collaborative Airspace Congestion Resolution (CACR) Benefits Analysis. *28th Digital Avionics Systems Conference*, Orlando, Florida,October 25-29, 2009.

Taylor, M. A. P. Travel Time Variability – The Case of Two Public Modes. *Transportation Science*, Vol. 16, No 4, pp. 507-521, November 1982.

Uniman David L., John Attanucci, Rabi G. Mishalani, and Nigel H. M. Wilson. Service Reliability Measurement Using Automated Fare Card Data Application to the London Underground. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2143, Transportation Research Board of the National Academies, Washington, D.C., pp. 92–99, 2010.

Vakili N. *Preference Based Fair Allocation of Limited Airspaces Resources*. Ph.D. Thesis, University of Maryland, 2009.

Vlachou, K. and D. J. Lovell. Mechanisms for equitable resource allocation when airspace capacity is reduced. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2325, pp. 97-102, 2013.

Wang, L. Z., L. Fang, and K. W. Hipel. Water Resources Allocation: A Cooperative Game Theoretic Approach. *Journal of Environmental Informatics*, Vol 2 (2), pp. 11-22, 2003.

Young P. H. *Equity in Theory and Practice*. Princeton University Press, Princeton, New Jersey, 1994.