Stochastic Queueing Models for Air Transportation Systems with Scheduled Arrivals

By

Anastasios Nikolaos Nikoleris

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Engineering – Civil and Environmental Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Mark Hansen, Chair
Professor Samer Madanat
Professor Rhonda Righter

Spring 2011
Abstract
Stochastic Queueing Models for Air Transportation Systems with Scheduled Arrivals

by
Anastasios Nikolaos Nikoleris

Doctor of Philosophy in Engineering – Civil and Environmental Engineering
University of California, Berkeley
Professor Mark Hansen, Chair

In this thesis we examine a queueing system with a single server under 4D trajectory (4DT) aircraft operations. In the definition of 4DT, we consider time as the fourth dimension. Thus, under 4DT’s aircraft are assigned scheduled times of arrival at a fix, which they meet with some stochastic error. We start our analysis by assuming a normal distribution for the error, and that aircraft enter service according to a First-Scheduled-First-Served (FSFS) queue discipline. We develop a recursive queueing model that employs the Clark approximation method to analytically estimate the mean and variance of aircraft delays. The accuracy of the approximation method is assessed through simulation experiments, which indicate good accuracy of the Clark method in estimating total system delays. Using a wide range of representative demand and capacity scenarios at seven major US airports, we compare the model estimates for average queueing delay per flight with those from a deterministic queueing model. We find that the estimates of expected queueing delay from the stochastic and the deterministic model are strongly correlated and very similar, except for cases of low airport utilization, characterized by average deterministic queueing delay over all arrivals smaller than one minute.

Next, we study a simplified situation in which a sequence of aircraft with the same 4DT execution accuracy are assigned scheduled times of arrival at a fix with constant excess time separation between them. Under these assumptions, the expected delay to a flight from imperfect trajectory adherence – which we term stochastic delay – depends on the excess time separation, or buffer, expressed as a ratio to the trajectory imprecision, as well as the place of the flight in the sequence. As the buffer goes up, the stochastic delay goes down, but at the cost of increased deterministic delay from reduced capacity. If stochastic delay costs more than deterministic delay, then the optimum buffer is greater than zero, but quite small under plausible cost ratios and trajectory precision levels.

We also explore the effect of queue disciplines that give priority to aircraft equipped with avionics that enable them to execute 4D trajectories with high precision. We find that by switching from a FSFS to a Best-Equipped-Best-Served (BEBS) policy, total delay in the system can remain at the same level, while achieving significant delay savings for equipped aircraft.
The basic queueing model is extended to analyze a system for aircraft landings at a single runway under 4D trajectory-based operations. The server of this queueing system is the runway threshold, at which aircraft are assigned scheduled times of arrival. In accord with evidence from a variety of sources, we assume a Gumbel distribution for aircraft’s stochastic lateness and for runway occupancy times (RT). We propose an approximate solution to analytically estimate the mean and variance of queueing delays, the accuracy of which is demonstrated through simulation experiments.

Next, we study again a simplified situation, in which a sequence of aircraft with the same capability of adherence to 4D trajectories, minimum headway, and runway occupancy time distribution are metered at a constant rate. We investigate the relationship between buffer time between scheduled arrivals and expected loss in system efficiency, defined as the weighted sum of delays due to imperfect trajectory adherence and varying RT – which we term stochastic delay – and due to reduced throughput. We find that if stochastic delay costs more than deterministic delay, then the minimum expected loss in efficiency is attained for buffer values greater than zero. Under conditions that we consider realistic, such a buffer would reduce planned throughput about 15% compared to the maximum possible under deterministic conditions. It is also shown that when RT is the binding factor that determines throughput, stochastic delays are substantially higher, compared to a situation where minimum headway is the principal constraint. For the former cases, it is found that the marginal benefit from improving trajectory precision diminishes when the standard deviation of adherence error is less than 0.5 times that of RT. Therefore, when runway throughput is controlled primarily by aircraft’s time to exit the runway, improving 4D trajectory precision yields delay savings, but at a decreasing rate.

In order to investigate the potential increase in runway throughput from implementation of 4D trajectories, we study the case of paired arrivals at the San Francisco International Airport. A mathematical model is developed that estimates the scheduling headway between two consecutive pairs of landing aircraft. This headway minimizes the time interval between consecutive arrival pairs, while allowing sufficient time for a pair of departing aircraft to take-off in the meantime. Results derived from a simplified case study indicate potential increases in landing throughput by two aircraft per hour for every second of decrease in the standard deviation of adherence error. Additionally, we compute the upper bound in landings brought by almost perfect adherence to 4DT’s, which appears as high as 81 aircraft/hour, an increase of 47% from current level of 55 aircraft/hour.
To my parents,
Varvara and Nikolaos,
and to my brother,
Vangelis
Contents

List of Tables ................................................................................................................... iv

List of Figures .................................................................................................................. v

Acknowledgements ......................................................................................................... vii

1. Introduction ................................................................................................................. 1

   1.1 Motivation and research goal .................................................................................. 1

   1.2 Background in air traffic management ..................................................................... 3

       1.2.1 Handling of a typical airline flight .................................................................. 3

       1.2.1 Major transformations on ATM introduced by NextGen ............................... 6

   1.3 Literature review and research contributions ......................................................... 6

2. Basic Model .................................................................................................................. 10

   2.1 General case .......................................................................................................... 10

       2.1.1 Model formulation ......................................................................................... 10

       2.1.2 Solution with the Clark approximation method .............................................. 13

       2.1.3 Approximation error ....................................................................................... 15

       2.1.4 Comparison with deterministic queueing model ............................................. 17

       2.1.5 System with two servers ................................................................................ 19

   2.2 Special case with constant buffer and uniform precision error ............................. 21

       2.2.1 Estimation of stochastic delay ......................................................................... 25

       2.2.2 Solution with the Clark approximation method .............................................. 28
2.3 Trade-offs between deterministic and stochastic delay .......................... 29
2.4 Evaluation of priority queue disciplines ............................................. 33
  2.4.1 Description of working scenarios ............................................... 33
  2.4.2 Results ......................................................................................... 35

3. Model with runway occupancy time .................................................... 38
  3.1 General model .................................................................................. 38
    3.1.1 Model formulation ................................................................. 39
    3.1.2 Approximate solution ............................................................. 41
    3.1.3 Approximation error tests ....................................................... 42
  3.2 Using buffers to mitigate the effects of stochastic runway occupancy time and
    precision error .................................................................................. 45
    3.2.1 A simplified case ................................................................. 45
    3.2.2 Results ......................................................................................... 48

4. Paired Arrivals at SFO .................................................................. 56
  4.1 Description of SFO terminal airspace ............................................ 56
  4.2 Model for estimating optimal headways ....................................... 58
  4.3 Numerical example ........................................................................ 60

5. Conclusions ....................................................................................... 63

Bibliography ......................................................................................... 66
# List of Tables

Table 2.1: Results of accuracy tests of the Clark approximation method…………………16

Table 2.2: Comparison of deterministic and stochastic models when $\sigma \in \{10, 20, 30, 40, 50, 60\}$ seconds………………………………………………………………………………………………18

Table 2.3: Comparison of deterministic and stochastic models when $\sigma \in \{10, 20, 30, 40, 50, 60\}$ seconds and $Q^D_k \leq 1$ minutes/flight……………………………………………………………………………………………………19

Table 2.4: Average percent change of total delays between FSFS and BEBS policies……..37

Table 3.1: Approximation error test results when 4DT precision has standard deviation $\sigma = 16$ seconds……………………………………………………………………………………………………44

Table 3.2: Approximation error test results when 4DT precision has standard deviation $\sigma = 8$ seconds……………………………………………………………………………………………………44

Table 3.3: Approximation error test results when 4DT precision has standard deviation $\sigma = 4$ seconds……………………………………………………………………………………………………44

Table 4.1: Results of numerical example……………………………………………………62
List of Figures

Figure 1.1: Idealized planar view of terminal area arrivals pattern.............................................4
Figure 1.2: Summary of the role that each of the three control facilities plays in ATM..............5
Figure 2.1a: Cumulative number of scheduled arrivals at CLT on 2/21/2010.............................22
Figure 2.1b: Cumulative number of scheduled arrivals at SLC on 2/16/2010.............................22
Figure 2.1c: Cumulative number of scheduled arrivals at IAH on 2/12/2010............................23
Figure 2.1d: Cumulative number of scheduled arrivals at ATL on 2/2/2010............................23
Figure 2.2: Queueing diagram with deterministic and stochastic delays.................................24
Figure 2.3: Expected delay $E(\hat{Z}_i)$ for several values of $\Delta$..................................................26
Figure 2.4: Normalized Expected Total Delay $E(W^*_n)$ for several values of $\Delta$......................27
Figure 2.5: Contour plot of $\epsilon(\%)$ in comparing approximate and numerical results for $E(\hat{Z}_i)$...29
Figure 2.6: Expected loss from deterministic and stochastic delay for N=20 aircraft...............30
Figure 2.7: Expected loss from deterministic and stochastic delay for N=40 aircraft................31
Figure 2.8: Expected loss from deterministic and stochastic delay for N=60 aircraft...............31
Figure 2.9: Optimal buffer as a function of $\beta$ and N...............................................................32
Figure 2.10: Average total delays under various queue disciplines and penetration rates..........36
Figure 3.1: Histogram for runway occupancy times measured at runway 17C at DFW............40
Figure 3.2: Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=1$, $N=100$, and $\Delta=\Theta+2$.  

Figure 3.3: Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=1$, $N=100$, and $\Theta=\Delta+2$.  

Figure 3.4: Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=1$, $N=100$, and $\Theta=\Delta$.  

Figure 3.5: Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=2$, $N=100$, and $\Delta=\Theta+2$.  

Figure 3.6: Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=2$, $N=100$, and $\Theta=\Delta+2$.  

Figure 3.7: Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=1$, $N=100$, and $\Theta=\Delta$.  

Figure 3.8: Minimum expected losses as a function of $\zeta$, for $\beta=1$, and $N=100$.  

Figure 3.9: Minimum expected losses as a function of $\zeta$, for $\beta=2$, and $N=100$.  

Figure 3.10: Comparison of two queueing systems with different capacities and unequal efficiency losses.  

Figure 4.1: Diagram of San Francisco International Airport.  

Figure 4.2: Rate of landings as a function of precision in executing 4D trajectories.
Acknowledgements

“I am indebted to my father for living, but to my teacher for living well”. ~ Alexander the Great

With these words in mind, I would like to express my deep gratitude to my advisor, Professor Mark Hansen, to whom I am indebted for thinking well. Mark, with his inventive thinking, has taught me how to get “In through the outdoor” anytime I set out to attack a problem, scientific or not. I am very fortunate to have studied under the guidance of an extremely talented and gifted academic researcher and teacher. I would like to thank Mark for being patient with me, understanding my character, and supporting me through believing in me. Finally, his unique sense of humor made every single day in McLaughlin special and provided me with numerous funny stories to remember.

I am also very grateful to Professor Alex Skabardonis for always being there for me. His support and friendship have defined my life in Berkeley. Things would have been very different without Alex.

My gratitude is extended to all Berkeley faculty members. Professor Samer Madanat’s encouragement to pursue a PhD played a pivotal role in my decision to return to graduate school. Moreover, I would like to thank him for his very useful comments throughout this research effort. Prof. Rhonda Righter’s comments and excellent review of this thesis are also greatly appreciated.

Being a member of the Berkeley transportation group provided me the opportunity to interact with many distinguished colleagues. To name a few, I thank my friends and lab-mates
Gurkaran Buxi and Bo Zou with whom I have shared a lot of great moments, and also Gautam Gupta, Amy Kim, Megan Smirti Ryerson, Claus Gwiggner, Xavier Prats, Yoonjin Yoon, Frank Ketcham, and Eric Gonzales. I would also like to thank Professor David Lovell for his constructive feedback during the legendary NASA teleconferences, and also Professor Dimitrios Tsamboulas for initially fostering my interest in transportation engineering studies.

I would like to express my gratitude to NASA Ames for providing me with the financial support needed to conduct this research. Special thanks goes to Todd Farley for his enthusiastic support. I am also thankful to the U.S. Department of Transportation Eisenhower Transportation Fellowship Program for awarding me travel grants to attend the Transportation Research Board Annual Meeting.

My years in Berkeley have been full of valuable and intense experiences, and I have had many amazing friends to share them with. I deeply thank Boriska Toth, Adonis Garefalakis, Dimitris Katsis, Panos Patatoukas, Eleni Christofa, Alex Dimakis, Lefteris Matsikoudis, Nikolas Geroliminis, and Laura Moussa. Special thanks goes also to Costas Daskalakis, Anthony Papavassiliou, George Pierrakos, Kostas Sarrigeorgidis, and Yannis Stefanou.

Finally, words remain trivial when it comes to thanking my parents, Nikos and Varvara, and my brother Vangelis, for their love and support. I dedicate this dissertation to them.
1. Introduction

The first chapter in this dissertation aims to delineate the context that motivated this research effort. It starts with an introduction to the Next Generation Air Transportation System and the main changes that it is expected to bring. In order to assist the reader in understanding the latter, a brief description of air traffic management practices in the US is provided. A thorough review of the literature on aircraft queueing model follows, and the chapter ends with highlighting the key contributions of this research.

1.1 Motivation and research goal

Recent studies by the Federal Aviation Administration (FAA) project a significant increase in demand for air travel for the next two decades, which may result in doubling or even tripling of air traffic operations (JPDO, 2010). At the same time, analysis of even the conservative growth estimate shows a significant lack of existing and planned capacity (Solomos et al., 2004). Airport capacity is primarily a function of weather conditions, number of available runways, and the air traffic control system in place. Since building additional runways constitutes a costly, lengthy, and often infeasible project due to scarcity of available land, the FAA has undertaken a mega-project to upgrade the existing air traffic control infrastructure. That will introduce a set of new technologies and processes, aiming to transform the current ground-based system of air traffic control into a satellite-based system of air traffic management, often referenced as the Next Generation Air Transportation System (NextGen).

This transformation aims at establishing a system where aircraft can fly their preferred trajectories, with high precision and reduced need for tactical air traffic control. To
leverage that, the FAA is planning to introduce a set of new air traffic control technologies. According to FAA’s proposed architecture of NextGen, precise GPS-based navigation, known as Required Navigation Performance (RNP), will enable aircraft to fly their preferred routes with high precision. GPS-based surveillance functions, such as Automatic Dependent Surveillance-Broadcast (ADS-B) technology, will facilitate the distribution of information on aircraft’s kinematic state to nearby aircraft and ground controllers, improving pilots’ ability of monitoring neighboring traffic and self-separating from it. Moreover, the Ground Based Augmentation System (GBAS) is an aircraft landing system based on real-time differential correction of the GPS signal. When in place, GBAS will provide guidance over an entire volume of airspace leading down to the runway, thus addressing a limitation of current Instrument Landing Systems (ILS) that define a single, straight-line final approach path (de Neufville & Odoni, 2003).

The above new technologies are considered by NextGen planners as necessary elements for system-wide application of trajectory-based aircraft operations. Under four-dimensional trajectory (4DT) operations aircraft fly an assigned 3D trajectory in space and meet scheduled times of arrival at points in airspace, with high precision. 4DT capability constitutes a cornerstone of FAA’s vision for NextGen, and particularly of an operational environment in which aircraft fly their preferred trajectories with reduced need for tactical air traffic control. A principal goal of 4DT operations is to increase aircraft flows, which will rely on controlled times of arrival at congested resources, including entry to terminal airspace areas and landing on runways (JPDO, 2010). Among other bottlenecks of the airspace system that can potentially stand to gain from NextGen are airports whose runway capacity is highly sensitive to weather conditions. For example, implementation of precise 4DT approaches to San Francisco International Airport (SFO) can yield significant benefits in runway throughput, enabling closely-spaced landings under instrument meteorological conditions.

In order for airlines to equip their fleet with avionics that enable 4DT operations, the investment in this technology should demonstrate benefits that outweigh costs. The Federal Aviation Administration is currently examining incentive mechanisms for airlines to invest to new technology, which include giving priority to aircraft equipped with NextGen avionics.

Even with the deployment of the very best 4D trajectory precision and navigation tools, adherence to 4D trajectories — in particular to scheduled arrival times at airspace fixes — will not be perfect. Sources of imprecision include variation in aerodynamic performance, errors in wind prediction, variations in flight crew technique, and varying degrees of exactitude in navigational performance (JPDO, 2010). As an order of magnitude, NextGen planners foresee accuracies of ±10 seconds in aircraft meeting scheduled times of arrival (Swenson, Barhydt, & Landis, 2006). This is well under existing separation requirements, which typically range between one and two minutes (Hansen, 2002), but sufficient to cause delays that may propagate backwards and affect upstream aircraft. That raises the question on how the throughput performance of airspace resources (e.g. runways, en-route sectors) is related to the level of trajectory precision.
One could tackle that question by performing a simulation. However, simulations have two main drawbacks: first, they typically require a significant amount of effort to set up, and second they have a long execution time.

A simplified and yet accurate way of estimating delays incurred by aircraft that must traverse a congested resource is by means of a queueing model. The limited input requirements and short computational times of analytical queueing models facilitate the exploration of a wide range of demand/capacity scenarios, which renders them as useful tools for strategic airport planning, as well as for applications that can be executed in real time.

Therefore, the central goal of this research is to develop an analytical queueing model for aircraft operations in NextGen, which accounts for different levels of 4D trajectory precision. That not only will allow for a fast and accurate analysis method on airspace delay and capacity issues, but also for an analysis of the distribution of delay savings among the users of the air transportation system.

1.2 Background in air traffic management

The infrastructure of the aviation system consists of two principal elements, airports and air traffic management (ATM) systems. The ATM system provides a set of services aimed at ensuring the safety and efficiency of air traffic flows (de Neufville & Odoni, 2003). In the following subsections we briefly describe selected concepts of ATM that serve as background knowledge for topics analyzed later in this dissertation.

1.2.1 Handling of a typical airline flight

The successive phases of a typical flight in the United States, flying under Instrumental Flight Rules (IFR) between two airports, are controlled by three basic types of facilities: the airport traffic control tower, the approach control facility or Terminal Radar Approach Control (TRACON), and the Air Route Traffic Control Centers (ARTCC). In brief, the airport tower manages ground operations at the airport and monitors the safety of landings and take-offs within the final approach to the airport. The role of the TRACON controllers is to generate runway assignments and landing times for each flight to maximize runway utilization while allowing for safe separation between flights (Meyn & Erzberger, 2005). The ARTCC (or Center) ensures safe separation between flights while airborne and detection of potential conflicts; additionally, a Center adjacent to a TRACON sequences flights into streams and hands them off to the TRACON at the arrival fixes (Figure 1.1). Figure 1.2 provides a summary of the role that each of the three control facilities plays in ATM.
Figure 1.1: Idealized planar view of terminal area arrivals pattern. (Source: Mundra (1989))
<table>
<thead>
<tr>
<th>Type of Facility</th>
<th>Terminal Area Facilities</th>
<th>En-Route Facilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlling Facility</td>
<td>Airport Traffic Control Towers</td>
<td>Approach Control Facilities</td>
</tr>
<tr>
<td>Type of Control</td>
<td>Ground Traffic Control Takeoff and Landing Control</td>
<td>Approach and Departure Control</td>
</tr>
<tr>
<td>Airspace</td>
<td>Airport Traffic Area</td>
<td>Approach Control (TRACON Area)</td>
</tr>
<tr>
<td>Airspace</td>
<td>Typically 5 nmi and 3000 ft AGL</td>
<td>Typically extending up to 40 nmi + 10000 ft from the airport</td>
</tr>
<tr>
<td>Transitional Phase</td>
<td>Cruising Phase</td>
<td>Up to 60000 ft</td>
</tr>
<tr>
<td>Typical Flight Time</td>
<td>Typical Ground Time 5-10 min</td>
<td>Typical Flight Time 10-20 min</td>
</tr>
</tbody>
</table>
1.2.2 Major transformations on ATM introduced by NextGen

According to FAA’s planning of NextGen (JPDO, 2010), the most significant changes in ATM will be facilitated by advances on trajectory-based operations, performance-based services (PBS), separation management, airspace allocation, and weather prediction. A short description of those follows:

**Trajectory-based Operations.** As discussed in Section 1.1, under 4DT operations access to congested resources of the airspace system is managed by assigning to each aircraft a specific time of arrival at an airspace point located inside or near the congested resource. This type of operations is also known as time-based metering, and it is currently used only in nine US airports, within approximately a 200 nmi radius from the airport and only for arrival streams (Landry, Farley, & Hoang, 2005).

**Performance-Based Services (PBS).** In NextGen, several operating environments will be available, and pilots will be able to select the service level appropriate for their flight according to aircraft performance requirements for each level. In this way, PBS will become a tool for controlling access to highly constrained resources or complex operating environments.

**Separation Management.** Currently, individual controllers monitor aircraft trajectories on a radar screen and issue voice instructions to the pilots for the tactical separation of flights. In NextGen, that will rely heavily on digital communications between controller and flight deck, and also on automation (e.g. decision support tools). That is expected by the FAA to enable reduced separation standards.

**Dynamic Airspace Configuration.** This is a capacity management technique to address imbalances on capacity and demand. Contrary to the current static configuration of airspace, changes to airspace configuration will be dynamic in NextGen, accommodating variations in air traffic demand.

**Weather Integration.** Enhanced probabilistic weather forecasting will enable the incorporation of information about en-route weather conditions into the 4D trajectory flight plan, thus minimizing disruptions due to unexpected weather phenomena.

Having discussed the key characteristics of air traffic operations, as they are expected to take place in NextGen, we proceed to examine whether existing queueing models can capture these features.

1.3 Literature review and research contributions

A review of the literature indicates that available aircraft queueing models are based upon assumptions that cannot capture key features of precise 4DT aircraft operations as well as
the impact of different levels of 4DT precision on system throughput. The classical work of Koopman (1972) provides systems of differential and difference equations for an \( M(t)/M(t)/1 \) and an \( M(t)/D(t)/1 \) queueing system respectively, that one has to solve numerically in order to obtain transient solutions of the queue length probability distribution. Based on Koopman’s work, Kivestu (1976) proposed an \( M(t)/E_k(t)/1 \) system where the service time follows an Erlang-\( k \) distribution. Common in these models is the assumption of a time-dependent Poisson point process for arrivals at the server, implying the variance in the total number of arrivals over a given time interval is constant and equal to its mean. However, in a system where customers have scheduled times of arrival at a server, which they meet with high precision, one should expect to see little variation in the total number of arrivals over a certain time interval. Moreover, as already discussed, we are concerned with an arrival process where aircraft are assigned scheduled times of arrival at airspace fixes. For such a process, Guadagni et al. (2010) show that when imprecision in meeting those times is large the resulting queue is very different from the Poisson case. Further, a model with Poisson arrivals cannot capture the effect of varying levels of precision in 4DT operations, because no controls are allowed for the reduction of this variance. The latter argument also holds for models that assume a deterministic number of arrivals over any time period (Peterson, Bertsimas, & Odoni, 1995). A fundamentally different approach was proposed by Newell (1979), who argues for data collection and for drawing graphs of the cumulative number of requests for landing or take-off, and the cumulative number of actual landings or take-offs. Such a model, at least in its original form, cannot incorporate precision as a model parameter and hence it cannot capture the impact of various levels of precision on aircraft delays.

Queueing models with scheduled arrivals have been proposed to study situations where customers are scheduled to arrive at the server at equal time intervals, but the actual arrival times exhibit some stochastic lateness (positive or negative). Mercer (1973) proposes a model where customers are served in a FIFO queue discipline, the lateness distribution is general, and customers do not join the system if they do not arrive in the interval for which they were scheduled. Sabria and Daganzo (1989) examine systems where customers must be served in a First-Scheduled-First-Served (FSFS) order. They develop a recursive equation for the customer departure times from the system that can be solved numerically, whereas exact transient solutions are obtained for the case when the lateness distribution is Gumbel with uniform scale parameter across all customers and service times are constant. Guadagni et al. (2011) focus on lateness distributions with large but finite standard deviation, finite support, constant service times, and a FIFO order of service. While the arrival process is shown to converge to a Poisson, the resulting queue is found to be very different.

Systems with light traffic, such that successive customers rarely interfere with each other, have been examined by Bloomfield and Cox (1972). Approximate expressions for several properties of the system are obtained for busy periods that involve two or three customers. For such traffic intensities, Sabria and Daganzo derive approximate closed-form expressions for the steady-state queueing delay assuming that stochastic latenesses are normally distributed, and service times follow a shifted type-2 Erlang distribution.
More recently, Jouini and Benjaafar (2010) consider systems where intervals between scheduled arrivals, lateness distributions, service time distributions (that belong to the exponential family), as well as probabilities of no-show are heterogeneous across customers. They also assume that customers arrive in the same order as their appointment times, by employing lateness distributions with finite and small enough support. The authors provide expressions for the expected time spent in the system of any customer that can be solved numerically, whereas closed-form solutions are derived if the lateness distribution is either uniform or triangular.

In this research we develop a queueing model with scheduled aircraft arrivals, while we extend previous work in several directions. We assume, in the first part of our analysis, a normal distribution for stochastic deviations from schedule, and we employ the Clark approximation method to derive an analytical transient solution to the general case, which includes heterogeneous headways between scheduled times of arrival, deterministic and non-constant service times, correlation between stochastic deviations from schedule, while it does not make any assumption on traffic intensity. Moreover, the model is also extended to account for the impact of runway occupancy time (RT) on delays by explicitly modeling it as a random variable. For this case, a Gumbel distribution is assumed for stochastic lateness of aircraft and RT distribution, while we relax the assumptions of uniform scale parameter and independency across random variables, by proposing an approximate solution for the maximum of correlated and non-identical Gumbel random variables.

In addition to developing a queueing model, we will investigate the effect on system performance of increasing the headway between scheduled times of arrivals, and how that effect varies as a function of dispersion in runway occupancy times and aircraft’s imprecision in meeting their assigned times of arrival at the runway threshold. For example, scheduling arrivals at the minimum allowed headway results in maximum throughput under saturated conditions, but any late arrivals cause delays that propagate backwards to upstream aircraft. Thus, one may want to insert some additional time separation, or “buffer”, between scheduled arrivals in exchange of a more stable system that can absorb delays due to unpunctual arrivals. Moreover, while precise 4DT operations will undoubtedly increase system throughput, their effect may be blunted as a result of variability in aircraft’s time to clear the runway. We investigate these trade-offs for a situation where model parameters are the same among all aircraft, and find the buffer size that minimizes the total cost of delay, defined as the weighted sum of delays due to unpunctual arrivals and exits from the runway, and due to losses in throughput.

Finally, aside from our analysis on the effect of 4DT precision on queueing delays, we are also interested in the effect of trajectory adherence error on runway throughput. For that, we study the case of paired arrivals at the San Francisco International Airport. We develop a mathematical model that estimates the optimal scheduling headway between two consecutive pairs of landing aircraft, as a function of aircraft’s precision in meeting assigned times of arrival. In this way, we quantify the benefit in landing throughput that is capacitated by precise 4DT operations.
The rest of the dissertation is organized as follows: Chapter 2 presents the basic form of our model and its approximate solution, and reports on experiments conducted to assess the accuracy of the model against simulation. The model is applied to investigate system performance under a Best-Equipped-Best-Served policy, which gives priority to aircraft that are equipped with new avionics. A special case is also examined, in which 4DT precision levels are the same for all aircraft. Working with a queueing model in this simplified setting enables us to determine an optimal metering rate that attains high throughput while keeping delays due to imprecise adherence to 4DT’s small. In Chapter 3 we extend our model to include runway occupancy times as a separate random variable. An approximate analytical solution for the mean and variance of flight delays is proposed, supported by simulation experiments on its accuracy. We again examine a simplified case, and investigate the trade-off between runway throughput and excess time separation, as a function of 4DT precision and stochastic variation in RT. Chapter 4 presents a methodology to assess benefits in landing throughput from increased 4DT precision for the case of closely-spaced arrivals at the San Francisco International airport. Finally, Chapter 5 summarizes our main findings and discusses the conclusions of this research.
2. Basic Model

In Section 2.1 we introduce the general form of the model and its approximate solution, and report on experiments conducted to assess the accuracy of the model against simulation. In addition, we present a comparison of the model estimates with those from a deterministic queueing model, under a wide range of demand and capacity scenarios at six major US airports. Section 2.1 ends with an extension of the model to the case of two servers. In Section 2.2 we consider a special case, in which 4DT precision levels are same among all aircraft. Working with a queueing model in this simplified setting enables us to determine (Section 2.3) an optimal buffer time to be inserted between arrivals that attains high throughput while keeping delays due to imprecise adherence to 4DT’s small. The chapter concludes with analysis of an alternative queue discipline that prioritizes aircraft equipped with 4DT adherence technology.

2.1 General case

We start by analyzing the general case, which accepts as inputs heterogeneous precision error distributions and an arbitrary schedule of aircraft arrivals.

2.1.1 Model formulation

Our queueing system consists of a single server, which is a fix (either a point in the airspace or a runway’s threshold), and of flights that must cross this fix. Aircraft are assigned scheduled times of arrival at the fix, and they fly 4D trajectories to arrive at the
fix at their scheduled times. However, due to imprecision in trajectory adherence, actual times of arrival at the fix have some stochastic deviation from the scheduled times. The sources of imprecision might include airframe-to-airframe variation in aerodynamic performance, limitations in wind prediction capability, variations in flight crew technique, and varying degrees of exactitude in navigational performance (JPDO, 2010). In addition, consecutive aircraft \( i - 1 \) and \( i \) must maintain a minimum headway \( h_{i-1,i} \), which can vary over pairs of arriving aircraft, for safety reasons. Since air traffic controllers, with guidance from separation rules, decide values for \( h_{i-1,i} \), we consider it as a deterministic variable in our model. Moreover, we assume that \( h_{i-1,i} \) is the binding constraint among all factors that may affect the minimum required separation between consecutive aircraft.

In the absence of queue, the actual time that airplane \( i \) would arrive at the fix, \( A_i \), consists of a deterministic and a stochastic portion, \( A_i = a_i + \tilde{A}_i \). The deterministic component \( a_i \) is its assigned arrival time at the fix, while the stochastic component \( \tilde{A}_i \) represents the lateness (positive or negative) with which the aircraft arrives at the fix, due to imprecision in trajectory adherence. It should be emphasized, though, that \( \tilde{A}_i \)'s do not represent factors such as departure delays, traffic management initiatives, severe weather, or en-route congestion, that cause significant amounts of delays; we assume that such factors have already been incorporated when calculating assigned arrival times \( a_i \).

Therefore, given a schedule of arrivals that is not subject to further major updates, our analysis aims to evaluate the performance of that system, applying tools from queueing theory. Determining how long in advance arrival times \( a_i \)'s are assigned depends on whether major operational disruptions are present. Nevertheless, in normal circumstances one can be confident that for airborne flights within a 200 nmi range from the airport their estimated time of arrival can be predicted with high accuracy. Current decision support tools for arrival traffic management, such as NASA’s TMA (Landry et al., 2003), assume that aircraft within a range of 150 nmi can meet their assigned times of arrival at the entry gates of terminal airspace with an error exhibiting 15 seconds of standard deviation (Green & Vivona, 1996). Actual flight trials in Stockholm’s terminal airspace indicated a standard deviation of 6 seconds in meeting controlled times of arrival at the initial approach fix and of 17 seconds for meeting runway threshold crossing times (Klooster, del Amo, & Manzi, 2009). In all cases the magnitudes of the errors appear well under existing airborne separation requirements, which are on the order of minutes. Furthermore, improved wind prediction technology in NextGen is expected to increase 4DT adherence, and thus decrease \( \tilde{A}_i \) deviations.

In such an operational context, where aircraft can meet their scheduled times of arrival with errors that are small compared to their separation requirements, serving flights on a First-Scheduled-First-Served (FSFS) order will not significantly increase delays. Under a FSFS queue discipline, the actual time airplane \( i \) departs from the fix, \( D_i \), would be \( A_i \) if the minimum headway is not binding, or the time the previous scheduled aircraft \( i - 1 \)
crossed the fix plus a minimum required separation headway $h_{i-1,i}$ between the two aircraft. The actual times that aircraft cross the fix under study would then be:

$$D_i = A_i$$
$$D_i = \max\left(A_i, D_{i-1} + h_{i-1,i}\right), \quad \forall i \geq 2$$

If there were no stochasticity in the system, the deterministic time of departure from the fix would be:

$$d_i = \max\left(a_i, d_{i-1} + h_{i-1,i}\right), \quad \forall i \geq 2$$

Accounting for stochasticity, the departure time from the fix of airplane $i$ is:

$$D_i = d_i + \tilde{D}_i$$

The distribution of the stochastic component $\tilde{D}_i$ clearly depends on $\tilde{A}_i$, which captures all stochastic effects that cause flight $i$ to arrive at a time other than its scheduled one $a_i$:

$$\tilde{D}_i = \tilde{A}_i \quad (2.1a)$$
$$\tilde{D}_i = \max\left(a_i + \tilde{A}_i, d_{i-1} + \tilde{D}_{i-1} + h_{i-1,i} - d_i\right), \quad \forall i \geq 2 \quad (2.1b)$$

We assume that the vector of stochastic errors $\tilde{A}_i$ follows a multivariate normal distribution with zero means (without loss of generality), standard deviations $\sigma_i$, and a covariance structure $\Sigma$: $\tilde{A} \sim \text{Normal}(0, \Sigma)$. The normality assumption stems from the observation that the probability distribution for $\tilde{A}_i$ is generated by convolving the individual distributions of low-correlated stochastic factors. It should be re-emphasized though, that $\tilde{A}_i$'s do not represent factors such as departure delays, traffic management initiatives, severe weather, or en-route congestion, that cause significant amounts of delays; we assume that such factors have already been incorporated in the determination of scheduled arrival times $a_i$. Evidence from detailed Monte Carlo simulations shows that trajectory control-time errors follow a bell-shaped distribution that can be approximated with a normal one (Mueller et al, 2004). Moreover, the same study reports a mean error of almost zero and a standard deviation of 24 seconds.

In practice, values for schedule deviation $\sigma_i$ could be grouped to represent classes of aircraft that have similar capabilities of adherence to 4D trajectories. For example, one
could assume two different values for the standard deviation, \( \sigma_A \) and \( \sigma_B \), in order to roughly represent aircraft with and without Required Navigation Performance (RNP) capabilities.

### 2.1.2 Solution with the Clark approximation method

The \textit{max} operation on normal random variables, in contrast to the \textit{add} operation, does not yield a normal random variable. A well-known result due to Clark (1961) derives analytical formulas for the mean and variance of the maximum of two normally distributed random variables. Let \( X \) and \( Y \) be normally distributed random variables, \( X \sim N(\mu_X, \sigma_X) \) and \( Y \sim N(\mu_Y, \sigma_Y) \), \( \rho \) represent the correlation coefficient between \( X \) and \( Y \), and \( Z \) be the maximum of \( X \) and \( Y \), \( Z \triangleq \max(X,Y) \). The mean \( \mu_Z \) and variance \( \sigma_Z^2 \) of \( Z \) are then:

\[
\mu_Z = \mu_X \Phi(\alpha) + \mu_Y \Phi(-\alpha) + \gamma \phi(\alpha) \tag{2.2}
\]

\[
\sigma_Z^2 = (\sigma_X^2 + \mu_X^2)\Phi(\alpha) + (\sigma_Y^2 + \mu_Y^2)\Phi(-\alpha) + (\mu_X + \mu_Y)\gamma \phi(\alpha) - \mu_Z^2 \tag{2.3}
\]

where

\[
\gamma \triangleq \left( \sigma_X^2 + \sigma_Y^2 - 2\rho \sigma_X \sigma_Y \right)^{1/2}
\]

\[
\alpha \triangleq \frac{\mu_X - \mu_Y}{\gamma}
\]

\[
\phi(x) \triangleq \frac{1}{\sqrt{2\pi}} \exp\left( -\frac{x^2}{2} \right)
\]

\[
\Phi(y) \triangleq \int_{-\infty}^{y} \phi(x) dx
\]

The coefficient of linear correlation between \( Z \) and a third normal random variable \( W \) can also be calculated, given that we know the coefficients of linear correlation between \( X \) and \( W \left( r_{x,W} \right) \), and between \( Y \) and \( W \left( r_{y,w} \right) \):

\[
r[W,Z] = \left( \sigma_X r_{x,W} \Phi(\alpha) + \sigma_Y r_{y,W} \Phi(-\alpha) \right) / \sigma_Z \tag{2.4}
\]

The above formulas give the exact mean and variance of \( Z \). The approximation is introduced by assuming that \( Z \) follows a normal distribution with mean \( \mu_Z \) and variance \( \sigma_Z^2 \).
In the context of our problem with scheduled aircraft arrivals, the Clark approximation method can be used for all \( i \geq 2 \) if we approximate \( \tilde{D}_t \)'s as normal random variables, and estimate their mean \( E(\tilde{D}_t) \) and variance \( Var(\tilde{D}_t) \) in a recursive manner:

\[
E(\tilde{D}_t) = a_i \Phi(\alpha_i) + \left[ d_{i-1} + E(\tilde{D}_{i-1}) + h_{i-1,i} \right] \Phi(-\alpha_i) + \gamma_i \phi(\alpha_i) - d_i
\]  \hspace{1cm} (2.5)

\[
Var(\tilde{D}_t) = \left( \sigma_i^2 + a_i^2 \right) \Phi(\alpha_i) + \left[ \Var(\tilde{D}_{i-1}) + \left[ d_{i-1} + E(\tilde{D}_{i-1}) + h_{i-1,i} \right]^2 \right] \Phi(-\alpha_i) + \left[ a_i + d_{i-1} + E(\tilde{D}_{i-1}) + h_{i-1,i} \right] \gamma_i \phi(\alpha_i) - \left[ E(\tilde{D}_i) \right]^2
\]  \hspace{1cm} (2.6)

\[
r[\tilde{\tilde{A}}_{i+1}, \tilde{D}_i] = \left[ \sigma_i \cdot \rho_i \Phi(\alpha_i) + \sqrt{\Var(\tilde{D}_{i-1})} \cdot \rho_2 \cdot \Phi(-\alpha_i) \right] / \sqrt{\Var(\tilde{D}_i)}
\]  \hspace{1cm} (2.7)

where

\[
\gamma_i \equiv \left( \sigma_i^2 + \Var(\tilde{D}_{i-1}) - 2 \cdot \rho \cdot \sigma_i \cdot \sqrt{\Var(\tilde{D}_{i-1})} \right)^{1/2}
\]

\[
\alpha_i = \frac{a_i - d_{i-1} - E(\tilde{D}_{i-1}) - h_{i-1,i}}{\gamma_i}
\]

and at each step \( i \) \( \rho = r[\tilde{A}_i, \tilde{D}_{i-1}] \), \( \rho_i = r[\tilde{\tilde{A}}_{i+1}, \tilde{A}_i] \), \( \rho_2 = r[\tilde{\tilde{A}}_{i+1}, \tilde{D}_{i-1}] \).

In the above equations, \( r[\tilde{A}_i, \tilde{D}_{i-1}] \) and \( r[\tilde{\tilde{A}}_{i+1}, \tilde{D}_{i-1}] \) are obtained through equation (2.7) in previous iterations. Effectively, the method is implemented by estimating at each step \( k \) \( r[\tilde{A}_i, \tilde{D}_k] \) for all \( i > k \). We obtain \( r[\tilde{\tilde{A}}_{i+1}, \tilde{A}_i] \) from covariance matrix \( \Sigma \). Equations (2.5)-(2.7) are easy to program and they are computationally efficient. Finally, for a stream of \( N \) flights scheduled to arrive at a fix, the total expected queueing delay is defined as:

\[
E[W_N] = E \left[ \sum_{i=1}^{N} D_i - A_i \right] = \sum_{i=1}^{N} E[D_i] - a_i
\]  \hspace{1cm} (2.8)

This completes the formulation of our queueing model. In summary, the model requires as inputs a schedule of arrival times \( a_i \), a vector of time separation headways \( h_{i-1,i} \), and a
covariance matrix of the standard deviations $\sigma_{i,j}$ of trajectory adherence errors. These, coupled with the assumption that $\sigma_i$'s are small enough to allow a first-scheduled, first-served policy, enable the estimation of expected flight delays through Clark's approximation method.

2.1.3 Approximation error

Although the maximum $Z$ of two normal random variables $X$ and $Y$ is not normally distributed, our model is based on approximating $Z$ with a normal random variable. In particular, in estimating $D_i = \max(A_i, D_{i-1} + h_{i-1,j})$ it is assumed that $D_{i-1}$ is normally distributed. That enables the estimation of the mean and variance of $D_i$, which is then also approximated as a normal random variable. However, each pair-wise operation introduces some error that is propagated and might affect the accuracy of our estimates. Horowitz et al. (1982) investigate the accuracy of the Clark approximation method for econometric estimation of a wide range of probit models for discrete choice probabilities, and find that the accuracy of the method varies greatly from model to model.

To test the accuracy of the Clark Approximation Method in the context of our analysis, several operational scenarios were considered. The estimates from the analytical queueing model were then compared against the average of $10^4$ Monte Carlo simulation runs, regarded as ground truth.

Each operational scenario was formulated as follows: a set of 120 aircraft must cross a fix, and the minimum required separation $h_{i-1,j}$ between any two successive aircraft takes values of 30, 60, or 90 seconds. To emulate real-world operations, a buffer time $b$ is added to $h_i$ when scheduling aircraft arrivals at the fix: $a_i = a_{i-1} + h_{i-1,j} + b$. Aircraft meet their scheduled times of arrival $a_i$ with some imprecision error that follows a normal distribution, has variance $\sigma^2$, and is independent across aircraft. A total of 90 scenarios were examined:

- 10 different sequences of $h_{i-1,j}$, where each sequence is determined randomly but given an even mix of 30, 60, and 90 second headway values
- $b = 0, 10, \text{ and } 20 \text{ seconds (held constant within each sequence)}$
- $\sigma = 10 \text{ seconds (uniform across all aircraft), } 30 \text{ seconds (uniform across all aircraft), and an even mix of both (with the order determined randomly)}$

For each scenario, we computed $E[D_i]_{\text{appr}}$ and $E[W_N]_{\text{appr}}$ using equations (2.5)-(2.8), and also $\bar{D}_i^{\text{sim}}$ and $\bar{W}_N^{\text{sim}}$, which denote the average estimates obtained through $10^4$ simulation runs. Three metrics for the approximation method accuracy were considered:
• Percentage Error in total Delay % (PE): \( \frac{E[W_N^{\text{appr}}] - \bar{W}_N^{\text{sim}}}{\bar{W}_N^{\text{sim}}} \cdot 100 \)

• Absolute Error in Total Delay (AE): \( |E[W_N^{\text{appr}}] - \bar{W}_N^{\text{sim}}| \)

• Flight Departure Time Mean Absolute Deviation (MAD): \( \frac{1}{N} \sum_{i=1}^{N} |E[D_i^{\text{appr}}] - \bar{D}_i^{\text{sim}}| \)

The first two metrics evaluate the accuracy of the approximation method in estimating the expected total queueing delay. The third metric provides with a measure of the error in predicted outcomes for individual flights.

The results are summarized in Table 2.1. Each entry in the table represents the average value for each metric across the ten scenarios of different \( h_{i+1,i} \) sequences, while the standard deviation is enclosed in parenthesis. In all cases, the Total Delay PE metric indicates that the Clark approximation method underestimates total delay in the system, but by no more than 8% as compared to simulation. Moreover, the absolute error in estimating total delay never exceeds 1.6 minutes. Also, the MAD metric indicates that the approximation method estimates the expected time of departure from the fix for each aircraft with accuracy better than 1 second, on average. The accuracy of the method slightly decreases when the fleet contains aircraft with different navigation capabilities. This must be due to heterogeneity in the variance of the normal distributions for \( A_i \) that enters in the \( \max \) operator in each step of the recursion.

**Table 2.1.** Results of accuracy tests of the Clark approximation method

<table>
<thead>
<tr>
<th></th>
<th>Buffer = 0 seconds</th>
<th></th>
<th>Buffer = 10 seconds</th>
<th></th>
<th>Buffer = 20 seconds</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PE (sec)</td>
<td>AE (sec)</td>
<td>MAD (sec)</td>
<td>PE (sec)</td>
<td>AE (sec)</td>
<td>MAD (sec)</td>
</tr>
<tr>
<td>( \sigma = 10 ) (sec)</td>
<td>-0.62% (0.17%)</td>
<td>13.78 (3.24)</td>
<td>0.14 (0.03)</td>
<td>-3.26% (0.35%)</td>
<td>9.17 (1.18)</td>
<td>0.09 (0.01)</td>
</tr>
<tr>
<td>( \sigma = 30 ) (sec)</td>
<td>-0.49% (0.11%)</td>
<td>36.50 (13.44)</td>
<td>0.35 (0.05)</td>
<td>-1.69% (0.17%)</td>
<td>40.92 (3.07)</td>
<td>0.35 (0.03)</td>
</tr>
<tr>
<td>Mixed</td>
<td>-1.52% (0.28%)</td>
<td>97.26 (14.89)</td>
<td>0.89 (0.14)</td>
<td>-5.74% (0.32%)</td>
<td>79.53 (4.33)</td>
<td>0.65 (0.04)</td>
</tr>
</tbody>
</table>

-5.74\% (0.32\%) | 79.53 (4.33) | 0.65 (0.04) | -7.70\% (0.21\%) | 54.07 (3.62) | 0.44 (0.01) |
In conclusion, these experimental results indicate that our proposed model accurately predicts operational consequences of metered operations with good but imperfect 4DT adherence, as might be expected in NextGen. It should be noted that the method has not been tested in the presence of correlation. However, positive correlation seems to improve the accuracy of the approximation, as discussed in (Daganzo, Bouthelier, & Sheffi, 1977).

2.1.4 Comparison with deterministic queueing model

Next, we compare the total expected delay computed through the stochastic model presented above with that obtained through a deterministic queueing model. The latter is a special case of the stochastic model, with aircraft capable of adhering perfectly to their assigned times of arrival at a fix. Therefore, in the deterministic model we have that $\hat{A}_i \equiv 0$ for every aircraft $i$.

The two models are compared under a wide range of realistic demand and capacity scenarios for seven major US airports: ATL, BOS, ORD, DFW, MIA, LGA, SFO. The model input data are those used by Hansen et al (2009) to compare a deterministic queueing model with the $M(t)/E_k(t)/1$ airport queueing model proposed by Kivestu (1976). The reader is referenced to Hansen et al (2009) for a detailed description of the input data. In brief, eight arrival flight schedules were considered for each airport, four peak and four off-peak days in 2007, obtained through runs from the Airspace Concept Evaluation System (Windhorst et al, 2006). The minimum headway requirements between successive flights are not aircraft-type specific, but estimated through airport acceptance rate (AAR) values, available in the Airspace System Performance Metric (ASPM) database. Clustering methodologies have been employed to extract, for different airports, representative capacity scenarios from these data (Liu, Mukherjee, and Hansen, 2006). For the current study, five capacity scenarios were employed for ATL, ORD, DFW, and LGA, while four were considered for MIA, and six for BOS and SFO. To translate capacity scenarios into minimum headway requirements, the day is divided into discrete time periods of equal duration $\tau$ and indexed with $t = 1, 2, \ldots, T$. For each time period $t$ a capacity, $C(t)$, defined as the maximum number of flights that can land in that period, is provided from the capacity scenarios. With the exception of LGA where $\tau$ is available on an hourly time scale, values for $C(t)$ every 15 minutes were available for all other airports. The minimum headway requirement between successive flights for every time period $t$ is then calculated as $h(t) = \tau / C(t)$.

A total of 288 scenarios were analyzed, each corresponding to a combination of a capacity profile at a given airport and an arrival flights schedule for an entire day at that airport. For each scenario $k$, we computed the average queueing delay across all $N_k$ scheduled arrivals through the deterministic model, $Q_k^D = \frac{1}{N_k} \sum_{i=1}^{N_k} (d_i - a_i)$, and the expected queueing delay averaged over all flights through the stochastic model,
To compute $E[D_i]'s$ we employed the Clark approximation method, and assumed that trajectory adherence errors $\tilde{A}_i$'s are i.i.d normal random variables with standard deviation $\sigma$ for all aircraft. Six cases for $\sigma$ were considered, in particular $\sigma \in \{10, 20, 30, 40, 50, 60\}$ seconds.

To compare $Q_k^S$ and $Q_k^D$, six linear regression models of the form $Q_k^S = \beta_0 + \beta_1 \cdot Q_k^D$ were estimated across all 288 data points $(Q_k^D, Q_k^S)$, one for each of the six values for $\sigma$. Table 2.2 summarizes the estimates from the six linear regression models.

**Table 2.2.** Comparison of deterministic and stochastic models when $\sigma \in \{10, 20, 30, 40, 50, 60\}$ seconds.

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>$\beta_0$ (minutes/flight)</th>
<th>$\beta_1$ (unit-less)</th>
<th>$R^2_{adj}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.023*</td>
<td>1.000*</td>
<td>1.00</td>
</tr>
<tr>
<td>20</td>
<td>0.086*</td>
<td>1.000*</td>
<td>1.00</td>
</tr>
<tr>
<td>30</td>
<td>0.181*</td>
<td>1.000*</td>
<td>1.00</td>
</tr>
<tr>
<td>40</td>
<td>0.300*</td>
<td>1.000*</td>
<td>1.00</td>
</tr>
<tr>
<td>50</td>
<td>0.437*</td>
<td>1.000*</td>
<td>1.00</td>
</tr>
<tr>
<td>60</td>
<td>0.587*</td>
<td>1.000*</td>
<td>1.00</td>
</tr>
</tbody>
</table>

* Significant at a 0.1% level

In all six cases, linear correlation between the estimates of the two models is extremely strong. Further, the intercept $\beta_0$ ranges between 1.4 seconds and 35 seconds of additional delay per flight, while the slope coefficient $\beta_1$ is almost equal to one. Therefore, both models yield very similar estimates of expected total delay, but for an additive constant. That suggests that delays are dominated by surges in arrivals, during which demand rates exceed capacity, whereas delays due to precision error are of secondary importance. This is in accordance with results in Hansen et al (2009), who find that estimates of expected queueing delay computed through a deterministic and an $M(t)/E_k(t)/1$ model have small differences. Since the stochastic model proposed in this paper represents intermediate levels of stochasticity between the above-mentioned models, one would expect its estimates to have even smaller differences from the deterministic than the $M(t)/E_k(t)/1$ model does.
The above findings, however, are no longer true if we restrict our analysis to scenarios of low average delay, where $Q_k^D \leq 1$ minutes/flight. These amount to 113 out of 288, or 39%, of all scenarios initially considered. Performing linear regression only on those scenarios yields the results displayed in Table 2.3. Estimates for $\beta_0$'s slightly increase, as compared to those in Table 2.2, but not more than two seconds, whereas the slope coefficient $\beta_1$ appears smaller than one and decreases as imprecision increases. Therefore, for low levels of utilization of the airport, characterized by average deterministic queueing delay smaller than one minute per flight, a discrepancy between the estimates of the deterministic and stochastic model is evident, and the impact of flight imprecision can no longer be an additive correction to a deterministic queueing model.

<table>
<thead>
<tr>
<th>$\sigma$ (seconds)</th>
<th>$\beta_0$ (minutes/flight)</th>
<th>$\beta_1$ (unit-less)</th>
<th>$R^2_{\text{adj}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.029*</td>
<td>0.980*</td>
<td>1.00</td>
</tr>
<tr>
<td>20</td>
<td>0.105*</td>
<td>0.932*</td>
<td>0.98</td>
</tr>
<tr>
<td>30</td>
<td>0.209*</td>
<td>0.881*</td>
<td>0.93</td>
</tr>
<tr>
<td>40</td>
<td>0.332*</td>
<td>0.832*</td>
<td>0.82</td>
</tr>
<tr>
<td>50</td>
<td>0.470*</td>
<td>0.783*</td>
<td>0.68</td>
</tr>
<tr>
<td>60</td>
<td>0.621*</td>
<td>0.735*</td>
<td>0.52</td>
</tr>
</tbody>
</table>

* Significant at a 0.1% level

Considering all 288 scenarios again, the relationship between standard deviation $\sigma$ and intercept $\beta_0$ can be represented by a polynomial of second degree, specifically $\beta_0 = 0.00785\sigma^2 + 0.1239\sigma$, where both $\sigma$ and $\beta_0$ take units of seconds.

In summary, the estimates of expected queueing delay from the stochastic and the deterministic model are strongly correlated and very similar, except for cases of low airport utilization, characterized by average deterministic queueing delay over all arrivals smaller than one minute.

### 2.1.5 System with two servers

In this section we present a formulation to model the progression of aircraft through a series of servers. That is often the case with operations in the terminal airspace area of large metropolitan airports, where aircraft are metered at entry fixes and must precisely fly an assigned trajectory throughout their descent to the runway. Such procedures are currently in place in PHL (Landry et al., 2003) and DFW (Swenson et al., 1997), and are
expected to predominate in NextGen under super-density arrival/departure operations (JPDO, 2010). We seek to estimate aircraft’s expected times of departure from each fix, given scheduled times of arrival at each fix as input.

The following analysis assumes that aircraft are assigned scheduled times of arrival at each fix, and that they cross each fix in the order specified by the schedule. Therefore, it suffices to consider only two fixes, as the extension to three or more is straightforward.

Let $D_{i,1}$ denote the time moment aircraft $i$ departs from upstream Fix 1, $D_{i,2}$ the moment when the same aircraft departs from downstream Fix 2, and $F$ the set of flights that traverse both fixes. Also, let $T_i$ be the unimpeded (from queueing effects) travel time of aircraft $i$ between the two fixes, and assume it is independent of the time it departs from fix 1, $D_{i,1}$. Consistent with our previous analysis, we assume that $T_i$ is normally distributed around $t_i$ with covariance structure $\Sigma$: $T \sim \text{Normal}(t, \Sigma)$. The departure time of aircraft $i$ from downstream Fix 2 can be expressed as:

$$D_{i,2} = D_{i,1} + T$$

(2.9a)

$$D_{i,2} = \max\{D_{i,1} + T_i, D_{i-1,2} + h_{i-1,i}\}, \quad \forall i \geq 2$$

(2.9b)

Our goal is to estimate $E[D_{i,2}]$ by employing (2.5)–(2.7). The main difficulty arises in (2.7), estimating the coefficient of linear correlation $r[D_{i,1} + T_i, D_{i-1,2}]$. That is addressed through a series of steps, described in the following algorithm:

**Step 0:** Estimate $E[D_{i,1}]$ for all aircraft departing from Fix 1 through (2.5)–(2.7). For each aircraft’s departure time $D_{i,1}$ estimate from equation 2.4 its coefficient of linear correlation with all preceding aircraft $k < i$:

$$r[D_{i,j}, D_{k,j}] = \frac{\sqrt{\text{Var}(D_{i,j})} \cdot r[D_{i,j}, D_{k,j}] \cdot (1 - \Phi(a_i))}{\sqrt{\text{Var}(D_{i,j})}}$$

**Step 1:** For the first aircraft departing from Fix 2, estimate $r[D_{i,1} + T_i, D_{i,2}]$. With use of (2.9a), that can be written as:

$$r[D_{i,1} + T_i, D_{i,2}] = \frac{\text{Cov}(D_{i,1}, D_{i,1}) + \text{Cov}(T_i, T_i)}{\sqrt{\text{Var}(D_{i,1}) + \text{Var}(T_i)} \cdot \sqrt{\text{Var}(D_{i,1}) + \text{Var}(T_i)}}$$

20
Step k: For all \( i \in F \) and \( i \geq k \), compute

\[
 r\left[ D_{i,1} + T_i, D_{k,2} \right] = \left[ \sqrt{\text{Var}(D_{k,1} + T_k)} \right] \cdot \rho_1 \cdot \Phi(a_k) + \\
+ \left[ \sqrt{\text{Var}(D_{k-1,2})} \cdot \rho_2 \cdot \Phi(-a_k) \right] / \sqrt{\text{Var}(D_{k,2})}
\]

where \( \rho_1 = r\left[ D_{i,1} + T_i, D_{k,1} + T_k \right] \) and \( \rho_2 = r\left[ D_{i,1} + T_i, D_{k-1,2} \right] \) are updated in every step \( k \).

To estimate \( \rho_1 \), first it can be easily shown that for any pair \((i, k)\):

\[ \text{Cov}\left[ D_{i,1} + T_i, D_{k,1} + T_k \right] = \text{Cov}\left[ D_{i,1}, D_{k,1} \right] + \text{Cov}[T_i, T_k] \]

Now, \( \text{Cov}\left[ D_{i,1}, D_{k,1} \right] \) is computed in Step 0, while \( \text{Cov}[T_i, T_k] \) is given as input in \( \Sigma \). Finally, \( \rho_2 \) is updated in step \( k - 1 \).

The reader will recognize that we have outlined a computational procedure for providing estimates of mean departure times from the downstream Fix 2. That can be used to estimate queueing delays for a set of flights that must traverse a series of fixes, subject to the limitation no merging of arrival streams is considered.

### 2.2 Special case with constant buffer and uniform precision error

We examine the case where average demand rate exceeds capacity over a period of time, after which it drops to zero. That can represent banked arrivals at busy hub airports, where surges of arrivals during rush hours are followed by time intervals with low number of scheduled arrivals. Figure 2.1 illustrates this pattern, through plots of cumulative number of scheduled arrivals on four typical weekdays in February 2010 at four busy US airports. At CLT, SLC, and IAH the pattern is recurrent throughout the day, whereas in ATL, the busiest airport in the US, banked arrivals are observed during the morning peak period. We, thus, focus to examine situations where demand is depicted by the simplified curve of Figure 2.2. Although not representative of cases where demand rate drops after the initial surge but not to a zero level, we consider the demand scenario in Figure 2.2 as a realistic case, which also has the advantage of reducing the number of parameters in our analysis.
Figure 2.1a. Cumulative number of scheduled arrivals at CLT on 2/21/2010 (source: www.bts.gov)

Figure 2.1b. Cumulative number of scheduled arrivals at SLC on 2/16/2010 (source: www.bts.gov)
Figure 2.1c. Cumulative number of scheduled arrivals at IAH on 2/12/2010 (source: www.bts.gov)

Figure 2.1d. Cumulative number of scheduled arrivals at ATL on 2/2/2010 (source: www.bts.gov)
To maximize throughput and minimize expected delay, \( a_i \) should be set equal to \( a_{i-1} + h_{i-1} \). It may, however, be better to set \( a_i \) slightly greater by inserting some buffer time \( b \), for at least two reasons. First, while inserting a buffer increases delay, it also makes the delay more predictable. Air traffic controllers value such predictability because it enables them to better plan operations. Second, the delays that result from excess separation \( b \) can be absorbed in a more efficient manner, by increasing time at higher altitudes and reducing speed slightly over a longer distance rather than dramatically as the aircraft approaches the metering fix. For these reasons, it may be desirable to accept an increase in overall delay in exchange for a reduction in the amount of that portion of overall delay that results from imperfect trajectory adherence.

In this section we quantify this trade-off under the following conditions: 1) The inserted buffer \( b \) between successive aircraft is constant and deterministic, 2) Stochastic errors \( \tilde{A}_i \) are i.i.d random variables with zero mean and standard deviation \( \sigma \). Having assumed that demand for traversing the fix exceeds capacity, aircraft can be scheduled to arrive at the fix at times \( a_i = a_{i-1} + h_{i-1} + b \). In this way we can distinguish two types of delay: the deterministic loss in efficiency from inserting a buffer \( b \) and the stochastic queueing delay due to imperfect adherence in scheduled times of arrival \( a_i \). The deterministic delay for aircraft \( i \) is simply \( i \cdot b \), while the stochastic delay is defined as the difference between aircraft’s times of arrival and departure from the fix, \( D_i - A_i \). A graphical illustration for a constant minimum headway \( h \), and therefore constant headway \( a \) between scheduled arrivals, is provided in the queueing diagram of Figure 2.2.

---

**Figure 2.2.** Queueing diagram with deterministic and stochastic delays
2.2.1 Estimation of stochastic delay

Since computation of deterministic delay is straightforward, we focus on estimating the expected value of stochastic delay. Following the notation of Section 2.1, the deterministic component of departure times from the fix becomes $d_i = a_i$, and thus Eq. (2.1) can be written as:

$$
\bar{D}_1 = \bar{A}_1
$$

(2.10a)

$$
\bar{D}_i = \max(\bar{A}_i, \bar{D}_{i-1} - b), \ \forall i \geq 2
$$

(2.10b)

where $\bar{A}_i \sim N(0, \sigma), \ \forall i \geq 1$. The queueing delay for each aircraft, therefore, becomes $\bar{D}_i$. We are interested in deriving an analytical expression for $E(\bar{D}_i)$, but this is a rather difficult task. Instead we define $\bar{Z}_i = \max(\bar{A}_i^*, \bar{Z}_{i-1} - \Delta)$, where $\bar{A}_i^* \sim N(0,1)$, and resort to dimensionless analysis by employing the following proposition:

**Proposition 2.1.** For a given level of constant $\Delta \triangleq b/\sigma$, the sequence of random variables $\bar{D}_i$ for $i > 1$ is the sequence of random variables $\bar{Z}_i$ multiplied by the standard deviation $\sigma$:

$$
\bar{D}_i = \sigma \cdot \bar{Z}_i
$$

(2.11)

**Proof.** Assume that for some $i \geq 2$, $\bar{D}_{i-1} = \sigma \cdot \bar{Z}_{i-1}$. Then we have:

$$
\bar{D}_i = \max(\bar{A}_i, \bar{D}_{i-1} - b)
$$

$$
= \max(\sigma \cdot \bar{A}_i^*, \sigma \cdot \bar{Z}_{i-1} - \sigma \cdot \Delta) = \sigma \cdot \bar{Z}_i
$$

But note that for $i = 2$:

$$
\bar{D}_2 = \max(\bar{A}_2, \bar{D}_1 - b)
$$

$$
= \max(\sigma \cdot \bar{A}_2^*, \sigma \cdot \bar{A}_1^* - \sigma \cdot \Delta) = \sigma \cdot \bar{Z}_2
$$

Hence, for $i \geq 3$ the result follows by induction. \[\therefore\]
Therefore, the problem of estimating \( E(D_i) \) is reduced to estimating \( E(\tilde{Z}_i) \). To estimate the mean of \( \tilde{Z}_i \) we first derive its cumulative distribution function:

\[
F_{\tilde{Z}_i}(x) = \Phi(x) \cdot \Phi(x + \Delta) \cdot \ldots \cdot \Phi(x + (i - 1)\Delta)
\]

The probability density function of \( \tilde{Z}_i \) is therefore:

\[
f_{\tilde{Z}_i}(x) = \sum_{m=1}^{i} \prod_{k=1}^{m} \Phi(x + (k - 1)\Delta)^{(\omega_k)} , \quad \omega_k = 1 \text{ for } m = k
\]
\[
\omega_k = 0 \text{ for } m \neq k
\]

\[
E[\tilde{Z}_i] = \int_{-\infty}^{\infty} x \cdot \sum_{m=1}^{i} \prod_{k=1}^{m} \Phi(x + (k - 1)\Delta)^{(\omega_k)} \, dx , \quad \omega_k = 1 \text{ for } m = k
\]
\[
\omega_k = 0 \text{ for } m \neq k \quad (2.12)
\]

Figure 2.3. Expected delay \( E(\tilde{Z}_i) \) for several values of \( \Delta \)

where the term \( (\omega_k) \) indicates a derivative. The mean of \( \tilde{Z}_i \) is therefore:
Therefore, it suffices to consider cases for $\Delta$ and compute $E(\tilde{Z}_i)$ through numerical integration; Figure 2.3 depicts several such curves for $E(\tilde{Z}_i)$. Note that $\tilde{Z}_i$'s eventually reach a steady state (except for the case $\Delta = 0$), after which each flight incurs the same amount of expected stochastic delay $E(\tilde{Z}_i) \cdot \sigma$. It should be noted that the results are exact, obtained through numerical integration and not by employing the Clark approximation method.

Finally, for $N$ aircraft in the arrival surge, the total expected stochastic delay is proportional to the standard deviation $\sigma$ of the stochastic error $\tilde{A}_i$:

$$E(W_N) = \sum_{i=1}^{N} E(\tilde{D}_i) = \sigma \cdot \sum_{i=1}^{N} E(\tilde{Z}_i)$$

We can therefore plot the normalized (i.e. with setting $\sigma = 1$) expected total delay $E(W_N^*)$ against the number of flights $N$ in the arrival stream. In Figure 2.4, values of $E(W_N^*)$ as a function of $N$ and for several cases of $\Delta$ are displayed. Thus, estimates of the expected total delay for a surge of $N$ aircraft can be easily obtained by selecting the appropriate curve from Fig. 2.4 and multiplying $E(W_N^*)$ by $\sigma$. Observe that for lower levels of the relative buffer $\Delta$, the system incurs higher expected delay due to its reduced capacity for absorbing stochastic deviations from schedule.

Figure 2.4. Normalized Expected Total Delay $E(W_N^*)$ for several values of $\Delta$
2.2.2 Solution with the Clark approximation method

In this section we are interested in the accuracy of Clark’s approximation method in the computation of $E(\tilde{Z}_i)$. While not exact, estimates of $E(\tilde{D}_i)$ when obtained through Clark’s method require significantly less computational time (1 second instead of 60 seconds, for problem size $N = 100$). The mean of $\tilde{D}_i$ estimated through Clark’s method is given by the following formulas:

$$E(\tilde{D}_i) = [E(\tilde{D}_{i-1}) - b] \Phi(-\alpha_i) + \gamma_i \phi(\alpha_i)$$

(2.14)

$$\text{Var}(\tilde{D}_i) = \sigma^2 \Phi(\alpha_i) + \left[ \text{Var}(\tilde{D}_{i-1}) + \left[ E(\tilde{D}_{i-1}) - b \right]^2 \right] \Phi(-\alpha_i)$$

$$+ \left[ E(\tilde{D}_{i-1}) - b \right] \gamma_i \phi(\alpha_i) - \left( E(\tilde{D}_i) \right)^2$$

(2.15)

where $\alpha_i = \left( -E(\tilde{D}_{i-1}) - b \right) / \gamma_i$ and $\gamma_i = \left( \sigma^2 + \text{Var}(\tilde{D}_{i-1}) \right)^{1/2}$.

We compare the values for $E(\tilde{D}_i)$ computed numerically and depicted in Figure 2.3, with those obtained using equations (2.14)-(2.15) above. In light of Proposition 2.1, it is sufficient to analyze the case of $\sigma = 1$. As a metric of comparison, the percent difference between approximate and numerical values is calculated:

$$\varepsilon(\%) = \frac{E(\tilde{Z}_i)^{\text{appr}} - E(\tilde{Z}_i)^{\text{num}}}{E(\tilde{Z}_i)^{\text{num}}} \cdot 100$$

Nine scenarios for $\Delta$ were considered, $\Delta \in \{0, 0.2, \ldots, 1.6\}$. The intensity of errors $\varepsilon$ is depicted in the contour plot of Figure 2.5. The estimates based on the Clark approximation are biased downward, while the magnitude of the errors increases, in percentage terms, as $\Delta$ increases.
2.3 Trade-offs between deterministic and stochastic delay

After developing methods for estimating each flight’s expected stochastic delay, we now analyze the trade-off between stochastic delay and throughput. In the context of our analysis, the maximum departure rate from the fix that can be attained is constrained by aircraft minimum separation requirements. Thus, maximum throughput is achieved by metering aircraft at the fix without any buffer times. However, that will result in maximum stochastic delay, as can be inferred from Figure 2.4 for the curve $\Delta = 0$. By inserting some buffer we decrease the expected amount of stochastic delay, since $\Delta$ increases with $b$, but at the same time we incur increases in deterministic queueing delay.

For a surge of $N$ aircraft arrivals, the expected loss from those two types of delay can be expressed as:

$$E[L] = \sum_{i=1}^{N} b \cdot (i - 1) + \beta \cdot \sum_{i=1}^{N} E[\tilde{Z}_i] \cdot \sigma$$
or

\[ E[L] = \left( \frac{1}{2} \cdot (N - 1) \cdot N \cdot \Delta + \beta \cdot \sum_{i=1}^{N} E[\tilde{Z}_i] \right) \cdot \sigma \]  \hspace{1cm} (2.16)

The coefficient \( \beta \) in the above relationship is the relative cost of stochastic delay over the delay due to reduced throughput. As noted earlier, delays due to reduced metering rate can be planned for well in advance, taken at a higher altitude, in less busy airspace, and with relatively small speed reductions. For illustration, we consider three values for the relative cost of stochastic over deterministic delay, \( \beta \in \{1, 2, 3\} \), as well three numbers of aircraft in the arrival surge, \( N \in \{20, 40, 60\} \). For each case, the normalized (after setting \( \sigma = 1 \)) expected loss \( E[L^*] \) as a function of \( \Delta \) is plotted in Figures 2.6-2.8.

![Figure 2.6](image-url)

**Figure 2.6.** Expected loss from deterministic and stochastic delay for \( N=20 \) aircraft
The values of the relative buffer that minimize $E[L^*]$, $\Delta^*$, range between 0 and 0.08. For $\beta = 1$ or 2 the minimum expected loss is achieved when the buffer is zero. When the cost of stochastic delay is triple that of deterministic, minimum loss is achieved for a non-
zero buffer $\Delta^* > 0$. For example, a value of $\Delta^* = 0.08$ translates to a buffer of 2 seconds between consecutive aircraft for a scenario where $\sigma = 20$ seconds. Assuming a value for $h$ on the order of a minute, such a buffer would reduce planned throughput about 3% compared to the maximum possible under deterministic conditions.

Next, we are interested in finding the values of $\Delta$ that minimize $E[L^*]$ for a wider range of the model parameters $\beta$ and $N$. Figure 2.9 displays such values, $\Delta^*$, for $\beta \in \{1, 2, \ldots, 10\}$ and $N \in \{20, 40, 60, 80, 100\}$. Note that $\Delta^*$ curves appear not smooth. This is actually an artifact of the $E[L^*]$ minimization process, since $\Delta^*$’s were obtained through searching a discrete set of $\Delta$’s, ranging from 0 to 1.0 in increments of 0.01.

For a certain number of aircraft $N$, the curve of optimal buffer $\Delta^*$ increases with $\beta$. This is because, as the unit cost of stochastic delay increases, a larger buffer is required to minimize losses. On the other hand, for a given $\beta$, the optimal buffer $\Delta^*$ decreases with the number of aircraft $N$, indicating that the loss from stochastic delays increases at a lower rate than the loss from deterministic delays, as the surge of aircraft becomes larger. That is expected since deterministic delays increase with $N^2$ (see Equation 2.16), while stochastic delays increase almost linearly with $N$ as can be observed in Figure 2.4.
2.4 Evaluation of priority queue disciplines

In order for airlines to equip their fleet with avionics that enable 4DT aircraft operations, the investment in this technology should demonstrate benefits that outweigh costs. The Federal Aviation Administration is currently examining incentive mechanisms for airlines to invest in new technology, which include giving priority to aircraft equipped with NextGen avionics. The so-called Best-Equipped-Best-Served (BEBS) policy provides priority access to capacity restricted airspace to those aircraft that are capable of achieving certain performance standards, thus ensuring highest possible utilization of system capacity. Such a policy constitutes a departure from today’s practice, which serves flights on a First-Come-First-Served (FCFS) or a First-Scheduled-First-Served (FSFS) basis, the latter usually under a Ground Delay Program.

Modeling the system performance under a BEBS policy and comparing it with current practice, provides performance metrics that can be employed in an economic analysis of the BEBS policy. Moreover, it can provide useful insights when designing such a policy, especially since details on implementation rules for BEBS have not been yet announced.

Here we seek to demonstrate how the queueing model presented in the section 2.1 can be used as a computational tool for that purpose. Our approach consists of formulating simplified, but realistic, working scenarios that provide us with insights on the mechanisms that control the effectiveness of a BEBS policy.

2.4.1 Description of working scenarios

We consider the case where average demand for landings exceeds runway capacity over a considerable time period. Responding to the demand/capacity imbalance, aircraft are metered at the runway, i.e. they are assigned Scheduled Times of Arrival (STA). Two classes of aircraft are distinguished: equipped with avionics to execute 4D trajectory-based operations with high precision, and non-equipped. We proceed by making the following assumptions:

- Demand for arrivals attains a steady rate of \( \lambda = 1 \) aircraft/min for a three-hour period and then it drops to zero. Therefore, the Earliest Time of Arrival (ETA) at the runway of the \( n^{th} \) flight is \( 60 \cdot n \) seconds.
- The required minimum time headway \( h \) between successive aircraft is constant and deterministic, with \( h = 90 \) seconds.
- Flight arrivals are metered at the runway at a constant headway of \( a \) time units. To maximize throughput and minimize expected delay, \( a \) is set equal to the minimum headway \( h, a = 90 \) seconds.
- Aircraft execute 4D trajectories with some imprecision error that follows a normal distribution with zero mean. The error standard deviation \( \sigma \) is 10 seconds and 60 seconds for equipped and non-equipped aircraft, respectively.
We consider five cases for the penetration rate of equipped aircraft in the system: 10%, 30%, 50%, 70%, and 90%. For each rate, we randomly generate 20 different sequences of equipped and non-equipped aircraft. For example, when the rate is 50%, 20 different sequences of equipped and non-equipped aircraft are created, with each sequence containing a total of 90 equipped aircraft. In this way, 100 demand scenarios are formulated.

To quantify the effect of a BEBS policy, we compare total delays for the system under a BEBS and a First-Scheduled-First-Served (FSFS) queue discipline. The latter is considered as the current service practice, reflecting the Ration-by-Schedule policy of the FAA.

While the definition of FSFS is straightforward, a concrete one for BEBS has been not yet established officially. As a result, we propose three alternatives of BEBS service that differ in the number of Maximum Position Shifts (MPS) denoted as \( p \). The following steps describe the process for assigning STA’s under our proposed BEBS queue disciplines.

- **Step 0**: Partition the schedule of ETA’s into two sets, for equipped – \( ETA^E \), and non-equipped aircraft – \( ETA^Neq \).

- **Step 1**: Set \( STA_i = \min\left( ETA_i^E, ETA_i^Neq \right) \).

- **Step \( i \)**: If \( \{ ETA_{r+1}^E - STA_{i-1} > h \text{ and } ETA_{k+1}^Neq < ETA_{r+1}^E \} \) or \( \{ i - N(i) > p \} \), then \( STA_i = \max\left( STA_{i-1} + h, ETA_{k+1}^Neq \right) \); otherwise, \( STA_i = \max\left( STA_{i-1} + h, ETA_{r+1}^E \right) \). The indices \( r \) and \( k \) denote the number of equipped and non-equipped aircraft, respectively, that have already been assigned an STA. The term \( N(i) \) denotes the cumulative number of non-equipped aircraft included in the set of the first \( i \) aircraft, sorted according to their ETA.

- **Continue until all flights have been assigned an STA.** Note that since we are metering at a constant rate \( 1/h \) that is lower than demand rate, we always have that \( STA_i = STA_{i-1} + h \).

The proposed priority policy attempts to schedule all equipped aircraft as close as possible to their ETA, while looking for gaps in the schedule to squeeze in non-equipped aircraft. The gaps are defined as those time intervals where the time difference between the last assigned STA and the ETA of the next available equipped aircraft is more than one headway \( h \). Additionally, we impose a maximum position shift constraint, which controls the shift in the landing order of any non-equipped aircraft between the original and the metered schedule of arrivals. Three values for that position shift, \( p \), are considered in this study: unlimited, 4, and 2. In the first case, we give highest priority to equipped aircraft, as the landing order of any non-equipped aircraft may be shifted further upstream by an unlimited number of positions. When \( p = 4 \) or 2, we constrain the maximum number of position shifts for the non-equipped aircraft.
The queueing model presented in section 2.1 is suitable for analyzing the above BEBS policy. The disaggregate structure of the model makes it feasible to track individual aircraft delays, and thus it provides delay estimates resulting from different sequences of STAs. Additionally, the model accounts for different levels of errors in trajectory execution, and therefore it can address the two classes of equipped and non-equipped aircraft.

2.4.2 Results

For the arrival demand and service conditions specified in the previous sub-section, the expected delay to each flight is computed using equations (2.5)-(2.7). Delay for flight \( i \) is defined as the difference between the expected time of landing, \( E(D_i) \), and its earliest time of arrival at the runway, \( E(TA_i) \). A total of 400 operational scenarios were analyzed that vary across three dimensions: penetration rate of equipped aircraft, sequences in the ETA schedule of equipped and non-equipped aircraft, and queue discipline as defined by number of MPS. For purposes of illustration, we average out the results along the dimension of sequencing. Therefore, the values displayed in Figure 2.10 are the average across the ten different initial schedules. Note that the delay to non-equipped (equipped) aircraft is the length of the red (blue) region in each bar.

Under all queue disciplines, the amount of delay absorbed by equipped aircraft increases with penetration rate. That is expected, as higher penetration rate implies increased number of equipped aircraft in the system. For example, at 90% penetration rate, equipped aircraft incur the largest portion of total system delays.

For the same penetration rate of equipped aircraft, total delays are almost the same across the four queue disciplines. That is because in all 400 scenarios the schedule of metered arrivals is identical, \( STA_i = STA_{i-1} + h \), and consequently the server is never idle. Although delay for each flight may differ across the four queue disciplines, total expected delay of the queueing system remains almost unchanged. The small discrepancies are due to different sequences of normal random variables with uneven standard deviations, varying along the four queue-service policies. In this way, our model can capture the stochastic effect of different priority rules, even if it appears minimal in the context of this analysis. Most important, our model demonstrates that for a given mix of equipped and non-equipped aircraft different priority queue disciplines result in almost the same expected total system delay.

At the same time, a significant transfer of delays from equipped to non-equipped aircraft is achieved through the proposed BEBS policies. Table 2.4 shows the percent changes with regards to the FSFS policy. In all cases, the BEBS policy causes a delay reduction to equipped aircraft and an increase to non-equipped aircraft. As the system moves toward higher penetration rates of equipped aircraft, the effect of the BEBS policy is diminished. For example, a BEBS policy with two position shifts at maximum, induces a 48% reduction in total delay when the proportion of equipped aircraft is 10%, but only 7% when that proportion is 90%. Therefore, under high penetration rates equipped aircraft cannot internalize the same amount of delay savings as in lower penetration rates, since at
high penetration rates only a small number of non-equipped aircraft are available to be delayed.

For lower values of MPS, while equipped aircraft continue to exhibit delay savings, the amount of additional delay to non-equipped aircraft is moderated. The delay allocation between equipped and non-equipped aircraft under a BEBS policy with MPS=2 is the closest to the pattern under FSFS, but still generates substantial delay savings for equipped aircraft.

Figure 2.10 – Average total delays under various queue disciplines and penetration rates
Table 2.4 – Average percent change of total delays between FSFS and BEBS policies

<table>
<thead>
<tr>
<th>delay FSFS (min)</th>
<th>MPS=Unlimited</th>
<th>MPS=4</th>
<th>MPS=2</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipped</td>
<td>894</td>
<td>-94.40%</td>
<td>-78.20%</td>
</tr>
<tr>
<td>Non-Equipped</td>
<td>7664</td>
<td>11.03%</td>
<td>8.95%</td>
</tr>
<tr>
<td>30%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipped</td>
<td>2547</td>
<td>-94.43%</td>
<td>-37.43%</td>
</tr>
<tr>
<td>Non-Equipped</td>
<td>5991</td>
<td>40.18%</td>
<td>15.67%</td>
</tr>
<tr>
<td>50%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipped</td>
<td>4258</td>
<td>-93.97%</td>
<td>-23.61%</td>
</tr>
<tr>
<td>Non-Equipped</td>
<td>4256</td>
<td>94.68%</td>
<td>23.36%</td>
</tr>
<tr>
<td>70%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipped</td>
<td>6001</td>
<td>-80.88%</td>
<td>-17.07%</td>
</tr>
<tr>
<td>Non-Equipped</td>
<td>2474</td>
<td>192.88%</td>
<td>40.93%</td>
</tr>
<tr>
<td>90%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipped</td>
<td>7549</td>
<td>-31.03%</td>
<td>-12.92%</td>
</tr>
<tr>
<td>Non-Equipped</td>
<td>837</td>
<td>272.52%</td>
<td>121.50%</td>
</tr>
</tbody>
</table>
3.

Model with runway occupancy time

In this chapter we examine a different queueing system than that studied before, by employing a more detailed model for the service process. More specifically, in Chapter 2 a generic deterministic minimum headway is employed, which captures both miles-in-trail separation and the requirement, likely to remain even under NextGen, that not more than one aircraft be on a runway at any given time. Here, we decompose the two constraints, and study the impact of runway occupancy time (RT) on delays by explicitly modeling it as a random variable. Consequently, the server in our queueing system is solely the threshold of a runway and not a general fix, a term that in the previous chapter included both points in the airspace and runway’s threshold.

The present chapter is structured as follows: Section 3.1 presents the general form of our model and its approximate solution, and reports on experiments conducted to assess the accuracy of the model against simulation. In Section 3.2 we consider a special case, in which aircraft are metered at a constant rate and the separation requirements as well as 4DT precision levels and RT distribution are the same for all aircraft. Working with a queueing model in this simplified setting enables us to determine an optimal metering rate that attains high throughput while keeping delays due to imprecise adherence to 4DT’s and varying runway occupancy time small.

3.1 General Model

Our modeling approach, similar to that in Chapter 2, features a disaggregate queueing model of scheduled arrivals. The model is further refined here, to include RT as a third random variable inside the maximum operator. However, a critical departure from the previous analysis is the assumption that precision errors, as well as RT, follow a Gumbel distribution. We discuss that in the following subsection.
3.1.1 Model formulation

Our queueing system consists of a single fix, which is a runway’s threshold, and of aircraft that have been assigned scheduled times of arrival at the fix. For each aircraft $i$ to traverse the threshold, two rules must be satisfied for safety reasons: there is not another aircraft on the runway, and there is a minimum longitudinal separation between aircraft $i$ and the next one to land $i+1$. Which constraint is binding often depends on the weather conditions present at the airport. For example, at US airports under visual meteorological conditions, aircraft frequently obtain clearance for visual approach to the runway, and single runway occupancy is then the binding constraint. Under instrument meteorological conditions though, consecutive aircraft must maintain a minimum headway $h_i$, which can vary over pairs of arriving aircraft but usually exceeds the average runway occupancy time (RT). Also, there may be situations where RT and $h_i$ are almost equal and the binding constraint varies among aircraft pairs. Since air traffic controllers, with guidance from separation rules, decide values for $h_i$, we consider it as a deterministic variable in our model. To the contrary, we allow stochastic variation in RT, caused by aircraft weight, speed at touchdown, condition of the runway surface, and the pilot’s choice of runway exit.

Undertaking an approach similar to that in Chapter 2, the actual time airplane $i$ would arrive at the runway threshold, $A_i$, consists of a deterministic and a stochastic portion, $A_i = a_i + \tilde{A}_i$. The deterministic component $a_i$ is its scheduled arrival time at that fix, while the stochastic component $\tilde{A}_i$ represents the lateness (positive or negative) with which the aircraft arrives at the fix, due to imprecision in trajectory adherence. It is again emphasized that $\tilde{A}_i$’s do not represent factors such as departure delays, traffic management initiatives, severe weather, or en-route congestion, that cause significant amounts of delays; we assume that such factors have already been incorporated in the determination of scheduled arrival times $a_i$.

This allows us to assume that the deviations $\tilde{A}_i$’s are small enough that serving aircraft on a First-Scheduled-First-Served (FSFS) order will not significantly increase delays. Under a FSFS queue discipline, the actual time airplane $i$ traverses the fix, $D_i$, would be the greatest of three times: $A_i$ if there is no delay caused by the single occupancy or minimum headway requirement, or the time the previous scheduled aircraft $i-1$ crossed the fix plus a minimum required separation headway $h_{i-1,i}$ between the two aircraft, or the time the previous scheduled aircraft cleared the runway $D_{i-1} + O_{i-1}$, where $O_{i-1}$ denotes the runway occupancy time of aircraft $i-1$ measured from the moment it crosses the runway threshold until the moment it clears the runway. The actual times that aircraft cross the fix under study would then be:

\[
D_1 = A_1
\]

\[
D_i = \max\left( A_i, D_{i-1} + h_{i-1,i}, D_{i-1} + O_{i-1} \right), \quad \forall i \geq 2
\]
The distribution of $D_i$ clearly depends on the distributions of $O_i$ and $\tilde{A}_i$. The specification of the RT distribution can be determined from field measurements, when available. For example, Xie, Shortle, and Donohue (2003) fit RT data from the Atlanta Hartsfield Airport with a normal distribution. Jeddi, Shortle, and Sherry (2006) analyze RT data from the Detroit Metropolitan Airport and fit the empirical data with distributions that exhibit slightly fatter tails than the normal and are skewed to the right.

Systems where RT is well approximated by a normal distribution can be modeled by assuming that $\tilde{A}_i$ is also normal and applying the Clark approximation method to solve (3.1). A detailed discussion on Clark’s method and its accuracy is presented in Chapter 2 of this dissertation.

It is often the case, however, that the RT distribution is not perfectly symmetric, reflecting situations when aircraft miss their standard runway exit and occupy the runway for an additional amount of time. This results in a positive skewness for the RT distribution, at least for certain runways, as shown empirically in Figure 3.1, as well as the previously cited work of Jeddi, Shortle, and Sherry (2006). To model those cases we must employ a distribution that is bell-shaped and has positive skewness. One such distribution, which also has some useful properties in connection with the $\max$ operator, is the Gumbel. The Gumbel distribution was derived to model the distribution of the maximum (or the minimum) of a number of samples from various distributions, a case that applies in (3.1). We thus assume that each aircraft’s runway occupancy time is a Gumbel random variable with location $\mu_i$, and scale $\xi_i$. That is:

$$F_{O_i}(x) = \exp\left\{-\exp\left(\frac{\mu_i - x}{\xi_i}\right)\right\}.$$  

![Figure 3.1. Histogram for runway occupancy times measured at runway 17C at DFW (source: ASDE-X database; dates of observation: 4/6/2008 - 7/11/2008)](image)

The probability distribution for $\tilde{A}_i$ is generated by convolving the individual distributions of low-correlated stochastic factors that cause flight $i$ to arrive at a time other than its
scheduled one \( a_i \). Therefore, one can expect each \( \tilde{A}_i \) to follow a normal or a similar bell-shaped distribution. This is supported by recent evidence from flight trials carried out in the Stockholm terminal maneuvering area that investigated how well aircraft could meet an assigned time of arrival to a feeder fix (Manzi, 2009). It was found that the adherence error distribution is bell-shaped and slightly skewed to the right. It is reasonable therefore to assume that stochastic error \( \tilde{A}_i \) follows a Gumbel distribution exhibiting location parameter zero (without loss of generality), and scale parameter \( \sigma_i \). While these distributional assumptions may not apply in all cases, they are likely to be valid for a substantial number.

This completes the formulation of our queueing model. In summary, the model requires as inputs a schedule of arrival times \( a_i \), vector of scale parameters \( \sigma_i \) of aircraft’s trajectory adherence errors, a vector of time separation headways \( h_{i-1,i} \), and RT distributions with location parameters \( \mu_i \) and scale parameters \( \xi_i \). Next, we are interested in obtaining estimates of the mean and variance of \( D_i \), the time each aircraft traverses the runway threshold.

### 3.1.2 Approximate solution

Ideally one would like to obtain exact analytical estimates of the first two moments of \( D_i \) for every \( i \); but this is a difficult task, even for \( i = 2 \) : to estimate the moments of \( D_2 \) one has to find the distribution of \( \max(D_1 + h_{1,2}, D_1 + O_i) \) which is the maximum of two correlated and non-identical random variables. However, approximate estimates can be obtained, the accuracy of which is discussed in Section 3.1.3.

While the sum of two independent Gumbel random variables does not yield another Gumbel, it can be approximated as Gumbel with an error that depends on the ratio of the scale parameters (Nadarajah & Kotz, 2008a). Here we assume that \( D_1 + O_i \) is distributed as a Gumbel random variable with, as can be easily shown, location \( \mu_i + \gamma \left( \xi_i + \sigma_i + \sqrt{\xi_i^2 + \sigma_i^2} \right) \) and scale parameter \( \sqrt{\xi_i^2 + \sigma_i^2} \), where \( \gamma \) is Euler’s constant \( (\gamma \approx 0.577) \).

The term \( Z_i = \max(D_1 + h_{1,2}, D_1 + O_i) \) then becomes the greatest of two correlated and non-identically distributed Gumbel random variables. The joint cumulative distribution function of \( D_1 + h_{1,2} \) and \( D_1 + O_i \) is given by the bivariate Oliveira distribution, which has the general form:

\[
F_{XY}(x,y) = \exp \left( - \left( e^{-m(x-a)/b} + e^{-m(x-c)/d} \right)^{1/m} \right), \quad \text{with} \quad m = 1 / \sqrt{1 - \rho}
\]
where $a$ and $c$ are the location parameters, and $b$ and $d$ are the scale parameters of the Gumbel random variables $X$ and $Y$, and $\rho$ is their correlation coefficient. Note that in our case $\rho = \left(\text{Var}[D_i]/(\text{Var}[D_i] + \text{Var}[O_i])\right)^{1/2}$. The distribution of their maximum, $Z = \max(X,Y)$, is then:

$$F_Z(z) = F_{X,Y}(z,z) = \exp\left(-\left(e^{-m(z-a)/b} + e^{-m(z-c)/d}\right)^{1/m}\right)$$

The $k^{th}$ moment of $Z$, if it exists, can be estimated as follows:

$$E[Z^{(k)}] = k \int_0^\infty z^{k-1}(1-F_Z(z))dz - k \int_0^\infty z^{k-1}F_Z(z)dz \quad (3.2)$$

We approximate $Z$ as a Gumbel random variable with moments given by (3.2). That enables us to find the moments of any $D_i$ in a recursive manner, approximating in each step the maximum of two Gumbel random variables with another Gumbel. It should be noted the moment estimation process involves numerical integration.

Finally, for a stream of $N$ flights scheduled for to arrive at a fix, the total expected delay is defined as:

$$E[W_N] = \sum_{i=1}^N (E[D_i] - a_i) \quad (3.3)$$

3.1.3 Approximation Error Tests

Our model is based on approximating the sum and the maximum of two Gumbel random variables with another Gumbel. In particular, in estimating $D_i = \max\left(A_i, D_{i-1} + h_{i-1,j}, D_{i-1} + O_{i-1}\right)$ it is assumed that both terms $D_{i-1} + O_{i-1}$ and $\max\left(D_{i-1} + h_{i-1,j}, D_{i-1} + O_{i-1}\right)$ yield Gumbel random variables. That enables the estimation of the mean and variance of $D_i$, which is then also approximated as Gumbel. However, each pair-wise operation introduces some error that is propagated and might affect the accuracy of our estimates. To test the accuracy of our method in the context of queueing modeling for air traffic, several operational scenarios were considered. The estimates from the analytical queueing model were then compared against the average of $10^4$ Monte Carlo simulation runs, which is considered as ground truth.

The scenarios were generated in such a way that a wide variety of combinations of factors that may affect the accuracy of our method are considered. Those factors can be identified as: a) the difference between average RT and minimum required separation
headway, b) the variance in RT and its relative magnitude over 4DT adherence error, and c) the scheduled time interval between consecutive arrivals. Each operational scenario was formulated as follows: a total of 120 aircraft must traverse a runway threshold, with an equal mix of small, large and heavy jets. The RT distribution across all aircraft follows a Gumbel distribution with mean 50 seconds and standard deviation 8 seconds, while the minimum required separation between any two successive aircraft \( h_{i-1,j} \) is scenario-specific, as will be discussed in the next paragraph. Each aircraft is assigned a scheduled time of arrival at the server \( a_i = a_{i-1} + \text{max}(50, h_{i-1,j}) + b \), where \( b \) denotes a buffer time inserted between the scheduled arrival times of any two consecutive aircraft. Aircraft arrive at the fix with some imprecision that follows a Gumbel distribution and has a standard deviation \( \sigma \) (uniform across all aircraft), while zero covariance was assumed across the aircraft arrival times at the fix, \( A_i \). A total of 270 scenarios were examined:

- Three different sets of \( h_{i-1,j} \) to capture situations where the binding constraint is RT, or the separation headways, or potentially either of them: \( h_{i-1,j} = \{30, 35, 40\} \) seconds, \( h_{i-1,j} = \{70, 100, 130\} \) seconds, or \( h_{i-1,j} = \{50, 65, 80\} \) seconds.
- For each set, 10 different sequences of \( h_{i-1,j} \), where each sequence is determined randomly but given an equal mix of the three headway values.
- \( b = 0, 10, \) and 20 seconds (held constant within each scenario).
- \( \sigma = 4, 8, \) and 16 seconds (held constant within each scenario).

For each scenario, we compute \( E[D_i]_{\text{appr}} - a_i \) and \( E[W_N]_{\text{appr}} \) using (3.2) and (3.3), and also \( \bar{D}_i^{\text{sim}} - a_i \) and \( \bar{W}_N^{\text{sim}} \), which denote the average flight delay and total delay, respectively, obtained by \( 10^4 \) simulation runs. Three metrics for the approximation method accuracy were considered:

- Percentage Error in total Delay % (PE): \( \left( E[W_N]_{\text{appr}} - \bar{W}_N^{\text{sim}} \right) / \bar{W}_N^{\text{sim}} \cdot 100 \)
- Absolute Error in Total Delay (AE): \( |E[W_N]_{\text{appr}} - \bar{W}_N^{\text{sim}}| \)
- Flight Departure Time Mean Absolute Deviation (MAD): \( \left( \sum_{i=1}^{120} |E[D_i]_{\text{appr}} - \bar{D}_i^{\text{sim}}| \right) / 120 \)

The first two metrics evaluate the accuracy of the approximation method in estimating the expected total aircraft delay against assigned scheduled times of arrival. The third
metric provides a measure of the error in predicted outcomes for individual flights. The results are summarized in Tables 3.1-3.3.

Table 3.1. Approximation error test results when 4DT precision has standard deviation $\sigma = 16$ seconds

<table>
<thead>
<tr>
<th>Buffer = 0 (sec)</th>
<th>Buffer = 10 (sec)</th>
<th>Buffer = 20 (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = {30, 35, 40}$</td>
<td>$h = {50, 65, 80}$</td>
<td>$h = {70, 100, 130}$</td>
</tr>
<tr>
<td>PE</td>
<td>$A_E$ (sec)</td>
<td>MAD (sec)</td>
</tr>
<tr>
<td>6.05%</td>
<td>535.99</td>
<td>4.47</td>
</tr>
<tr>
<td>(0.60%)</td>
<td>(57.85)</td>
<td>(0.47)</td>
</tr>
</tbody>
</table>

$h = \{30, 35, 40\}$ | $h = \{50, 65, 80\}$ | $h = \{70, 100, 130\}$ |
| PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) |
| -4.46% | 502.73 | 4.69 | -0.44% | 8.23 | 0.21 | -0.07% | 3.15 | 0.16 |
| (0.51%) | (63.55) | (0.53) | (0.21%) | (3.69) | (0.04) | (0.32%) | (2.49) | (0.05) |

$h = \{30, 35, 40\}$ | $h = \{50, 65, 80\}$ | $h = \{70, 100, 130\}$ |
| PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) |
| -1.76% | 119.74 | 1.07 | -0.32% | 11.96 | 0.22 | -0.25% | 5.35 | 0.17 |
| (0.82%) | (42.25) | (0.19) | (0.76%) | (7.27) | (0.09) | (0.42%) | (2.09) | (0.05) |

Table 3.2. Approximation error test results when 4DT precision has standard deviation $\sigma = 8$ seconds

<table>
<thead>
<tr>
<th>Buffer = 0 (sec)</th>
<th>Buffer = 10 (sec)</th>
<th>Buffer = 20 (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = {30, 35, 40}$</td>
<td>$h = {50, 65, 80}$</td>
<td>$h = {70, 100, 130}$</td>
</tr>
<tr>
<td>PE</td>
<td>$A_E$ (sec)</td>
<td>MAD (sec)</td>
</tr>
<tr>
<td>9.43%</td>
<td>610.93</td>
<td>5.09</td>
</tr>
<tr>
<td>(9.89%)</td>
<td>(63.18)</td>
<td>(0.53)</td>
</tr>
</tbody>
</table>

$h = \{30, 35, 40\}$ | $h = \{50, 65, 80\}$ | $h = \{70, 100, 130\}$ |
| PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) |
| -3.23% | 335.46 | 2.95 | -4.34% | 3.65 | 0.09 | -0.64% | 0.99 | 0.07 |
| (6.52%) | (57.50) | (0.42) | (0.38%) | (1.53) | (0.01) | (0.18%) | (6.55) | (0.00) |

$h = \{30, 35, 40\}$ | $h = \{50, 65, 80\}$ | $h = \{70, 100, 130\}$ |
| PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) |
| -2.61% | 91.32 | 0.76 | -6.83% | 7.86 | 0.14 | -0.10% | 0.85 | 0.07 |
| (0.31%) | (11.57) | (0.10) | (1.79%) | (8.62) | (0.08) | (0.19%) | (6.45) | (0.01) |

Table 3.3. Approximation error test results when 4DT precision has standard deviation $\sigma = 4$ seconds

<table>
<thead>
<tr>
<th>Buffer = 0 (sec)</th>
<th>Buffer = 10 (sec)</th>
<th>Buffer = 20 (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = {30, 35, 40}$</td>
<td>$h = {50, 65, 80}$</td>
<td>$h = {70, 100, 130}$</td>
</tr>
<tr>
<td>PE</td>
<td>$A_E$ (sec)</td>
<td>MAD (sec)</td>
</tr>
<tr>
<td>13.26%</td>
<td>739.22</td>
<td>6.17</td>
</tr>
<tr>
<td>(2.25%)</td>
<td>(123.57)</td>
<td>(1.65)</td>
</tr>
</tbody>
</table>

$h = \{30, 35, 40\}$ | $h = \{50, 65, 80\}$ | $h = \{70, 100, 130\}$ |
| PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) |
| -3.15% | 285.26 | 2.51 | -1.96% | 6.30 | 0.09 | 0.35% | 2.60 | 0.07 |
| (1.44%) | (8.15) | (0.02) | (1.34%) | (2.40) | (0.03) | (1.28%) | (1.10) | (0.04) |

$h = \{30, 35, 40\}$ | $h = \{50, 65, 80\}$ | $h = \{70, 100, 130\}$ |
| PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) | PE | $A_E$ (sec) | MAD (sec) |
| -3.41% | 65.26 | 0.60 | 1.64% | 4.52 | 0.07 | 2.15% | 4.76 | 0.07 |
| (1.39%) | (25.53) | (0.19) | (3.25%) | (7.39) | (0.09) | (2.37%) | (3.02) | (0.04) |

Each entry in the tables represents the average value for each metric across the ten scenarios of different $h_{i-1,j}$ sequences, while the standard deviation is enclosed in parenthesis. The presence of buffer affects the approximation error in a substantial way. When $b = 10$ or 20 seconds, the percent and absolute error in estimating total delay never exceed 5% and 0.6 minutes, respectively, while the MAD metric indicates that the approximation method estimates the expected delay of each aircraft with accuracy better than 0.5 second, on average. However, when no buffer is added between scheduled

*While performing numerical integration of (2) in MatLab some abnormal results were obtained, either in the form of "NaN" indication or as extremely high approximation errors that were different by 10 or even 100 times than other estimates under the same model parameters. After investigating the error propagation, it was concluded that the cause of these outliers was unstable performance of MatLab's `quadgk` function, and for that these cases were not included in the set of results summarized in Tables 1-3."
arrivals, the accuracy of the approximation decreases. The Total Delay PE metric exhibits an average value of 13% when \( \sigma = 16 \) seconds, and 9.5% when \( \sigma = 8 \) seconds. Moreover, the error in estimating total delay in those cases is biased upward by from 10 minutes to 13 minutes—around 6 seconds per flight.

In summary, these experimental results indicate that our proposed approximation technique accurately predicts average deviations from schedule when there is a buffer between scheduled arrivals of more than 10 seconds, but caution must be exercised in applying the method when this buffer is zero. The accuracy of the technique improves as the standard deviation of trajectory precision error becomes larger, compared to that of RT. A similar argument holds when separation headways are the binding constraints or at least potentially equally binding with single runway occupancy rule.

### 3.2 Using buffers to mitigate the effects of stochastic runway occupancy time and precision error

Similar to our analysis in Section 2.2, we consider again the case where average demand for service exceeds capacity over a considerable period of time, after which demand drops to zero. We differentiate between delays from imperfect trajectory adherence and varying RT (‘stochastic’), and delays due to reduced throughput (‘deterministic’). In this section we employ the model presented in 3.1.1 to quantify trade-offs between the two types of delay.

#### 3.2.1 A simplified case

We focus on a single fix, and make the following assumptions:

- Flight arrivals are uniformly scheduled during the analysis period, with metering headway \( a \) (we assume that demand for landings is persistent and a constant metering headway can be maintained).
- The required minimum time headway \( h \) between successive aircraft is constant and deterministic.
- Stochastic errors \( \hat{A}_i \) are \( i.i.d. \) Gumbel random variables with zero location and scale parameter \( \sigma_1 \).
- Runway occupancy times \( O_i \) are \( i.i.d. \) Gumbel random variables with location \( \mu \) and scale parameter \( \sigma_2 \).
• \( \bar{A}_i \)'s and \( O_i \)'s are uncorrelated.

Let \( s^* \) denote the binding constraint – in an expected value sense – between aircraft minimum separation requirements and mean RT, \( s^* = \max(h, \mu + \sigma \cdot \gamma) \). In a fully deterministic environment, where \( \sigma_1 = \sigma_2 = 0 \), there will be no delays at the fix if \( a \geq s^* \), and every flight would traverse the fix at a deterministic time \( d_i = i \cdot a \). Accounting for random effects, if an airplane arrives later than its scheduled arrival time, then it might cause delay to airplanes upstream of it. The actual time each aircraft traverses the fix can be expressed as \( D_i = d_i + \tilde{D}_i \), where \( \tilde{D}_i \) is the stochastic delay due to imprecise adherence to 4DT's and varying RT. We are interested in quantifying the expected amount of stochastic delay \( \tilde{D}_i \) that each airplane in the stream incurs. The model described by (3.1) in Section 3.1.1 thus becomes:

\[
\tilde{D}_1 = \bar{A}_1
\]  

(3.4a)

\[
\tilde{D}_i = \max(\bar{A}_i, \tilde{D}_{i-1} + h - a, \tilde{D}_{i-1} + O_{i-1} - a), \quad \forall i \geq 2
\]  

(3.4b)

Deriving an analytical expression for each \( E(\tilde{D}_i) \) is intractable. Instead, we resort to dimensionless analysis and examine stochastic delays that occur when \( \sigma_2 = 1 \):

\[
\tilde{D}_i^* \triangleq \max(\bar{A}_i^*, \tilde{D}_{i-1}^* + h - a, \tilde{D}_{i-1}^* + O_{i-1}^* + \mu - a)
\]

where \( \bar{A}_i^* \sim \text{Gumbel}(0, \zeta) \), \( O_i^* \sim \text{Gumbel}(0,1) \), and \( \zeta = \sigma_1 / \sigma_2 \). Estimating stochastic delays for cases other than \( \sigma_2 = 1 \) is straightforward, through the following statement:

**Proposition 1.** For given values of the relative buffers \( \Delta = (a - h) / \sigma_2 \), and \( \Theta = (a - \mu) / \sigma_2 \), the sequence of random variables \( \tilde{D}_i \) is the sequence of random variables \( \tilde{D}_i^* \) multiplied by the scale of the runway occupancy time distribution:

\[
\tilde{D}_i = \sigma_2 \cdot \tilde{D}_i^*
\]  

(3.5)

**Proof.** For \( \sigma_2 = 1 \) we write:

\[^* \text{Since RT is assumed a Gumbel random variable, even if its mean is greater than the minimum headway, } \mu + \sigma_\gamma > h, \text{ its distribution may yield values that are lower than } h. \text{ Thus, the term } \text{binding} \text{ here implies that } \text{most of the distribution mass is located at values greater than } h. \]
\[
\tilde{D}_i^* = \max\left(\tilde{A}_i^*, \tilde{D}_{i-1}^*, \Delta, \tilde{D}_{i-1}^* + O_{i-1}^* - \Theta\right)
\] (3.6)

Assume that for some \(i \geq 3\), \(\tilde{D}_{i-1} = \sigma_2 \cdot \tilde{D}_{i-1}^*\). Then, employing a known property of the Gumbel distribution, we have:

\[
\tilde{D}_i = \max\left(\sigma_2 \cdot \tilde{A}_i^*, \sigma_2 \cdot \tilde{D}_{i-1}^* - \sigma_2 \cdot \Delta, \sigma_2 \cdot \tilde{D}_{i-1}^* + \sigma_2 \cdot O_{i-1}^* - \sigma_2 \cdot \Theta\right) = \sigma_2 \cdot \tilde{D}_i^*
\]

But note that for \(i = 2\):

\[
\tilde{D}_2 = \max\left(\sigma_2 \cdot \tilde{A}_2^*, \sigma_2 \cdot \tilde{A}_1^* - \sigma_2 \cdot \Delta, \sigma_2 \cdot \tilde{A}_1^* + \sigma_2 \cdot O_1^* - \sigma_2 \cdot \Theta\right) = \sigma_2 \cdot \tilde{D}_2^*
\]

Hence, for \(i \geq 3\) the result follows by induction. \(\because\)

In this way, the problem of estimating \(E(\tilde{D}_i)\) is reduced to estimating \(E(\tilde{D}_i^*)\). That can be performed either by employing the approximate solution presented in Section 3.1.2, or through simulation. For a surge of \(N\) aircraft arrivals, the expected loss \(L\) from those two types of delay is:

\[
E[L] = \frac{1}{2} \cdot N \cdot \left(N \cdot a - N \cdot s^*\right) + \beta \cdot \sum_{i=1}^{N} E[\tilde{D}_i] \cdot \sigma_2
\]

which can also be expressed as:

\[
E[L] = \left(\frac{1}{2} \cdot N^2 \cdot \Delta + \beta \cdot \sum_{i=1}^{N} E[\tilde{D}_i]\right) \cdot \sigma_2 \tag{3.7a}
\]

or

\[
E[L] = \left(\frac{1}{2} \cdot N^2 \cdot \Theta + \beta \cdot \sum_{i=1}^{N} E[\tilde{D}_i]\right) \cdot \sigma_2 \tag{3.7b}
\]

depending on whether minimum required headway or RT is binding.

Following a notation similar to that of Section 2.2, the coefficient \(\beta\) in equations 3.7a and 3.7b is the relative cost of stochastic delay over the delay due to reduced throughput. As noted earlier, delays due to reduced metering rate can be planned for well in advance,

\* If \(X\) follows a Gumbel distribution with location parameter \(\mu\) and scale parameter \(\xi\), \(F_X(x) = \exp\left(-\exp\left((\mu - x)/\xi\right)\right)\), then multiplying \(X\) by a constant \(\delta > 0\) would result in another Gumbel distribution: \(\delta X \sim Gumbel(\delta \cdot \mu, \delta \cdot \xi)\).

47
taken at a higher altitude, in less busy airspace, and with relatively small speed reductions. Therefore $\beta \geq 1$.

### 3.2.2 Results

After developing methods for estimating each flight’s expected stochastic delay, we now analyze the trade-off between stochastic delay and throughput. The parameter space of our model has five dimensions: $\Delta$, $\Theta$, $\zeta$, $\beta$, and $N$. For illustration in two dimensions, we consider two values for the relative cost of stochastic over deterministic delay, $\beta = 1$ and $\beta = 2$, an arrival surge of $N = 100$ aircraft, and three cases for the relationship between $\Delta$ and $\Theta$: $\Delta = \Theta + 2$, $\Theta = \Delta + 2$, and $\Theta = \Delta$. Those correspond to situations where the binding constraint is RT, or the minimum separation headway, or neither. For example, $\Delta = \Theta + 2$ implies that average RT $\mu$ is greater than minimum headway $h$ by $2\sigma_2$. For each case, the normalized (after setting $\sigma_2 = 1$) expected loss $E[L^*]$ as a function of $\Delta$, $\Theta$, and $\zeta$ is plotted in Figures 3.2–3.7. All curves for $E[L^*]$ were obtained through simulation.

![Figure 3.2. Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=1$, $N=100$, and $\Delta=\Theta+2$](image)

Figure 3.2. Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=1$, $N=100$, and $\Delta=\Theta+2$
Figure 3.3. Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=1$, $N=100$, and $\Theta=\Delta+2$.

Figure 3.4. Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=1$, $N=100$, and $\Theta=\Delta$.

The first set of Figures, 3.2–3.4, includes all formulations where the unit costs of deterministic and stochastic delay are equal, $\beta = 1$. In all three cases, minimum expected losses are attained when there is no buffer between scheduled arrivals. However, when average RT is either greater than or equal to the minimum headway (Fig. 3.2 and 3.4),
one can achieve efficiency losses almost equal to minimal by allowing a buffer up to $0.4 \cdot \sigma_2$. In such circumstances, which generally reflect favorable weather, one might want to set the buffer to at least this value without further inquiry about the value of $\beta$. Also, a system where minimum headway is the binding constraint is more efficient, as can be inferred by comparing Fig. 3.3 with Figures 3.2 and 3.4. While minimum losses range between 800–1300 when $\Theta = \Delta + 2$, they almost triple when RT is the binding constraint, ranging between 2900-3300 (see Fig. 3.2), or nearly quadruple when both constraints are equally binding, ranging between 4000–4300 (see Fig. 3.4). In the latter two cases, the term $\tilde{D}_{i-1} + O_{i-1}^* - \Theta$ is large enough to affect the value of $\tilde{D}_i^*$ in equation 3.6, therefore $O_{i-1}^*$ contributes a stochastic effect in every step of the recursion, which in turn results in greater efficiency loss.

Figure 3.5. Expected loss from deterministic and stochastic delays, for alternative values of $\zeta$ and for $\beta=2$, $N=100$, and $\Delta=\Theta+2$
Figures 3.5–3.7 include results for the cases where the cost of stochastic delays is double that of deterministic delays, $\beta = 2$. Here the effect of stochastic delays becomes more
prominent, and the values of relative buffer that minimize $E[L^+]$ range between 0.1 and 0.8. A value of $\Theta = 0.8$ translates to a buffer of 8 seconds between consecutive aircraft when $\sigma_2 = 10$ seconds, a realistic value. Assuming a value for $\mu$ on the order of a minute, such a buffer would reduce planned throughput about 15% compared to the maximum possible under deterministic conditions. As is the case when $\beta = 1$, the system suffers significantly smaller efficiency loss when minimum headway is the binding constraint. Moreover, for the same values $\zeta$ of dispersion ratio of precision error to RT, the minimum point of $E[L^+]$ is attained for a substantially smaller value of relative buffer, when $h$ is the binding constraint. Examining the case where $\zeta = 2$, the optimal relative buffer equals 0.6 when RT is the binding constraint (see Fig. 3.5), and 0.8 when both constraints are of the same magnitude (see Fig. 3.7), while it drops to 0.1 when $h$ is the binding constraint (see Fig. 3.6). Therefore, when average RT is approximately equal to or greater than $h$, variation in RT causes a substantially greater amount of stochastic delay that cannot be absorbed with an optimal buffer as small as in the case where $h$ is binding. As a result in these cases, a larger relative buffer is justified for the system. Also, it should be noted that for non-zero buffers the approximation method yields accurate estimates, as discussed in Section 3.1.3.

One can also examine the effect of improving aircraft’s capability of adhering to 4DT’s by plotting the minimum values of $E[L^+]$ as a function of $\zeta$, the relative magnitude of $\sigma_1$ over variation in RT $\sigma_2$. Figures 3.8–3.9 summarize the results for models with $\beta = 1$ and with $\beta = 2$.

In both figures, minimum expected losses are significantly lower when $h$ is the binding constraint, which confirms and generalizes the previous finding. For the formulation with $\beta = 1$ in Fig. 3.8, minimum losses when $h$ is the binding constraint are three times lower compared to the case with RT as binding constraint, and four times lower when compared to the situation with both $h$ and RT as equally binding. However, losses in efficiency should not be confused with system throughput. Figure 3.10 depicts two scenarios for runway throughput, the first representing an airport under good weather conditions where RT is the binding constraint, $\mu > h$, while in the second case Instrumental Flight Rules (IFR) are in place due to adverse weather and minimum headway is the binding constraint, $\mu' < h'$. Although capacity in the second case is lower compared to the case with good weather, $1/h' < 1/\mu$, efficiency losses are smaller in the case with adverse weather because minimum headway is the binding constraint.
Figure 3.8. Minimum expected losses as a function of $\zeta$, for $\beta=1$, and $N=100$

Figure 3.9. Minimum expected losses as a function of $\zeta$, for $\beta=2$, and $N=100$
Figure 3.10. Comparison of two queueing systems with different capacities and unequal efficiency losses

We also observe in both Figures 3.8 and 3.9 that for $\zeta \leq 0.5$ the two upper curves decrease at a lower rate. In other words, the marginal benefit from improving 4DT adherence reduces as $\sigma_1$ becomes smaller than or equal to $0.5 \sigma_2$. For a typical value of RT standard deviation, $\sigma_2 = 10$ seconds, we conclude that reducing the standard deviation of adherence error beyond 5 seconds yields a diminished benefit in $E[L^*]$. Moreover, the benefit from reducing $\sigma_1$ is greater, both in absolute and in percentage terms, when $h$ is the binding constraint. In Figure 3.9 for example, $E[L^*]$ drops from 2400 to 1100 as $\zeta$ ranges from 2 to 0.1, a reduction of 54%. When RT is binding constraint, the reduction in $E[L^*]$ is 1000 or 20%, and when $h$ and RT are both binding the decrease is 1000 or 17%. Therefore, runway occupancy time, when a binding constraint, moderates the potential benefit from reduced 4D trajectory imprecision.

Finally, it should be noted that trajectory adherence error and uncertainty in RT are complementary. Thus, if one defines the inverse of $\zeta$, $\zeta' = \sigma_2 / \sigma_1$, and conducts the same analysis as that presented in Figures 3.8-3.9, he would observe similar patterns in
the relationship between expected loss in efficiency $E[L']$ and $\zeta'$. Therefore, if uncertainty about RT were reduced – for example through dissemination of information about aircraft weight to the air traffic control system – the loss in system efficiency due to stochastic factors would be monotonically decreasing as a function of RT’s standard deviation $\sigma_2$. 
Paired Arrivals at SFO

Terminal airspace areas surrounding large metropolitan airports that require intensive air traffic control (ATC) resources are expected to stand to gain from NextGen. In this chapter we are concerned with parallel closely-spaced landings at the San Francisco International Airport (SFO) in an operational environment where aircraft execute 4D trajectories with high precision. In order to quantify the potential increase in runway throughput from implementation of 4DT’s, we find the required headways between consecutive arrival pairs as a function of aircraft’s 4D trajectory precision.

4.1 Description of SFO terminal airspace

Due to a variety of factors (geographic location, runway configuration, nature of demand and weather), San Francisco International Airport has a unique set of characteristics that make it an ideal case study for potential NextGen-associated benefits. The runway configuration at SFO is comprised of two sets of parallel runways spaced 750 feet apart and intersecting mid length at right angles, as illustrated in Fig. 4.1. Approximately 83% of the time aircraft depart on either Runway 01L or 01R and arrive on either Runway 28L or 28R (San Francisco International Airport, 2010). This operation is called the “West Plan.” During visual meteorological conditions (VMC) and with westerly winds, aircraft obtain clearance to fly visual approaches to runways 28L and 28R. This work focuses on SFO arrivals under the west plan, landing on Runways 28 Left (28L) and 28 Right (28R) in VMC.
In periods of high demand for arrivals and departures, and due to SFO’s intersecting runway configuration, it is most efficient to have aircraft land as pairs touching down on 28R and 28L with near simultaneous runway occupancy time. In turn, departing aircraft use runways 1L and 1R with similar simultaneous occupancy. Arriving aircraft meet at the final approach point, approximately 5 nautical miles (nmi) from the runways, and traverse the final approach in a staggered arrangement, with the trailing aircraft approximately 1 nmi behind the leading one. In this way, the trailing aircraft will be at a...
safe distance in case the leading executes a missed-approach procedure. As a result of this landing arrangement, the trailing aircraft clears the points where arrival runways intersect with departure runways shortly after the leading aircraft, allowing enough time for two departures to take-off before the next pair of landing aircraft touch down. It is under this mode of operations when the runway system of SFO attains greatest throughput, serving on average 95 to 99 aircraft per hour, 55 of which are typically arrivals (San Francisco International Airport, 2010). The smaller proportion of departures implies that aircraft pairing is not always achieved and in those cases departing aircraft are held on the ground to give priority to landing aircraft. For example, if the trailing aircraft of pair \( k \) arrives slightly late at the final approach fix, there will not be enough time separation between arrival pairs \( k \) and \( k + 1 \) for departures to take-off. In this case, departures must wait on the ground until the pair \( k + 1 \) clears the runways, resulting in a lost slot for take-off.

For aircraft to be paired on arrival under existing practice at SFO requires significant ATC resources. Air traffic controllers use extensive path-stretching and speed assignment, also known as “vectoring”, to ensure that aircraft are paired up and on speed as they cross final fixes for landing. However, there is a high degree of efficiency variance among individual controllers. It is this pairing of aircraft required with the SFO west plan, therefore, that provides an opportunity to substitute controllers’ initiatives with series of waypoints and execution of 4D trajectories. With 4D trajectory control being a future prospect, one may envision an environment where aircraft are assigned scheduled times of arrival at the merging point, which they meet with high precision and with minimal air traffic control involvement. Thus, our objective is to quantify the benefits from precise paired arrivals at SFO, capacitated by 4DT operations.

4.2 Model for estimating headways between pairs of arrivals

A key parameter for the combined arrivals-departures system throughput at SFO is the headway between two consecutive arrival pairs. That should allow sufficient time for arrivals to clear the runways and departures to take-off, before the next pair of arrivals cross the runway thresholds. Currently, air traffic controllers use a 4 nautical miles separation between pairs of arrivals, which includes a buffer to account for all stochastic factors that may prevent coordination of operations and result in delaying take-offs in order to prioritize landings. Precise trajectory-based operations will provide the capability of assigning headways between arrival pairs that eliminate buffer times and ensure a low rate of missed departure slots.

To quantify those benefits, we develop a mathematical model that computes the desired headways between arrival pairs. Aircraft are assigned scheduled times of arrival at the merging point, in particular the final approach point, which they meet with some normally distributed error. Let \( ST_j \) and \( A_j \) be the scheduled and actual arrival time of aircraft \( j \) at the merging point, respectively, and \( O_j \) the time it occupies the runway from the threshold until it passes the intersection point with runway 1L. Also, each aircraft must be fully configured for landing in the final approach segment. That means it must
have landing gear extended, flaps deployed, a stabilized sink rate and on an established Velocity Reference Speed (Vref), a specific final approach speed based on aircraft type, flap setting, landing weight and environmental conditions. We, therefore, consider the time needed by aircraft \( j \) to traverse the final approach segment, \( t_j \), as deterministic. Air traffic controllers in the airport control tower observe incoming arrival traffic and decide whether to hold departures waiting for take-off, in order to secure single runway occupancy by the time arrivals traverse the runway threshold. To reach their decision, they assume a certain amount of time needed for both departures to clear the intersection point with runway 28R. Therefore, that variable should enter our model formulation as a deterministic quantity that we will denote as \( \tau_k \), with index \( k \) representing the pair of aircraft waiting for take-off.

We focus on two pairs of arriving aircraft \( \{i, i+1\} \) and \( \{i+2, i+3\} \), with the lowest index in each pair denoting the aircraft scheduled to arrive first at the merging point, independent of its assigned landing runway. The variable of interest is the headway between the scheduled times of arrival of the two leading aircraft:

\[
ST_{i+2} = ST_i + H 
\]

where we assume that \( A_j \sim \text{Normal}\left(ST_j, \sigma_j\right) \), for \( j = i, i+1, i+2, i+3 \), and \( ST_{i+2} = ST_i + x \).

A relationship remains to be defined between the scheduled times of arrival, \( ST_i \) and \( ST_{i+1} \), of the two aircraft landing in parallel. This must ensure that when aircraft arrive at the merging point they are separated by at least \( b \) time units for aircraft to be in a proper staggered arrangement. If we specify that event to occur with probability at least \( 1 - p' \), an offset \( F \) can be defined as:

\[
F = \left\{ y > 0 : P(A_{i+1} - A_i < b) \leq p' \right\}
\]

where \( A_j \sim \text{Normal}\left(ST_j, \sigma_j\right) \), for \( j = i, i+1 \), and \( ST_{i+1} = ST_i + y \).

Denoting by \( \Phi(x) \) the standard normal cumulative distribution function, we can write:

\[
F = b - \Phi^{-1}(p') \cdot \sqrt{\sigma_i^2 + \sigma_{i+1}^2}
\]
The solution for $H$ in equation 4.1 depends on the type of probability distribution assumed for the random variables $O_i$. As discussed in chapter 3, it is reasonable to assume that runway occupancy time (ROT) follows a normal distribution. Thus, let $O_i$ be normally distributed with mean $\mu_i$ and standard deviation $\xi_i$, and independent of precision errors. Under these conditions, the term $A_j + t_j + O_j$ is also a normal random variable with mean $ST_j + t_j + \mu_j$ and variance $\sigma^2 + \xi^2_j$. However, the random variables $\max(A_i + t_i + O_i, A_{i+1} + t_{i+1} + O_{i+1})$ and $\min(A_{i+2} + t_{i+2}, A_{i+3} + t_{i+3})$ are not normal and, moreover, it may even be inaccurate to approximate them as normal for the purpose of this analysis. It should be emphasized that, contrary to the analysis in section 2.1.3, here we are concerned not only for the accuracy in approximating the mean, but the entire probability distribution. As demonstrated in (Daganzo, Bouthelier, & Sheffi, 1977), the accuracy of the approximation is satisfactory when the means of the random variables entering the $\max$ (or $\min$) operator are different, and the ratio of their standard deviations is as high as two. Nadarajah and Kotz (2008b) arrive at similar conclusions. In our application, the random variables $A_i + t_i + O_i$ and $A_{i+1} + t_{i+1} + O_{i+1}$ have different means, while the ratio of their standard deviations does not exceed two. Similar arguments hold for the terms entering the $\min$ operator, and therefore we will approximate $Z = \max(A_i + t_i + O_i, A_{i+1} + t_{i+1} + O_{i+1})$ and $V = \min(A_{i+2} + t_{i+2}, A_{i+3} + t_{i+3})$ as normal random variables. The mean and variance of $Z$ are provided by equations (2.2) and (2.3). Also, let $X$ and $Y$ be normally distributed random variables, $X \sim N(\mu_X, \sigma_X)$ and $Y \sim N(\mu_Y, \sigma_Y)$, $\rho$ represent the correlation coefficient between $X$ and $Y$, and $W$ be the minimum of $X$ and $Y$, $W = \min(X, Y)$. The mean $\mu_W$ and variance $\sigma^2_W$ of $W$ are then (see Nadarajah & Kotz, 2008b):

$$\mu_W = \mu_X \Phi(-\alpha) + \mu_Y \Phi(\alpha) - \gamma \phi(-\alpha)$$

$$\sigma^2_W = \left(\sigma^2_X + \mu^2_X\right) \Phi(-\alpha) + \left(\sigma^2_Y + \mu^2_Y\right) \Phi(\alpha) - \left(\mu_X + \mu_Y\right) \gamma \phi(-\alpha) - \mu^2_Z$$

where $\alpha = (\mu_X - \mu_Y) / \gamma$ and $\gamma = \left(\sigma^2_X + \sigma^2_Y - 2 \rho \sigma_X \sigma_Y\right)^{1/2}$. Assuming that $Z$ and $V$ are independent, their difference follows a normal distribution with mean $\mu_Z - \mu_V$ and variance $\sigma^2_Z + \sigma^2_V$. Thus, (4.1) can be written as $H = \{x > 0 : P(Z-V > -\tau_x) \leq \rho\}$. Therefore, through use of equations (2.2)-(2.3) and (4.1)-(4.5), one can obtain estimates of the scheduled headway $H$, provided that the ROT distributions for arrivals are normal.

### 4.3 Numerical example

Next, we are interested in obtaining insight into the effect of precision on the scheduled headway $H$ between consecutive pairs of arrivals. We proceed by examining a simplified
case, in which all model parameters do not vary across aircraft. Thus, a probability level of \( p = 0.1 \) is allowed for delaying departures to avoid simultaneous runway occupancy, and a level of \( p' = 0.1 \) for not achieving the desired minimum separation at the merging point, \( b = 10 \) seconds. Also, all landing aircraft have a mean runway occupancy time of 30 seconds and standard deviation of 5 seconds, while controllers assume \( \tau_i = 40 \) seconds. Since the time from merging point to runway threshold has been assumed identical for all aircraft, it cancels out from the calculations. For a range of standard deviations \( \sigma \) that are representative of aircraft precision in executing 4D trajectories in NextGen, namely for \( \sigma = 1 \) to \( \sigma = 15 \) seconds, Table 4.1 presents the estimates for scheduled headway between arrival pairs \( H \), scheduled headway between aircraft of the same pair \( F \), and the rate of landings \( L \). Furthermore, Fig. 4.2 shows the relationship between \( \sigma \) and the rate of landings.

The results in Table 4.1 suggest the potential from reducing the aircraft trajectory imprecision level \( \sigma \). To illustrate that we compare the cases when \( \sigma = 13 \) and \( \sigma = 8 \) seconds. The former case may well capture current conditions at SFO, since the associated rate of landings is 56 aircraft/hour and the headway between aircraft of the same pair is 34 seconds, which translates to a separation of approximately one nautical mile, typically observed in aircraft at a staggered landing arrangement. With a 5 second reduction in imprecision, a drop of 18 seconds in the headway \( H \) between arrival pairs is achieved. In turn, that increases landing throughput to 65 aircraft/hour, a 16% increase from today’s rate.

The graph in Fig. 4.2 indicates an almost linear relationship between precision \( \sigma \) and the rate of landings. A linear regression model was estimated using adherence error’s standard deviation \( \sigma \) as explanatory variable, which is summarized through the expression \( L = -2.0039\sigma + 81.547 \) (with \( R^2 = 0.99231 \)). Therefore, reducing \( \sigma \) by one second yields an increase in the rate of landings of two aircraft per hour. For super-density arrival operations in NextGen that require high levels of trajectory precision (JPDO, 2010), we assume a level of \( \sigma = 5 \) seconds. That results to a landing rate of 71 aircraft/hour, which translates to a 29% increase in landing throughput. Finally, almost perfect adherence to 4DT’s, corresponding to a situation when \( \sigma = 1 \) second, yields an upper limit of 81 landings per hour, or 47% higher than the current level of 55 aircraft per hour at SFO. The latter represents an environment where air traffic controllers assign maneuvers to aircraft to achieve paired landings, whereas the former an environment where aircraft self-separate and merge, while flying their preferred trajectory to the runway. It should be noted that even under 81 landings per hour, the separation headway of 89 seconds is reasonable for aircraft landings, as discussed in related studies (Hansen, 2002).
Table 4.1 Results of numerical example

<table>
<thead>
<tr>
<th>$\sigma$ (s)</th>
<th>$F$ (s)</th>
<th>$H$ (s)</th>
<th>Landings (aircraft/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>89</td>
<td>81</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>92</td>
<td>78</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>95</td>
<td>76</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>98</td>
<td>73</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>102</td>
<td>71</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>104</td>
<td>69</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>108</td>
<td>67</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>111</td>
<td>65</td>
</tr>
<tr>
<td>9</td>
<td>27</td>
<td>115</td>
<td>63</td>
</tr>
<tr>
<td>10</td>
<td>29</td>
<td>119</td>
<td>61</td>
</tr>
<tr>
<td>11</td>
<td>31</td>
<td>122</td>
<td>59</td>
</tr>
<tr>
<td>12</td>
<td>32</td>
<td>125</td>
<td>58</td>
</tr>
<tr>
<td>13</td>
<td>34</td>
<td>129</td>
<td>56</td>
</tr>
<tr>
<td>14</td>
<td>36</td>
<td>133</td>
<td>54</td>
</tr>
<tr>
<td>15</td>
<td>38</td>
<td>136</td>
<td>53</td>
</tr>
</tbody>
</table>

Figure 4.2. Rate of landings as a function of precision in executing 4D trajectories
5. Conclusions

In this dissertation a queueing system with a single server under 4D trajectory-based aircraft operations is examined. Aircraft are assigned scheduled times of arrival at a fix, which they meet with some stochastic error. We began our analysis by assuming a normal distribution for the error, and that aircraft enter service according to a First-Scheduled-First-Served (FSFS) queue discipline. We developed a recursive queueing model that employs the Clark approximation to analytically estimate the mean and variance of aircraft delays. The accuracy of the approximation method was assessed through simulation experiments, which indicated good accuracy of the Clark method in estimating total system delays. Using a wide range of representative demand and capacity scenarios at seven major US airports, we compared the model estimates for average queueing delay per flight with those from a deterministic queueing model. It was found that the estimates of expected queueing delay from the stochastic and the deterministic model are strongly correlated and very similar, except for cases of low airport utilization, characterized by average deterministic queueing delay over all arrivals smaller than one minute.

Next, we studied a simplified situation in which a sequence of aircraft with the same 4DT execution accuracy are assigned scheduled times of arrival at a fix with constant excess time separation between them. We found that, under these assumptions, the expected delay to a flight from imperfect trajectory adherence – which we term stochastic delay – depends on the excess time separation, or buffer, expressed as a ratio to the trajectory imprecision, as well as the place of the flight in the sequence. As the buffer goes up, the stochastic delay goes down, but at the cost of increased deterministic delay from reduced capacity. If stochastic delay costs more than deterministic delay, then the optimum buffer is greater than zero, but quite small under plausible cost ratios and trajectory precision levels.
We also explored the effect of queue disciplines that give priority to aircraft equipped with avionics that enable them to execute 4D trajectories with high precision. It was found that by switching from a FSFS to a Best-Equipped-Best-Served (BEBS) policy, total delay in the system can remain at the same level, while achieving significant delay savings for equipped aircraft. Moreover, a BEBS policy that allows only two position shifts can still yield substantial delay savings for equipped aircraft, while minimizing the additional delays imposed on non-equipped aircraft.

The basic queueing model was extended in chapter 3 to analyze a system for aircraft landings at a single runway under 4D trajectory-based operations. The server of this queueing system is the runway threshold, at which aircraft are assigned scheduled times of arrival. In accord with evidence from a variety of sources, we assumed a Gumbel distribution for aircraft’s stochastic lateness and for runway occupancy times (RT). We employed a recursive queueing model and proposed an approximation method to analytically estimate the mean and variance of aircraft delays. The accuracy of the approximation method was demonstrated through simulation experiments.

Similar to the analysis in chapter 2, we studied a simplified situation, in which a sequence of aircraft with the same capability of adherence to 4D trajectories, minimum headway, and runway occupancy time distribution are metered at a constant rate. We investigated the relationship between buffer time between scheduled arrivals and expected loss in system efficiency, defined as the weighted sum of delays due to imperfect trajectory adherence and varying RT – which we term stochastic delay – and due to reduced throughput. It was found that if stochastic delay costs more than deterministic delay, then the minimum expected loss in efficiency is attained for buffer values greater than zero. Under conditions that we consider realistic, such a buffer would reduce planned throughput about 15% compared to the maximum possible under deterministic conditions. It was also shown that when RT is the binding factor that determines throughput, stochastic delays are substantially higher, compared to a situation where minimum headway is the principal constraint. For the former cases, it was found that the marginal benefit from improving trajectory precision diminishes when the standard deviation of adherence error is less than 0.5 times that of RT. Therefore, when runway throughput is controlled primarily by aircraft’s time to exit the runway, improving 4D trajectory precision yields delay savings, but at a decreasing rate.

In order to investigate the potential increase in runway throughput from implementation of 4D trajectories, we studied the case of paired arrivals at the San Francisco International Airport. A mathematical model was developed that estimates the scheduling headway between two consecutive pairs of landing aircraft. This headway minimizes the time interval between consecutive arrival pairs, while allowing sufficient time for a pair of departing aircraft to take-off in the meantime. Results derived from a simplified case study indicate potential increases in landing throughput by two aircraft per hour for every second of decrease in the standard deviation of adherence error. Additionally, we computed the upper bound in landings brought by almost perfect adherence to 4DT’s, which was as high as 81 aircraft/hour, an increase of 47% from current level of 55 aircraft/hour.
The effect from reducing aircraft’s imprecision in adhering to assigned times of arrival is twofold. In a direct way, it reduces stochastic queueing delays, as discussed in chapter 2. Moreover, it results in larger relative buffers $\Delta$ between scheduled arrivals, which in turn creates an opportunity for increasing the rate of metering aircraft at the server, and thus reducing deterministic delay. The queueing model developed in this dissertation enables us to conduct quantitative assessment of the above improvements.

For situations where aircraft are assigned times of arrival at the runway threshold, trajectory imprecision is of secondary importance to system delay as compared to variability in runway occupancy time, when RT is the binding factor that determines throughput. Reducing trajectory imprecision reduces system delay, but at a diminishing rate when the ratio of standard deviations of trajectory adherence and RT is smaller than 0.5. Most important, our analysis in chapter 3 indicates that reducing variance in RT per one unit yields higher savings in stochastic delay than a per unit reduction in trajectory precision error’s variance.

This dissertation focused on a situation where aircraft are served in the order in which they were scheduled to arrive. However, it might well be the case that a stream of aircraft flying 4D trajectories includes a few aircraft without such capability. Due to imprecise trajectory control of the non-equipped aircraft, those flights might deviate significantly from their scheduled time of arrival, and continuing to serve aircraft in a first-scheduled first-served order may cause prohibitive delays. In this case, serving flights on a first-come first-served basis would result in significantly lower delays. Under such a queue discipline, the realized order of aircraft arrivals at the server is different than the initially scheduled, and therefore one has to estimate the average arrival time of the first, second, etc aircraft that will actually arrive at the server. Future research will attempt to tackle this problem by employing tools from the field of ordered statistics.

Moreover, the analysis on the effect of buffer on system performance, presented in sections 2.2–2.3 and 3.2, assumes a simple shape of the aircraft demand curve. Future work will attempt to explore cases where the slope of the demand curve decreases after a certain time moment, but does not reach zero.

Finally, chapter 3 focuses on a situation where all aircraft have the same capability of flying 4D trajectories. However, it might well be the case that two distinct classes of aircraft, with regards to their trajectory adherence capability, are scheduled to arrive at an airport with two runways. Investigating the resource allocation and efficiency of such a system will be the focus of future research.
Bibliography


