Data-Driven Modeling of the Airport Runway Configuration Selection Process Using Maximum Likelihood Discrete-Choice Models

by

Jacob Bryan Avery

B.S., University of Illinois at Urbana Champaign (2013)

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Master of Science in Aeronautics and Astronautics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2016

Signature redacted

Author .....
Department of Aeronautics and Astronautics
January 15th, 2016

Signature redacted

Certified by .....
Hamsa Balakrishnan
Associate Professor of Aeronautics and Astronautics
Thesis Supervisor

Signature redacted

Accepted by .....
Paulo C. Lozano
Associate Professor of Aeronautics and Astronautics
Chair, Graduate Program Committee

Data-Driven Modeling of the Airport Runway Configuration Selection Process Using Maximum Likelihood Discrete-Choice Models

by

Jacob Bryan Avery

Submitted to the Department of Aeronautics and Astronautics
on January 15th, 2016, in partial fulfillment of the
requirements for the degree of
Master of Science in Aeronautics and Astronautics

Abstract

The runway configuration is a key driver of airport capacity at any time. Several factors, such as wind speed, wind direction, visibility, traffic demand, air traffic controller workload, and the coordination of flows with neighboring airports influence the selection of the runway configuration.

This paper identifies a discrete-choice model of the configuration selection process from empirical data. The model reflects the importance of various factors in terms of a utility function. Given the weather, traffic demand and the current runway configuration, the model provides a probabilistic forecast of the runway configuration at the next 15-minute interval. This prediction is then extended to obtain the probabilistic forecast of runway configuration on time horizons up to 6 hours.

Case studies for Newark (EWR), John F. Kennedy (JFK), LaGuardia (LGA), and San-Francisco (SFO) airports are completed with this approach, first by assuming perfect knowledge of future weather and demand, and then using the Terminal Aerodrome Forecasts (TAFs). The results show that given the actual traffic demand and weather conditions 3 hours in advance, the models predict the correct runway configuration at EWR, JFK, LGA, and SFO with accuracies 79.5%, 63.8%, 81.3% and 82.8% respectively. Given the forecast weather and scheduled demand 3 hours in advance, the models predict the correct runway configuration at EWR, LGA, and SFO with accuracies 78.9%, 78.9% and 80.8% respectively. Finally, the discrete-choice method is applied to the entire New York Metroplex using two different methodologies and is shown to predict the Metroplex configuration with accuracies of 69.0% on a 3 hour prediction horizon.

Thesis Supervisor: Hamsa Balakrishnan
Title: Associate Professor of Aeronautics and Astronautics

Acknowledgments

First and foremost, I would like to thank my advisor, Professor Hamsa Balakrishnan, for her support and direction while completing this research and my graduate degree. I feel incredibly fortunate to have worked with an advisor who encouraged me to question my thoughts, motivated me to test new hypotheses, inspired me think creatively about a problem, and pointed me in the right direction when I faltered. Hamsa's leadership instilled in me the ability to think about challenging problems as a true engineer, and for that I am eternally grateful.

Additionally, I would like to thank my fellow ICAT students and friends, Yashovardhan Chati, Sandeep Badrinath, Karthik Gopalakrishnan, Zebulon Hanley, Michael Kasperski, Patrick McFarlane, and Daniel Schonfeld, not only for listening to my ramblings on everything from Star Wars to Neural Networks, but for helping me sort through some of my most difficult research challenges.

I also benefited greatly from conversations with Monica Alcabin (Boeing), Dan Bueno (FAA), Richard DeLaura (MIT Lincoln Labs), Larry Goldstein (TRB), Belinda Hargrove (TransSolutions), Renee Hendrics (FAA), Richard Jordan (MIT Lincoln Labs), Leo Prusak (FAA), and Varun Ramanujam (MIT) during the course of this research. Their insight, suggestions, and guidance was invaluable when developing the proposed models. I would also like to acknowledge the assistance of Mr. Andrea Pergola in providing archived TAF data, and Graham Davis (MIT) in the processing of the TAF data.

I would like to thank my family for their words of encouragement, financial support, and unwavering belief in me. Finally, and most importantly, I would like to thank my fiancé and best friend for her patience, love, and constant support over the past several years. She kept me sane during the long nights and the hard times. Without her, none of this would have been possible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Airport congestion leads to significant flight delays at the busiest airports around the world. Fundamentally, congestion is caused by an imbalance between demand (airport operations) and supply (airport capacity) within the air transportation system. Airport expansion projects can increase the runway capacity at an airport, but are expensive and take many years to complete; by contrast, the better utilization of existing airport capacity is a less expensive approach to mitigating airport congestion. The key driver of airport capacity at a given time is the active runway configuration [1], which is the combination of runways being used to handle the arrival and departure flows at the airport under consideration.

When selecting a runway configuration, air traffic control personnel must consider meteorological and operational factors such as wind speed, wind direction, arrival demand, departure demand, noise mitigation, and inter-airport coordination. A comprehensive understanding of the runway configuration selection process by air traffic personnel is necessary for the future development of decision support tools under SESAR and NextGen initiatives. A keen understanding of this process has the potential to increase the operational efficiency of airport capacity utilization and provides a key step toward airport capacity prediction. Airport capacity predictions are important inputs needed for air traffic flow management [2, 3], airport surface operations scheduling [4], and system-wide simulations [5]. Since the capacity of an airport depends heavily on the runway configuration being used, the forecast of the

runway configuration is a key step toward predicting the capacity of an airport.

This paper develops a data-driven model of the runway configuration selection process using a discrete-choice modeling framework for Newark (EWR), John F. Kennedy (JFK), LaGuardia (LGA), and San-Francisco (SFO) airports. It also extends this discrete choice approach to model the configurations of the entire New York Metroplex which includes EWR, LGA, and John F. Kennedy (JFK) airports. The models infer the utility functions that best explain (that is, maximize the likelihood of) the observed decisions. The utility functions give insight on the relative importance of the different decision factors to air-traffic control personnel when selecting a runway configuration. The resultant model yields a probabilistic prediction of the runway configuration at any time, given a forecast of the influencing factors.

1.1 Related Work

1.1.1 Prescriptive Models

Two types of models have previously been developed for the runway configuration selection problem: prescriptive and descriptive models. Prescriptive models account for the weather and other operational constraints to recommend an optimal runway configuration. An early example of a prescriptive model is the Enhanced Preferential Runway Advisory System (ENPRAS) that was developed for Boston Logan International Airport (BOS) [6]. The ENPRAS was created to mitigate noise impacts from aircraft arrivals and departures at BOS and provide noise relief to nearby communities by selecting optimal runway configurations (usually over water). The optimal configuration was determined considering weather, demand, and runway conditions. Both long-term aggregate noise pollution impacts and short-term impacts to nearby communities were considered when developing the system. So far, the ENPRAS has successfully helped operators at BOS lower the noise levels to surrounding communities, and future implementations will help airport operators plan for optimal runway use over their entire shift period.

Also motivated by aircraft noise considerations, runway allocation systems were designed for Sydney and Brisbane airports [7]. These systems allowed for a more robust method for detecting aircraft noise profiles by using time-stamped aircraft movement composite-year data sets instead of the conventionally used day-average movement data sets. These new composite-year data sets were then tested within runway allocation models and was shown to predict the runway allocation at a higher level of confidence than previously used methods.

More recently, several authors have considered the problem of optimally scheduling runway configurations, taking into account different models of weather forecasts and the loss of capacity during configuration switches [8, 9, 10, 11, 12]. Airports are assumed to operate under capacity envelopes that govern the amount of arrivals and departures that an be handled at an airport. These capacity envelopes directly depend on the runway configuration being used. High-level approaches using these concepts have attempted to describe the benefits to the entire National Airspace System (NAS) when operating in a certain runway configuration. More recent approaches drill deeper and examine airports individually, focusing on their unique complexities and attributes to predict future runway configuration changes [12]. Additionally, models were also developed to determine the optimal relative sequencing of runway configurations by applying mixed integer programming models. These models were developed to reach an optimal balance of arriving aircraft and departing aircraft at an airport over time [11]. Other models dealt with capacity loss at an airport during a runway configuration switch, assigning transition penalties to mimic real-world operating conditions during a runway configuration switch [10].

1.1.2 Descriptive Models

Descriptive models use data-mining approaches to predict the runway configuration selection based on historical data. These models describe the decision selection processes of decision makers and use those processes to make predictions rather than simply recommending an optimal runway configuration. Descriptive models have received less attention than prescriptive models, but recently research in this field has

grown significantly. An example of a descriptive modeling approach uses data-mining methods to forecast Airport Arrival Rates (AARs) using Terminal Aerodrome Forecast (TAF) data. The TAF data is used to develop capacity profiles for airports using different stochastic approaches such as k-means clustering and dynamic time warping. A design of experiments methodology was taken by assigning the cost of delay to an objective function. The stochastic capacity profiles forecast AARs by minimizing the cost of delay objective function using the real-time data available to airport operators when making decisions. This methodology also makes the approach beneficial for Ground Delay Program (GDP) planning [14].

A 24-hour forecast of runway configuration was developed for Amsterdam Schiphol airport using a probabilistic weather forecast [15]. The method used modified two-dimensional Gaussian distribution with wind speed and wind direction as degrees of freedom to develop probabilities of selecting a runway configuration. The intent of this model is to provide airport operators with decision support tools based on future weather forecasts. The model can also keep nearby civilians who live close to the airport informed of the likely impacts that weather will have on airport operations - and in turn the noise over their communities. In some cases, these models achieve accuracies of up to 70%.

A logistic regression based approach was used to develop a descriptive model of runway configuration selection at LGA and JFK, although this was not a predictive model [16]. These models have been shown to have accuracies of up to 75%, however, a difficult task has been modeling the observed resistance to configuration changes from air traffic control personnel. Recent research using discrete-choice models of the runway configuration selection process for year 2006 at EWR and LGA airports have taken this observation into account with similar accuracies [17, 18]. The model parameters in these discrete-choice models were set using standard operating procedures at EWR and LGA, which are subject to change and can sometimes disagree with the data.

1.1.3 Extension of Previous Work

This paper will extend the aforementioned discrete-choice models at EWR and LGA for years 2011-2012 with a more data-driven approach. The methodology will also be applied to SFO and JFK for years 2011-2012. Finally, the discrete-choice approach will be applied to the entire New York Metroplex for years 2011-2012. A key novelty in this paper is that the constraints pertaining to the maximum allowable tailwinds and crosswinds are learned from the actual data rather than the FAA operating manuals.

The utility functions within the discrete-choice framework capture the importance of wind speed and direction, air traffic demand, noise abatement procedures, and the coordination of flows with neighboring airports. An advantage of the discrete-choice approach is its ability to account for the resistance to configuration switches by air-traffic control, called operational "inertia" in this paper. Switching a runway configuration requires increased coordination among airport stakeholders which lowers airport throughput, and consequently makes air traffic control personnel to resist frequent configuration changes. While the influence of inertia is arguably less important on long forecast horizons (when the key factors are likely to be wind conditions, visibility, and demand), the resistance to configuration changes play a much more prominent role on short forecast horizons (such as 3-hours ahead). Without accounting for inertia, tools that suggest possible choices for the optimal runway configuration recommend significantly more frequent changes than were seen in actual operations [12]. The discrete-choice modeling framework helps to accommodate the effect of inertia, in addition to the other influencing factors.

This paper illustrates the proposed approach using case studies for EWR, JFK, LGA, SFO, and the New York Metroplex, first assuming a knowledge of the actual weather conditions and the traffic demand 3 hours ahead, and then using the most recent Terminal Aerodrome Forecast (TAF) available 3 hours in advance.

1.2 Airports Considered

1.2.1 EWR

Newark Liberty Internal Airport is one of three major airports in the New York Metroplex. EWR handles approximately 35 million passengers a year on three runways: 4L/22R, 4R/22L, and 11/29 [19]. EWR serves as a hub for United Airlines, which handles approximately 70% of its passenger traffic [19].

An airport layout of EWR is shown in Figure 1.2.1. Typically, runway 4L/22R is used for departures and 4R/22L is used to handle arrivals. Runway 11/29 is not usually preferred because it is not capable of instrument landing approaches, however during very strong crosswinds runway 11/29 may be used to handle either arrivals or departures. 25 different runway configurations at EWR were reported in year 2011. Table 1.2.1 shows the frequencies with which the most commonly-used configurations at EWR were observed.

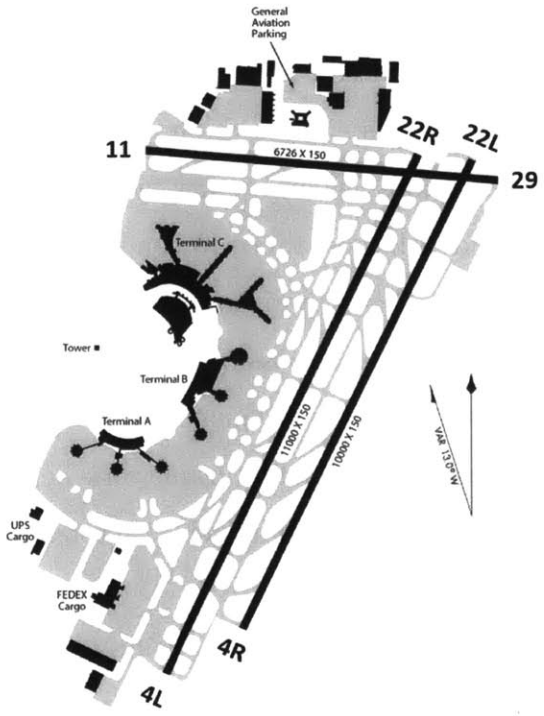

Figure 1.2.1: Layout of EWR airport.

1.2.2 JFK

John F. Kennedy International Airport is another of the three major airports in the New York Metroplex. JFK handles approximately 55 million passengers a year and is operationally the largest international airport in the United States [20]. The airport serves as a hub for American Airlines, Delta Airlines, and JetBlue.

JFK has four runways: 13R/31L, 4R/22L, 4L/22R, and 13L/31R [21]. An airport layout of JFK is shown in Figure 1.2.2. Commonly, arrivals are handled using runways 31L/13R or 4L/22R depending on the specific conditions, and departures are typically handled with 13R/31L. 43 different runway configurations at JFK were reported in 2011. Table 1.2.1 shows the frequencies with which the most commonly-used configurations at JFK were observed.

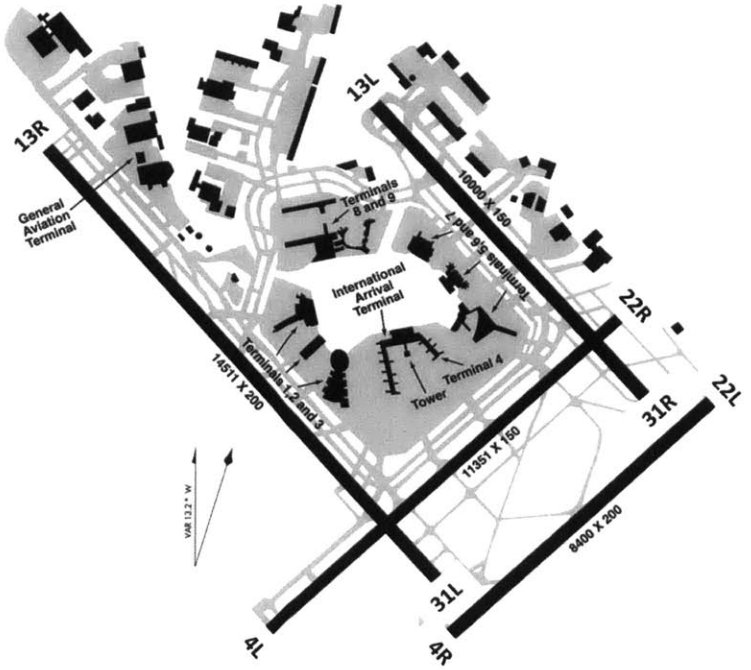

Figure 1.2.2: Layout of JFK airport.

1.2.3 LGA

LaGuardia Airport is the final major airport in the New York Metroplex. LGA handles approximately 25 million passengers a year on two runways: 4/22 and 13/31

[22]. LaGuardia acts as a hub for Delta Airlines.

An airport layout of LGA is shown in Figure 1.2.3. The active runways for arrivals and departures depends on the weather conditions, but typically, arrivals at LGA are handled on runway 4/22 and departures are handled on runway 13/31. 27 different runway configurations at LGA were reported in 2011. Table 1.2.1 shows the frequencies with which the most commonly-used configurations at LGA were observed.

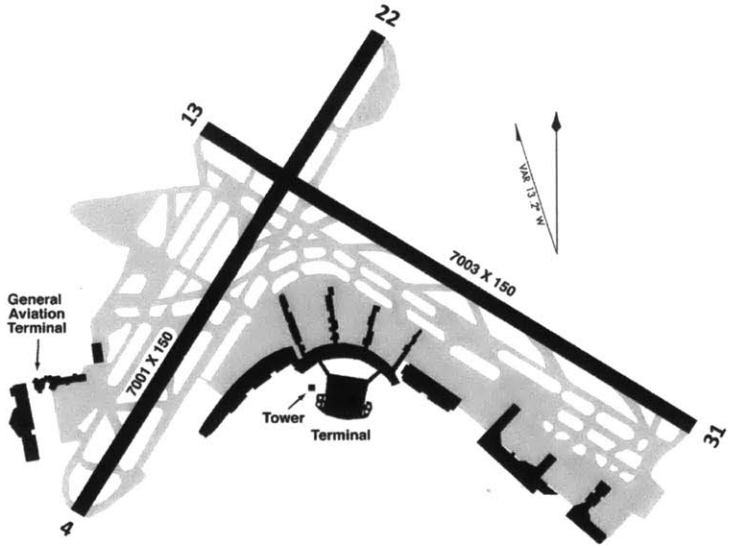

Figure 1.2.3: Layout of LGA airport.

1.2.4 SFO

San Francisco International Airport is the major airport in the San Francisco Bay Area, handling approximately 47 million passengers annually [23]. SFO acts as a hub for United Airlines and Virgin America.

As shown in Figure 1.2.4, SFO has four runways: 10L/28R, 10R/28L, 01L/19R, and 01R/19L [24]. Handling arrivals and departures at SFO is uniquely challenging for air traffic control because the centerlines of runways 10L/28R and 10R/28L are only separated by 750 feet. According to FAA regulations parallel approaches cannot be allowed during poor weather conditions at such a small separation. This effectively lowers the capacity of the airport during overcast weather, and will be shown

to present a significant modeling challenge in this paper. 24 different runway configurations at SFO were reported in 2011. Table 1.2.1 shows the frequencies with which the most commonly-used configurations at SFO were observed.

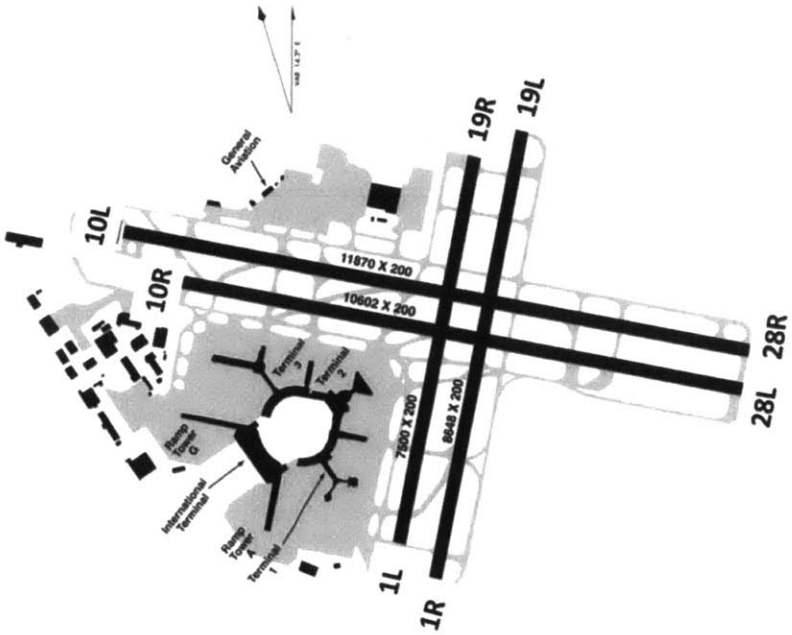

Figure 1.2.4: Layout of SFO airport.

1.2.5 New York Metroplex

The New York Metroplex is comprised of three major airports within a relatively close proximity to another - EWR, JFK, and LGA. A map layout of the New York Metroplex is shown in Figure 1.2.5 In come cases, the large general airport TEB is included, but it will be left out for the purposes of this research. The New York Metroplex is a very large bottleneck on the NAS because it has the largest number of arrival and departure operations in the United States [35, 36]. Consequently, it is the most congested airspace in the NAS [34]. Delays that occur in the New York Metroplex propagate throughout the rest of the system and can have a heavy impact on the overall delay state of the NAS.

Prior research has suggested that the New York Metroplex is beginning to reach its maximum airspace capacity given the current operational landscape and physical

Table 1.2.1: Frequent configurations observed at EWR, JFK, LGA, and SFO in year 2011.

Airport	Configuration	Frequency	% Frequency
EWR	21L,11|22R	4,214	13%
	22L|22R	16,559	50%
	22L|22R,29	353	1%
	4L,4R|4L	528	2%
	4R,11|4L	1,576	5%
	4R|4L	10,221	31%
JFK	13L,22L|13R	2,174	6%
	13L|13R	1,395	4%
	22L,22R|22R	2,579	7%
	22L,22R|22R,31L	1,785	5%
	22L|22R	3,528	10%
	22L|22R,31L	4,085	12%
	31L,31R|31L	7,064	20%
	31R|31L	3,233	9%
	4L,4R|4L	1,927	5%
	4L,4R|4L,31L	1,987	6%
	4R|4L	2,592	7%
	4R|4L,31L	1,835	5%
LGA	22|13	6,846	24%
	22|31	5,556	19%
	22,31|31	852	3%
	31|31	2,676	9%
	31|4	7,608	26%
	4|13	4,113	14%
	4|4	1,372	5%
SFO	19R,19L|10R,10L	957	3%
	28R,28L|01R,01L	24,871	74%
	28R,28L|28R,28L	3,000	9%
	28R|01R,01L	467	1%
	28L|01R,01L	4,244	13%

layout of EWR, JFK, and LGA [32]. Because of the large expected increase in air travel demand over the next few decades, a major component of NextGen research has been devoted to examining possible capacity enhancements for the New York Metroplex. This paper will also model the New York Metroplex using a discrete choice approach. The utility functions learned from these models could help future models with capacity predictions or defining objective functions for decision support

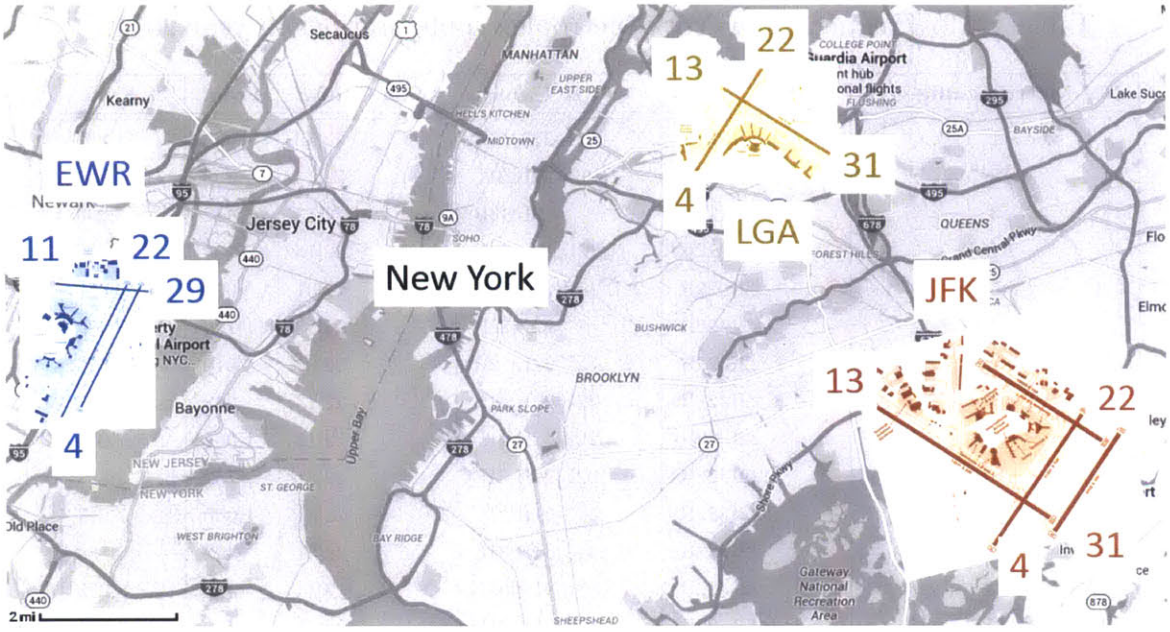

Figure 1.2.5: Layout of EWR, JFK, and LGA in the New York Metroplex.

tools.

Just as the previously mentioned airports were shown to have runway configurations, the New York Metroplex is considered to have certain "configurations" as well. The New York Metroplex runway configurations are simply the combined runway configurations of EWR, JFK, and LGA at any given time. By the nature of this definition, the New York Metroplex configurations switch much more often and have a lower number of occurrences compared to the configurations at each individual airport. Table 1.2.2 shows the most frequent (filtered at 1%) configurations seen in the New York Metroplex during 2011.

As shown, in Table 1.2.2, many configurations seem similar to others and typically occur less frequently than most of the individual airport configurations. For these reasons the New York Metroplex configurations are difficult to model. To help manage these difficulties, prior research used the configurations at each individual airport to develop overarching configurations for the New York Metroplex [34]. This paper will draw from that work and define similar overarching runway configurations using the frequent configurations shown in Table 1.2.2. The overarching New York Metroplex runway configurations follow eight structures

Table 1.2.2: Frequent New York Metroplex configurations in year 2011.

NY Metro Config. ID	EWR Config.	JFK Config.	LGA Config.	Freq.	% Freq.
1	11,22L|22R	13L,22L|13R	22|13	565	1.6%
2	11,22L|22R	22L|22R,31L	22|13	358	2.6%
3	22L|22R	13L|13R	22|13	442	3.2%
4	22L|22R	22L,22R|22R	22|13	884	6.3%
5	22L|22R	22L,22R|22R,31L	22|31	628	4.5%
6	22L|22R	22L|22R	22|13	1,149	8.2%
7	22L|22R	22L|22R	22|31	516	3.7%
8	22L|22R	22L|22R,31L	13,22|13	368	2.6%
9	22L|22R	22L|22R,31L	22|13	1,061	7.6%
10	22L|22R	22L|22R,31L	22|31	739	5.3%
11	22L|22R	31L,31R|31L	22|31	1,060	7.6%
12	22L|22R	31L,31R|31L	31|31	392	2.8%
13	22L|22R	31L,31R|31L	31|4	416	3.0%
14	22L|22R	31R|31L	22|31	356	2.5%
15	4R,11|4L	31L,31R|31L	31|4	518	3.7%
16	4R|4L	31L,31R|31L	31|4	1,273	9.1%
17	4R|4L	31R|31L	31|4	580	4.2%
18	4R|4L	4L,4R|4L	4|13	750	5.4%
19	4R|4L	4R|4L	31|4	425	3.0%
20	4R|4L	4R|4L	4|13	623	4.5%
21	4R|4L	4R|4L,31L	31|4	487	3.5%
22	4R|4L	4R|4L,31L	4|13	371	2.7%

1. Southern aircraft flow under VMC with an arrival priority.
2. Southern aircraft flow under IMC.
3. Southern aircraft flow under VMC with a departure priority.
4. Mixed southern and northern flow with an emphasis on northern flow.
5. Mixed southern and northern flow with an emphasis on southern flow.
6. Northern aircraft flow under VMC with an arrival priority.
7. Northern aircraft flow under IMC.
8. Northern aircraft flow under VMC with a departure priority.

The overarching New York Metroplex configurations and their components are shown in Table 1.2.3.

Table 1.2.3: Combined NY Metro. configurations.

NY Metro Config. ID	EWR Config.	JFK Config.	LGA Config.	Freq.	% Freq.
1	11,22L|22R	13L,22L|13R	22|13	565	1.6%
3	22L|22R	13L|13R	22|13	442	3.2%
South Flow - VMC - Arrival Priority (S-VMC-AP)				**2,794**	**20%**
2	11,22L|22R	22L|22R,31L	22|13	358	2.6%
4	22L|22R	22L,22R|22R	22|13	884	6.3%
6	22L|22R	22L|22R	22|13	1,149	8.2%
8	22L|22R	22L|22R,31L	13,22|13	368	2.6%
9	22L|22R	22L|22R,31L	22|13	1,061	7.6%
South Flow - IMC (S-IMC)				**2,033**	**14.6%**
5	22L|22R	22L,22R|22R,31L	22|31	628	4.5%
7	22L|22R	22L|22R	22|31	516	3.7%
10	22L|22R	22L|22R,31L	22|31	739	5.3%
South Flow - VMC - Departure Priority (S-VMC-DP)				**1,883**	**13.5%**
11	22L|22R	31L,31R|31L	22|31	1,060	7.6%
14	22L|22R	31R|31L	22|31	356	2.5%
Mixed North and South Flow with South Emphasis (Mixed S to N)				**1,416**	**10.1%**
12	22L|22R	31L,31R|31L	31|31	392	2.8%
13	22L|22R	31L,31R|31L	31|4	416	3.0%
Mixed North and South Flow with North Emphasis (Mixed N to S)				**808**	**5.8%**
15	4R,11|4L	31L,31R|31L	31|4	518	3.7%
16	4R|4L	31L,31R|31L	31|4	1,273	9.1%
17	4R|4L	31R|31L	31|4	580	4.2%
North Flow - VMC - Arrival Priority (N-VMC-AP)				**2,371**	**17.0%**
18	4R|4L	4L,4R|4L	4|13	750	5.4%
20	4R|4L	4R|4L	4|13	623	4.5%
22	4R|4L	4R|4L,31L	4|13	371	2.7%
North Flow - IMC (N-IMC)				**1,744**	**12.5%**
19	4R|4L	4R|4L	31|4	425	3.0%
21	4R|4L	4R|4L,31L	31|4	487	3.5%
North Flow - VMC - Departure Priority (N-VMC-DP)				**912**	**6.5%**

1.3 Notation

Runway configurations are typically designated in the form of 'A1,A2 | D1, D2' where A1 and A2 are the arrival runways, and D1 and D2 are the departure runways. The numbers for each active runway are reported based on their bearing from magnetic north (in degrees) divided by 10. Pairs of parallel runways are differentiated by 'R' and 'L'. For instance, if Newark, shown in Figure 1.3.1, is operating in runway

configuration 22L|22R, aircraft arrivals are handled on runway 22L which faces 220 degrees from magnetic north and departures are handled on parallel runway 22R which also faces 220 degrees from magnetic north.

Theoretically, an airport with N runways has O(6N) possible configurations, since each runway can be used for arrivals, departures or both, and in either direction. Throughout a year, many different runway configurations are seen, however, typically 5-10 runway configurations are used the majority of the time. In addition, due to the additional coordination required during switches, runway configurations only change 1-3 times per day on average.

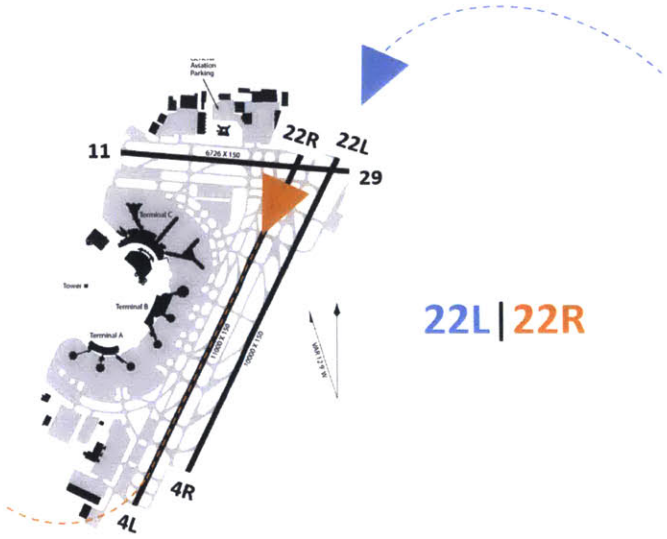

Figure 1.3.1: Example of runway configuration notation.

Chapter 2

Methodology

2.1 Discrete-Choice Modeling Framework

Discrete-choice models are behavioral models that describe the choice selection of a decision maker, or the nominal decision selection among an exhaustive set of possible alternative options, called the choice set [26]. Each alternative in the choice set is assigned a utility function based on defining attributes that are related to the decision selection process. At any given time, the feasible alternative with the maximum utility is assumed to be selected by the decision maker.

The utility function is modeled as stochastic random variable, with an observed (deterministic) component, V, and a stochastic error component, ϵ. For the n^{th} selection, given a set of feasible alternatives C_n, the utility of choice $c_i \in C_n$ is represented as

$$U_{n,i} = V_{n,i} + \epsilon_{n,i}. \tag{2.1.1}$$

The decision maker selects the alternative with maximum utility, that is, $c_j \in C_n$ such that

$$j = \underset{i:c_i \in C_n}{\text{argmax}}(U_{n,i}) \tag{2.1.2}$$

The observable component of the utility function is defined as a linear function of the observed vector of attributes, $\vec{X}_{n,i}$. The attributes include the different factors that can influence the decision. They are weighted by the values in vector, $\vec{\beta}_{n,i}$, and

include alternative specific constants, $\alpha_{n,i}$, as follows:

$$V_{n,i} = \alpha_{n,i} + [\vec{\beta}_{n,i} \cdot \vec{X}_{n,i}]. \tag{2.1.3}$$

The random error component of the utility function reflects all measurement errors, including unobserved attributes, variations between different decision-makers, proxy variable effects, and reporting errors. The error term is assumed to be distributed according to a Type I Extreme Value (or Gumbel) distribution with a location parameter of zero, that is:

$$f(x) = \mu e^{-\mu(x-\eta)} e^{-e^{\mu(x-\eta)}} \tag{2.1.4}$$

where μ is the scale parameter and η is the location parameter. The location parameter is set to zero when defining the discrete choice models. The Gumbel distribution is used to approximate a normal distribution due to its computational advantages. The Multinomial Logit (MNL) model assumes that the error components of each utility function are independent from one another, as shown in Figure 2.1.1.

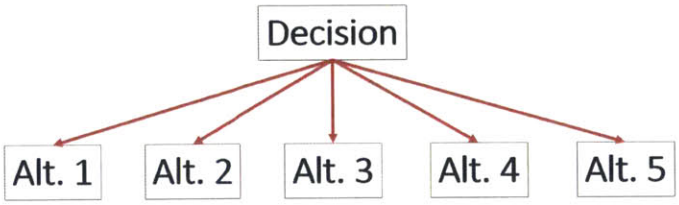

Figure 2.1.1: Example of a MNL model structure.

Under the assumptions of the MNL model, the probability that choice i is chosen during the n^{th} selection is given by

$$P_{n,i} = \frac{e^{V_{n,i}}}{\sum_{j:c_j \in C_n} [e^{V_{n,j}}]}. \tag{2.1.5}$$

The independence among the error terms of each utility function in the MNL model assumes that all correlation among alternatives has been captured by the attributes included in the utility function [26]. The Nested Logit (NL) model relaxes

this assumption by grouping alternatives into subsets, or nests (denoted B_k), which have correlation between their error terms (Figure 2.1.2).

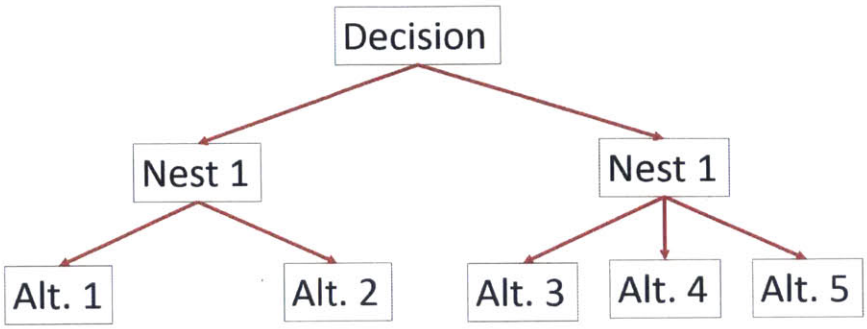

Figure 2.1.2: Example of a NL model structure.

The NL model splits the observable part of the utility function into a component that is common among the alternatives within a nest, and a component that varies between the different alternatives in a nest. The NL model can then be treated as nested MNL models using conditional probabilities. The probability that a specific alternative is chosen is given by the probability that its nest is chosen, multiplied by the probability that the specific alternative is chosen from among the alternatives in that nest. In other words

$$P_n(c_i) = P_n(c_i|B_k)P_n(B_k), \tag{2.1.6}$$

$$\text{where} \quad P_n(B_k) = \frac{\exp(I_{n,k})}{\sum_{l=1}^{K}[\exp(I_{n,l})]} \tag{2.1.7}$$

$$P_n(c_i|B_k) = \frac{\exp(\mu_k V_{n,i})}{\sum_{j\in B_k}[\exp(\mu_k V_{n,j})]} \tag{2.1.8}$$

$$I_{n,k} = \frac{1}{\mu_k}\ln(\sum_{j\in B_k}\exp(\mu_k V_{n,i})). \tag{2.1.9}$$

Equation (2.1.7) has an additional term in the numerator called the inclusive value, that acts as a bridge between the "lower level MNL structures" within each nest, and the "upper level MNL structure" comprised of the nests themselves.

2.2 Maximum-Likelihood Estimation of Model Parameters

Maximum-likelihood estimates of the linear weighting parameters, alternative specific constants, and scale parameters are estimated from the training data. The maximum-likelihood function is defined as the joint probability that the vector of sample data will occur, given a vector of parameters $\vec{\theta} =< \alpha, \vec{\beta}, \mu >$ as follows,

$$\mathcal{L}(\vec{\theta}) = P(\vec{\mathbf{X}}; \vec{\theta}). \tag{2.2.1}$$

The estimated parameters are those that maximize the likelihood of the observations:

$$(\widehat{\alpha}, \widehat{\vec{\beta}}, \widehat{\mu}) =_{\alpha,\vec{\beta},\mu} (\mathcal{L}(\alpha, \vec{\beta}, \mu)). \tag{2.2.2}$$

The resulting nonlinear optimization problem is solved computationally using an open-source software package called BIOGEME [27].

2.3 Statistical Tests

The discrete-choice models for EWR, JFK, LGA, SFO, and the New York Metroplex were realized iteratively, adding or removing variables based on their statistical significance as determined by the Students t-test. The significance of each attribute to the overall model was tested using likelihood ratio testing. Different nested logit model tree structures were also evaluated for statistical significance using likelihood-ratio testing [26].

Chapter 3

Pre-processing The Data

3.1 Training and Testing Data

The training and test datasets for each airport were taken from the FAA's Aviation System Performance Metrics (ASPM) database [28]. The data is reported in 15-minute intervals and includes the active runway configuration, the arrival and departure demand, cloud ceiling, visibility conditions, wind speed, and wind direction. Training data for SFO, LGA, and EWR was taken from year 2011 and the test datasets were taken from year 2012. Results from the prediction models that use ASPM test data assume perfect knowledge of the wind, visibility, and demand for the subsequent three hour interval. Test data sets will also be created using Terminal Aerodrome Forecast (TAF) data and schedule demand data which does not assume precise knowledge of the weather and demand over the next three hours. Prediction results using the TAF and schedule demand data effectively simulate predictions using data that air traffic control personnel would have in real time.

3.2 Attribute Selection

The utility function is assumed to be a linear function of the observed vector of attributes, or factors, that influence the decision. For the runway configuration selection problem, the following attributes were considered.

3.2.1 Inertia

During a runway configuration switch it is necessary to have increased coordination among all stakeholders (airlines, air traffic control, ground operations,etc.), which reduces the airport throughput [11]. Operational "inertia" is a term used to describe the preference of air traffic control to resist runway configuration changes, due to the high operational effort required during the switch. Tools designed to suggest an optimal runway configuration have been shown to recommend significantly more frequent runway configuration switches than is actually observed in practice. An inertia variable was added as an attribute in the models to reflect the preference of air traffic controllers to resist configuration changes. The inertia variable adds a positive contribution to the utility function of the incumbent configuration.

3.2.2 Wind Speed and Wind Direction

Wind speed and direction are key factors that influence the choice of runway configuration. High tailwinds and crosswinds are operationally unsafe in many circumstances, and as a result, render certain runways unusable. The Federal Aviation Administration (FAA) has specified the maximum allowable tailwind and crosswind thresholds for the safe operation of a runway in standard operating procedures (SOPs). Prior work on runway configuration selection based runway availability on SOPs in the model. In this paper, the threshold values of runway tailwinds and crosswinds are directly backed out of the ASPM year 2011 training data sets for each airport.

The ASPM dataset gives both the wind direction with respect to true north, θ and wind speed, v, for each 15-minute interval. Figure 3.2.1 illustrates that the headwind and crosswind components are given by,

$$Headwind = v\cos(\varphi - \theta) \tag{3.2.1}$$

$$Crosswind = v\sin(\varphi - \theta) \tag{3.2.2}$$

where φ denotes the orientation of the runway with respect to true north. Note

that φ does not correspond to the angle implied by the runway number, which is given with respect to magnetic north. The runway angle relative to true north is calculated by adding an offset that depends on the geographic location of the airport, shown in Table 3.2.1. Tailwinds occur when the headwind function takes a negative value.

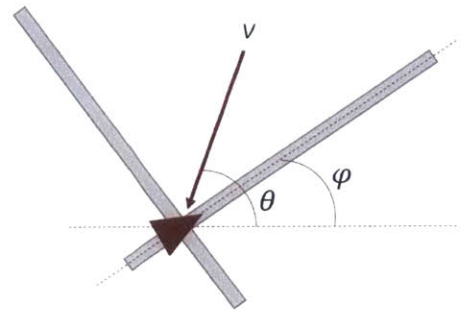

Figure 3.2.1: Determination of the headwind and crosswind components.

Table 3.2.1: Runway angles at EWR, JFK, LGA, and SFO with respect to true north and magnetic north.

Airport	Runway #	Angle TN [deg]	% Angle MN [deg]
EWR (Mag. Variation 13W)	4	26	39
	11	95	108
	22	206	219
	29	275	288
JFK (Mag. Variation 14W)	4	31	45
	13	121	135
	22	211	225
	31	301	315
LGA (Mag. Variation 12W)	4	32	44
	13	122	134
	22	212	224
	31	302	314
SFO (Mag. Variation 14E)	01	27	13
	10	117	103
	19	207	193
	28	297	283

Figures 3.2.2, 3.2.3, 3.2.4, 3.2.5 show the identified ranges of feasible wind speed and wind direction for each runway at EWR, JFK, LGA, and SFO respectively. These figures also show the observed wind speed and wind direction combinations for

year 2011 at these airports. The identified ranges of tailwind and crosswind values for runway feasibility were learned directly from the 2011 ASPM data and are also shown on the plot. Note that air traffic controllers prefer headwinds to tailwinds and crosswinds, and while there is no headwind threshold for runway feasibility, headwind thresholds were plotted at 40 knots in the diagrams for better illustration. The tailwind and crosswind limits were taken on a "per-runway" basis, and were calculated via the following procedure.

1. Aggregate all tailwind and crosswind values for each runway at EWR, JFK, LGA, and SFO.

2. For each runway list, remove tailwind and crosswind combinations from time periods when the active configurations did not include any operations on the runway.

3. Take tailwind thresholds at 90th percentile and crosswind thresholds at the 95th percentile to remove possible reporting errors. In the plots, the solid line corresponds to the aforementioned thresholds for each runway, and the dashed lines correspond to the 75th percentile.

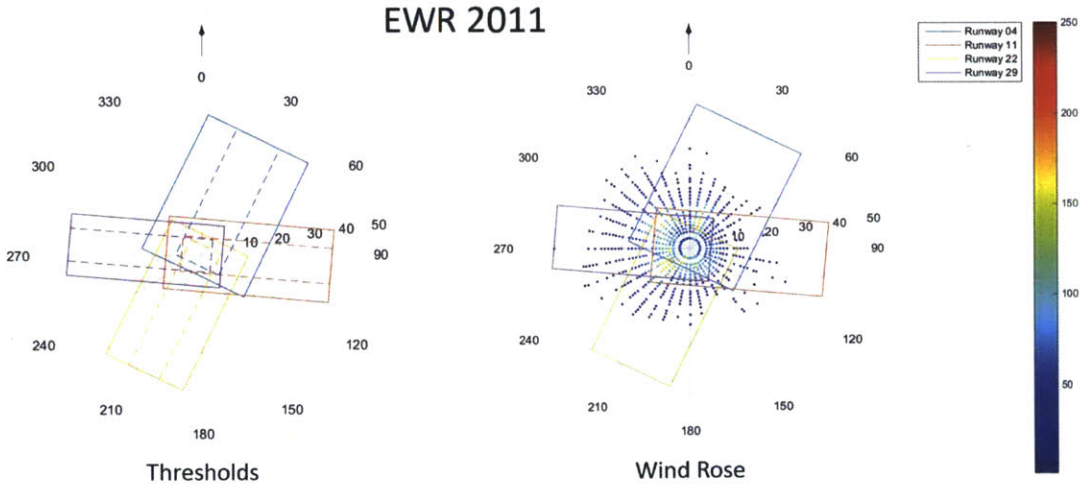

Figure 3.2.2: Wind speed, wind direction, and wind thresholds learned from year 2011 EWR data.

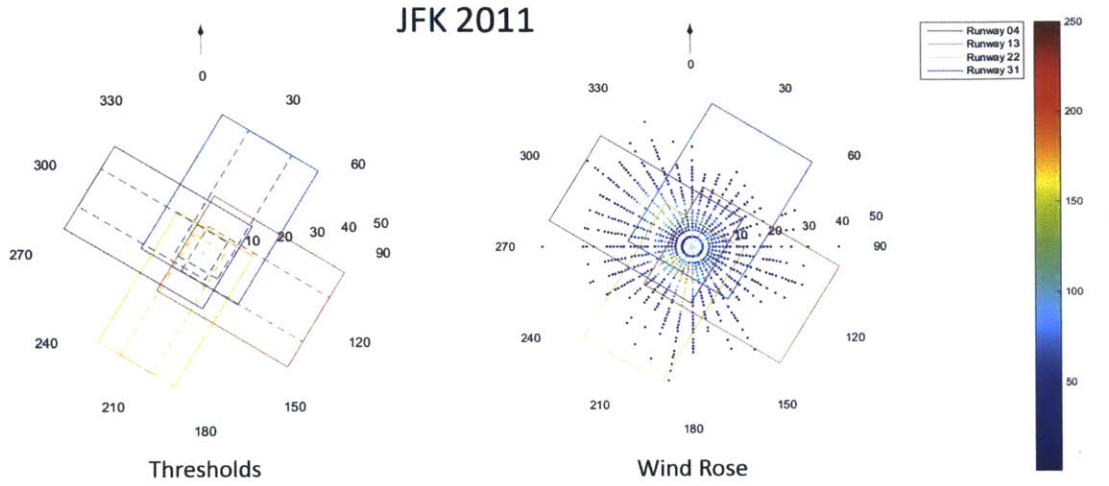

Figure 3.2.3: Wind speed, wind direction, and wind thresholds learned from year 2011 JFK data.

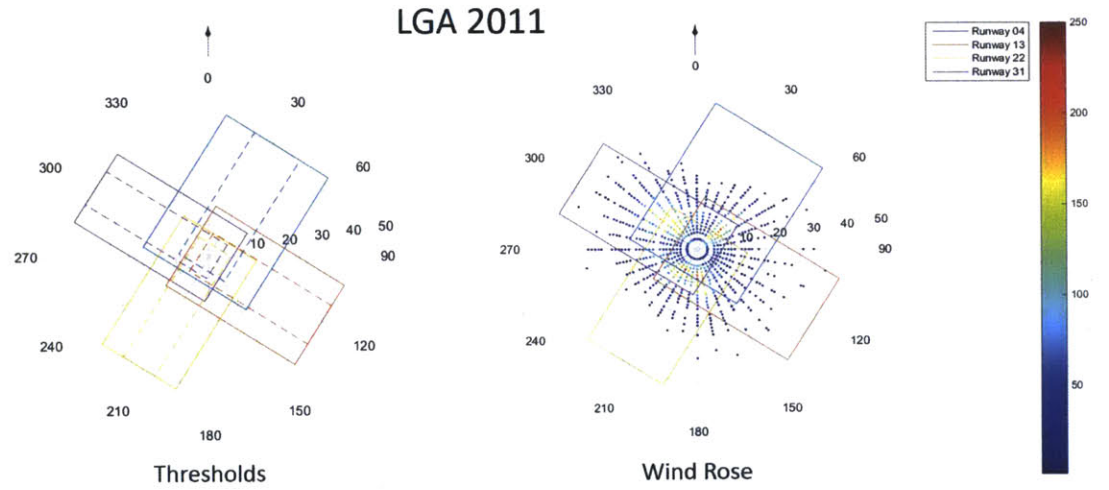

Figure 3.2.4: Wind speed, wind direction, and wind thresholds learned from year 2011 LGA data.

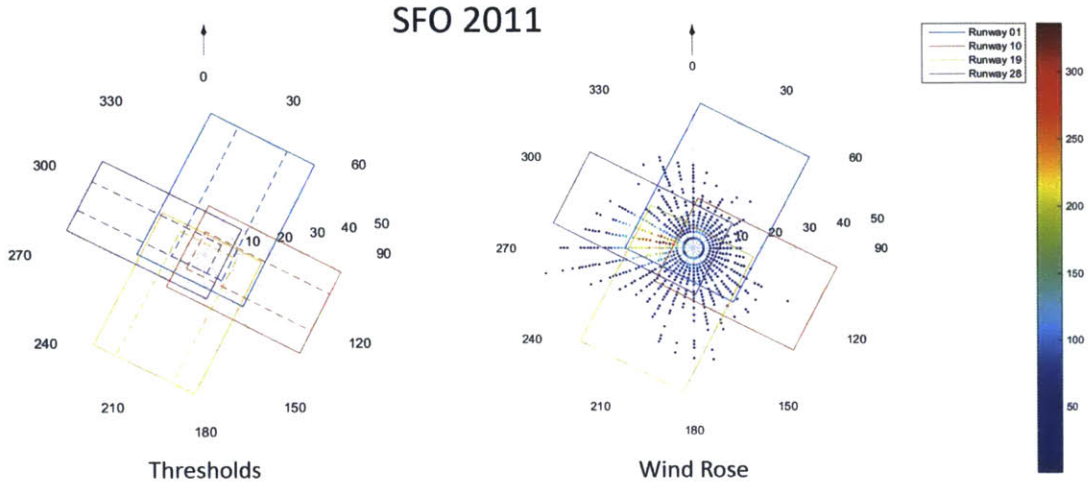

Figure 3.2.5: Wind speed, wind direction, and wind thresholds learned from year 2011 SFO data.

The available choice set during a given selection period can change in a discrete choice framework. Even with the maximum tailwind and crosswind limitations, several runway configurations may be feasible at any time. Runway feasibility was used to directly govern the available subset of runway configurations in the discrete choice model during each time interval. If the wind speed and wind direction combination fell outside any of the thresholds for a runway, all configurations using that runway were removed from the available choice set during the given decision selection period. Figures 3.2.2, 3.2.3, 3.2.4, and 3.2.5 show that the majority of points correspond to conditions in which all the runway configurations are considered feasible.

Headwinds are expected to add a positive contribution to the utility functions and tailwinds are expected to add a negative contribution. Significantly high headwinds, however, could potentially have an adverse effect on airport operations by decreasing the space between aircraft arrivals during landing - a phenomenon known as compression [29]. Shown in Figure 3.2.6, compression occurs when there are significantly higher headwinds near the ground than at cruise altitude during the arrival approach. If the ground headwinds are high enough, the relative spacing between consecutive arrivals will begin to decrease as the flights begin their descent toward the airport, which can cause safety problems and congestion.

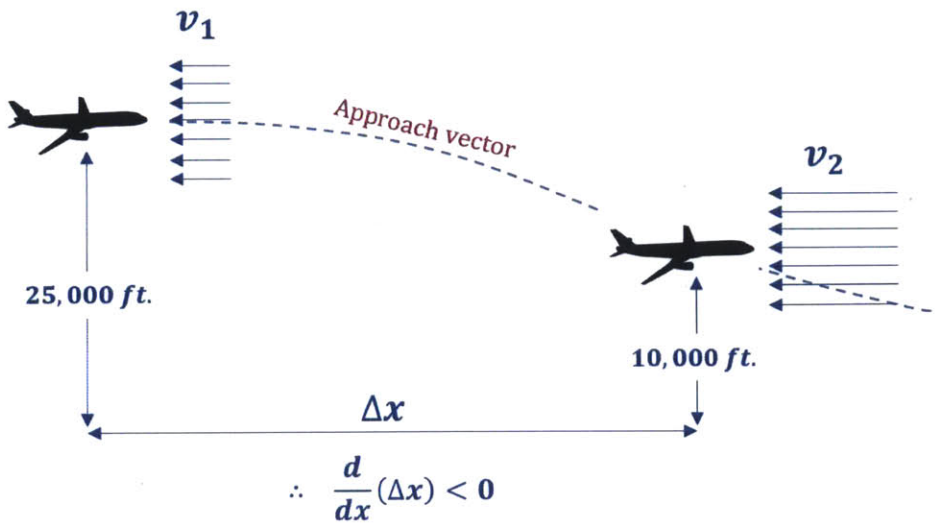

Figure 3.2.6: Diagram explaining aircraft compression upon arrival.

To account for the effects of compression, headwinds above the 85th percentile were treated as "high headwinds". Variables for "normal headwinds" (below the 85th percentile) and "tailwinds" were added to each model as well.

3.2.3 Demand

Airport arrival and departure demand play a significant role when selecting the runway configuration. Specifically, in high demand situations, high capacity configurations are preferred. These typically include an extra arrival or departure runway.

Figures 3.2.7, 3.2.8, 3.2.9, and 3.2.10 show the active arrival demand variation throughout the time of day at EWR, JFK, LGA and SFO airports during year 2011. Demand typically peaks around 9:00 - 11:00 AM because it is the most convenient time to arrive for business travel. Because of the time-dependence of arrival demand at each airport, adding demand into the utility functions of the runway configurations also brings time-of-day effects into the model. One would assume that if demand typically peaks at an airport during a certain time of day, the high capacity configurations are more likely to be used during that time of day.

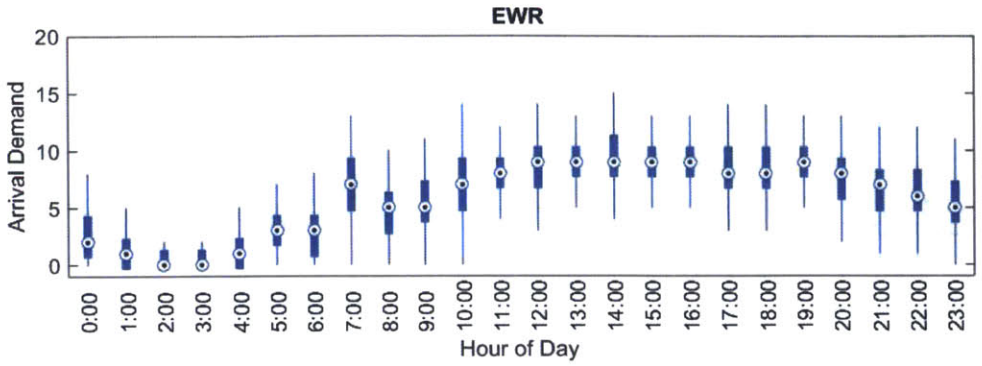

Figure 3.2.7: Temporal active arrival demand profile at EWR airport for year 2011.

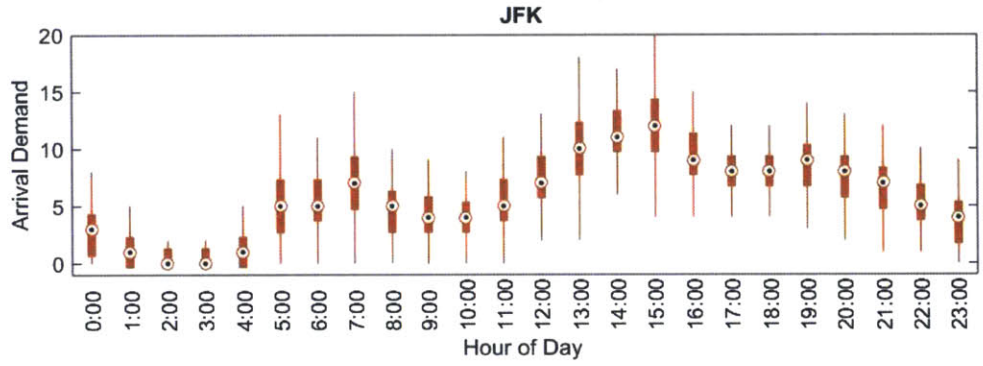

Figure 3.2.8: Temporal active arrival demand profile at JFK airport for year 2011.

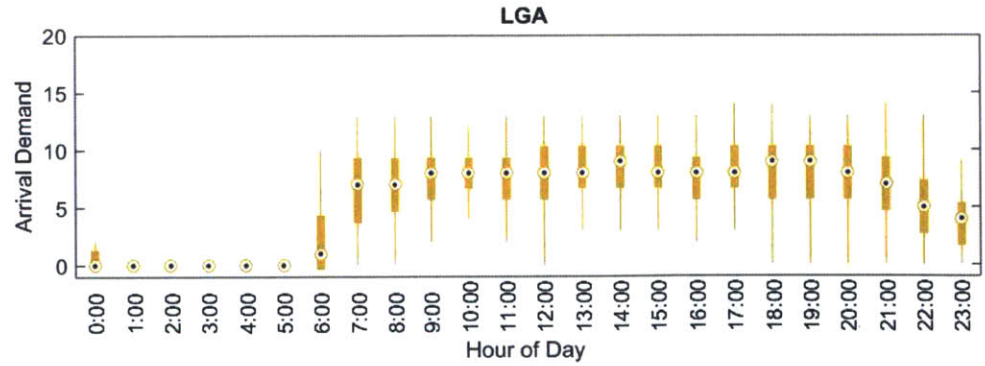

Figure 3.2.9: Temporal active arrival demand profile at LGA airport for year 2011.

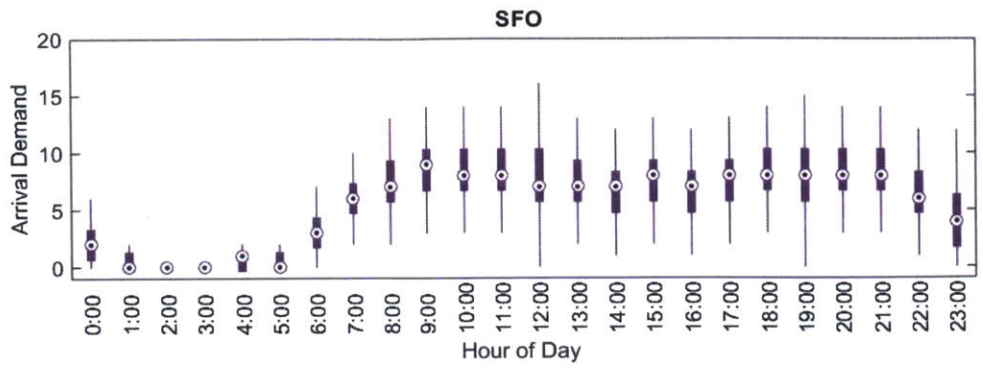

Figure 3.2.10: Temporal active arrival demand profile at SFO airport for year 2011.

3.2.4 Noise Abatement Procedures

Noise abatement procedures are used at many major airports to reduce the impacts of noise on communities in the vicinity of the airport, especially during early morning and nighttime hours. At EWR, JFK, and LGA, configurations with flight paths over the city and away from populated areas are preferred during the nighttime. At SFO, runway configurations that arrive and depart over the water are preferred to configurations that fly over populated areas. Variables were included in each model to account for these effects.

3.2.5 Cloud Ceiling and Visibility

Meteorological conditions, as represented by the cloud ceiling and visibility, are important considerations for air traffic control personnel when selecting the runway configuration. Visual Meteorological Conditions (VMC) refer to times when the visibility is sufficient for pilots to maintain visual separation from the ground and other aircraft. Instrument Meteorological Conditions (IMC) refer to times when pilots are required to use their flight instruments. IMC is defined by a visibility less than 3 miles and a ceiling less than 1,000 feet [30]. The runway configuration selection could depend on whether VMC or IMC is implemented. The EWR and LGA models incorporates variables for each utility function corresponding to visual and instrument conditions. In this manner, these variables will capture any preferences for one configuration over another configuration under VMC or IMC.

The optimal capacity configuration at SFO is 28R,28L|1R,1L. The runways used in this operation are closely-spaced at 750 feet apart. As per FAA regulations, simultaneous arrivals are not allowed under IMC [30]. Therefore, one would expect that the 28R|1R,1L or 28L|1R,1L configurations (which involve using only one of the two runways for arrivals) would be favored under IMC. Table 3.2.2 shows the relative use of each of these configurations under VMC and IMC. As shown, even though the single arrival runway configurations are used a greater fraction of the time in IMC than in VMC, the runway configuration 28R,28L|1R,1L is still used a majority of the time in IMC. During IMC, simultaneous "side-by" landings are not possible, and the airport operates as if it would be in a single arrival runway configuration using a "staggered" arrival approach. Operationally, the staggered 28R,28L|1R,1L configuration under IMC may have a small capacity benefit over the 28R|1R,1L and 28L|1R,1L configurations. In order to evaluate these effects, new variables that combine the effects of visibility and demand were used in the SFO model. Four categorical variables were defined for periods of:

1. IMC + low demand,
2. IMC + high demand,
3. VMC + low demand,
4. VMC + high demand,

where a low demand period was defined as less than 5 arrivals per 15-minute period and high demand was defined as greater than 8 arrivals per 15-minute interval.

Table 3.2.2: Occurrences of 28R,28L|1R,1L and 28R/L|1R,1L at SFO in 2011.

Configuration	VMC	IMC
28R,28L|1R,1L	19,832 (82.7%)	4,161 (17.3%)
28R/L|1R,1L	3,466 (65.3%)	1,844 (34.7%)

3.2.6 Coordination With Surrounding Airports

The four airports in the New York Metroplex - Newark (EWR), John F. Kennedy International Airport (JFK), LaGuardia (LGA), and Teterboro (TEB) - are all in

a very close proximity to one another. Air traffic controllers at each airport must therefore coordinate their aircraft arrival and departure flows with the other neighboring airports. In the EWR and LGA individual airport models, the impacts of TEB are ignored and categorical variables were added to account for operations at JFK. JFK was chosen because of its large volume of operations, which was expected to have a significant impact on the runway configurations at EWR and LGA. The JFK discrete-choice model includes coordination variables for both EWR and LGA, but also ignores the impacts from TEB.

3.2.7 Switch Proximity

If the airport conditions necessitate a runway configuration switch, certain configuration switches require more coordination from airport stakeholders than others. For instance, the addition of an extra arrival runway may be easier to implement than a complete reversal in the direction of operations. To account for these effects, variables were added to weight each utility function differently depending on the previous configuration. The switch proximity variables are fundamentally the same as the inertia variables, but are applied only to the utility functions of the runway configurations that were not seen in the previous time interval instead of the incumbent configuration. In this sense, they do not account for the low likelihood of switching between two runway configurations that require a high amount of operational effort once the decision to switch the configuration has been made.

3.2.8 New York Metroplex Model Data Processing

The New York Metroplex model uses slightly different definitions than the attributes mentioned above. Since each overarching New York Metroplex configuration consisted of several different runway configurations at multiple airports, runway availability plays a much smaller role than it does in the individual airport models. The wind parameters were processed to align with the primary flow of arrivals and departures (i.e. North or South flow). Arrival and departure demand variables were taken as

the total arrival or departure demand from EWR, JFK, and LGA. Additionally, the VMC/IMC classifiers from EWR, JFK, and LGA were all averaged to create an overall New York Metroplex VMC/IMC classifier. The averaging allows the utility bonus or penalty to be reduced for a New York Metroplex configuration if EWR, JFK, and LGA are not all in VMC or IMC.

3.3 Runway Configuration Filtering

Runway configurations that were seen less than 1% of the time throughout the year were removed at EWR, JFK, and SFO to reduce possible reporting errors or cases of special operations that would make it difficult to reliably estimate the weights and predict with the models. For LGA, the filter was set to 2% for the individual airport model to further help reduce reporting errors that occurred during nighttime hours. Similarly, the New York Metroplex configuration model uses a 1% cutoff.

Chapter 4

Estimated Discrete-Choice Utility Functions

4.1 EWR

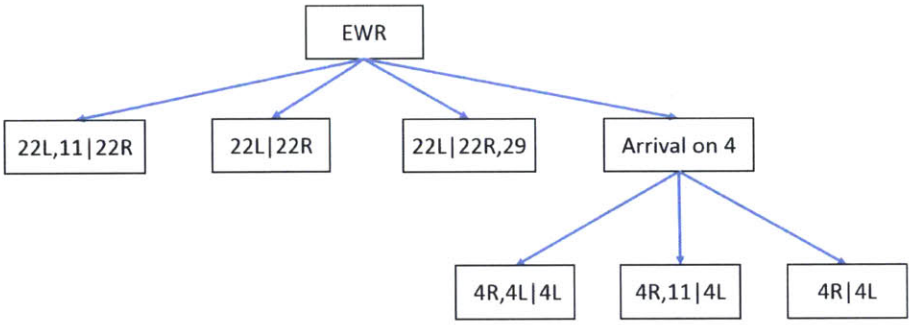

Figure 4.1.1: EWR model specification.

In the EWR model, configurations were removed if they were not seen at least 1% of the decision selection periods throughout the year. Many model structures were tested, and the final final model was chosen as a nested logit structure with nest containing all alternatives using runway 4 for arrivals with a scale parameter of $\mu_{Arr4} = 1.22$, shown in Figure 4.1.1.

The estimated weighting parameters for the utility function for the EWR model are shown along with their standard errors and t-statistics in Table 4.1.1. Any pa-

rameter that was not statistically significant, i.e. the t-statistic had an estimated value less than 1.96, was removed from the model except in cases when removing the variable could bias the predictions.

As expected, the estimated weights indicate that inertia is the most important factor to air traffic controllers when selecting a runway configuration at EWR. Because configurations within a common nest have a very similar directional aircraft flow structure, they were constrained to have the same inertia value. This helped prevent biases during the prediction.

Compression effects at EWR were not statistically significant during estimation and are therefore not shown in Table 4.1.1. Additionally, headwinds and tailwinds showed a linear correlation. As a result, the high headwind, normal headwind, and tailwind variables were all linearly constrained in the final model. The wind variables for the departure runways at EWR were not statistically significant, which may indicate that air traffic control prioritizes arrivals over departures when selecting a runway configuration.

Arrival demand was shown as statistically significant for runway configurations with extra arrival runways; namely, 22L,11|22R , 4R,4L|4L, and 4R,11|4L. As arrival demand increased, these configurations received a utility bonus which encourages the addition of an arrival runway to handle the increasing demand. The threshold values used in the model were determined as a nuisance parameter. Interestingly, a similar variable dealing with the departure demand was not statistically significant for runway configuration 22L|22R,29.

4.2 JFK

Configurations that were not seen within at least 1% of the decision selection periods in 2011 were removed from the JFK model. The final model was specified as a nested logit structure with nests containing arrivals and departures on the 4 runways, 13 runways, the 22 runways, and the 31 runways, shown in Figure 4.2.1. The scale parameters are $\mu_{4s} = 1.02$, $\mu_{13s} = 1.00$, $\mu_{22s} = 1.04$, and $\mu_{31s} = 1.07$ respectively.

Table 4.1.1: Estimated utility function weights for EWR.

Parameters	Value	Std. error	t-statistic
Inertia parameters			
Config. 22L,11|22R	4.70	0.329	14.29
Config. 22L|22R	4.70	0.329	14.29
Config. 22L|22R,29	4.70	0.329	14.29
Config. 4L,4R|4L	5.04	0.251	20.10
Config. 4L,11|4L	5.04	0.251	20.10
Config. 4R|4L	5.04	0.251	20.10
Wind parameters			
High headwind on arrival runway	0.107	0.0114	9.38
Normal headwind on arrival Runway	0.107	0.0114	9.38
Tailwind on arrival runway	-0.107	0.0114	-9.38
Tailwind on extra arrival runway	-0.295	0.0284	-10.39
VMC/IMC parameters			
VMC on 4R,4L|4	-2.48	0.187	-13.28
Noise parameters			
Arrival demand on 22L,11|22R (thresh =12)	1.97	0.819	2.41
Arrival demand on 4R,4L|4L (thresh=12)	2.26	0.419	5.41
Arrival demand on 4R,11|4L (thresh=12)	4.09	4.09	5.57

Tables 4.2.1 - 4.2.3 show the estimated values of the attribute weights in the utility functions, their standard errors, and their t-statistics. The estimated weights show that as before, inertia is the most important factor for a runway configuration selection at JFK, however the other parameters such as demand and visibility play a much more significant role than at EWR. Headwinds and especially tailwinds are still very significant with high t-statistics. In most cases, the arrival demand and departure demand variables give a higher positive bonus to higher capacity configurations as would be expected. 27 of the 132 possible switch proximity variables for JFK converged for this model. It appears that the switch proximity variables favor configuration switches that require less operational coordination, such as 22L,22R|22R,31L to 22L|22R. The exception seems to be runway switches to configurations involving the 31 Runways, which are always very likely to have a configuration switch. Additionally, coordination variables between EWR and LGA with JFK were extremely important in this model. In many cases, utilities in the JFK model receive a positive bonus if the configuration effectively aligns the arrival or departure flows with the

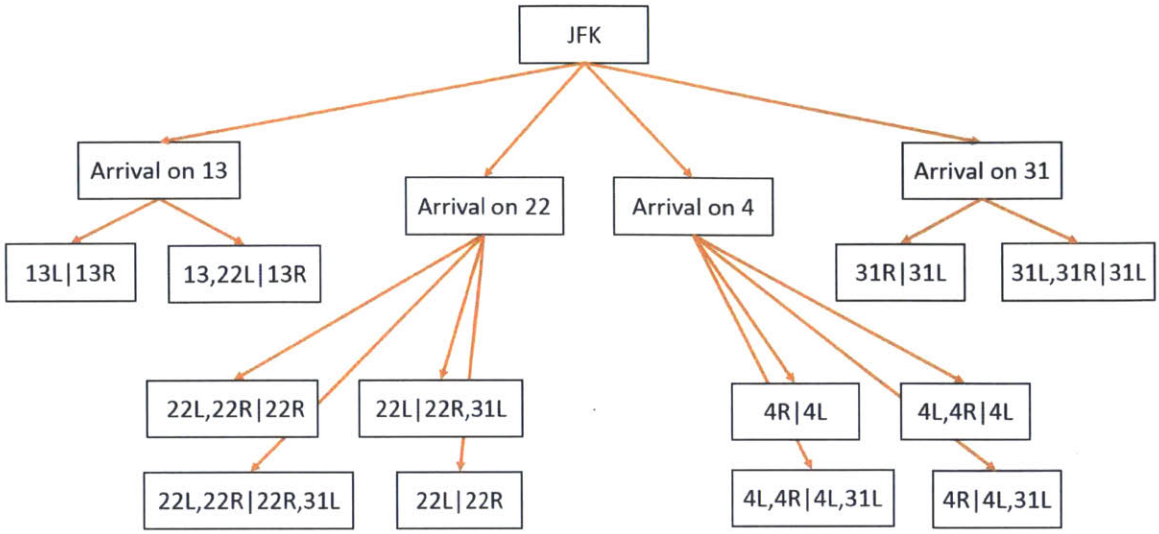

Figure 4.2.1: JFK model specification.

other airports and receives a negative bonus if it does not align arrival or departure flows.

Many of the filtered configurations are extremely similar in structure, such as 22L|22R and 22L,22R|22R, which causes confusion within the prediction models since many of the attributes give the same value within their respective utility functions. The higher uncertainty with the demand variables makes it difficult to distinguish between these two configurations except in extreme cases. Additionally, the ASPM data does not give any indication if configurations are reported as side-by or parallel landings and takeoffs which causes more confusion within the prediction models. Since in many cases, the FAA would consider these configurations the same when predicting on certain horizon, the runway configurations that were very similar in structure are combined and the utilities were re-estimated in a new combined configuration model for JFK. The combined configurations are shown in Table 4.2.4 and the updated model specification is shown in Figure 4.2.2. The prediction models shown later use the combined configurations.

The updated utility parameters estimated on the combined model is shown in Table 4.2.5. The utility weights in the combined model are a reflection of the utility weights from the other model. Inertia is still shown as the most important parameter.

Table 4.2.1: Estimated utility function weights for JFK - Part I.

Parameters	Value	Std. error	t-statistic
Inertia parameters			
Config. 13L,22L|13R	4.62	0.233	19.82
Config. 13L|13R	4.97	0.261	19.06
Config. 22L,22R|22R	6.21	0.286	21.72
Config. 22L,22R|22R,31L	6.28	0.38	16.53
Config. 22L|22R	6.43	0.338	19.03
Config. 22L|22R,31L	6.57	0.372	17.64
Config. 31L,31R|31L	6.01	0.335	17.94
Config. 31R|31L	5.58	0.235	23.73
Config. 4L,4R|4L	4.17	0.171	24.44
Config. 4L,4R|4L,31L	5.83	0.193	30.27
Config. 4R|4L	5.94	0.224	26.52
Config. 4R|4L,31L	6.61	0.352	18.76
Wind parameters			
headwind on arrival runway	0.0365	0.0096	3.8
Tailwind on arrival runway	-0.138	0.0171	-8.06
Demand parameters			
Arrival demand on 13L,22L|13R	0.257	0.022	11.65
Arrival demand on 4L,4R|4L,31L	0.106	0.0312	3.4
Arrival demand on 31L,31R|31L	0.151	0.0196	7.69
Arrival demand on 4L,4R|4L	0.124	0.0307	4.04
Departure demand on 4L,4R|4L,31L	0.307	0.0286	10.73
Departure demand on 4R|4L	0.139	0.0306	4.55
Departure demand on 4R|4L,31L	0.361	0.0255	14.19
Departure demand on 13L|13R	0.187	0.0298	6.28
Departure demand on 22L,22R|22R	0.0981	0.0300	3.27
Departure demand on 22L,22R|22R,31L	0.346	0.0245	14.12
Departure demand on 22L|22R	0.154	0.0248	6.23
Departure demand on 22L|22R,31L	0.348	0.0191	18.24
Departure demand on 31R|31L	0.201	0.0248	8.13
VMC/IMC parameters			
4L,4R|4L,31L under VMC	-1.05	0.248	-4.24
4R|4L under VMC	-0.754	0.232	-3.24
4R|4L,31L under VMC	-0.866	0.226	-3.83
4L,4R|4L under VMC	-0.709	0.239	-2.97
Noise parameters			
Morning/evening arrivals on runway 4	0.418	0.137	3.06
Morning/evening departures on runway 22	0.418	0.137	3.06

Headwinds give positive contributions to the utility functions of each runway configuration. Tailwinds contribute a larger negative bonus to the utility functions, reflecting

Table 4.2.2: Estimated utility function weights for JFK - Part II.

Parameters	Value	Std. error	t-statistic
Switch proximity parameters			
31L,31R|31L to 4L,4R|4L,31L	0.945	0.325	2.91
4L,4R|4L to 4L,4R|4L,31L	1.78	0.446	3.99
4R|4L,31L to 4R|4L	3.53	0.428	8.26
31L,31R|31L to 4R|4L	1.76	0.431	4.09
4L,4R|4L to 4R|4L,31L	1.18	0.424	2.79
13L|13R to 13L,22L|13R	1.7	0.458	3.71
22L|22R to 13L,22L|13R	0.923	0.365	2.53
22L,22R|22R,31L to 22L,22R|22R	2.21	0.472	4.68
22L|22R to 22L,22R|22R	1.71	0.342	4.98
22L,22R|22R to 22L,22R|22R,31L	1.68	0.543	3.09
31L,31R|31L to 22L,22R|22R,31L	1.02	0.283	3.58
22L,22R|22R to 22L|22R	1.39	0.475	2.93
22L,22R|22R,31L to 22L|22R	1.42	0.503	2.83
22L|22R,31L to 22L|22R	1.38	0.258	5.34
22L,22R|22R to 22L|22R,31L	1.3	0.422	3.07
4R|4L to 31L,31R|31L	1.18	0.402	2.93
4R|4L,31L to 31L,31R|31L	2.31	0.461	5.01
13L|13R to 31L,31R|31L	1.49	0.557	2.68
22L,22R|22R to 31L,31R|31L	1.38	0.525	2.63
22L,22R|22R,31L to 31L,31R|31L	1.59	0.431	3.68
22L|22R to 31L,31R|31L	1.48	0.348	4.25
22L|22R,31L to 31L,31R|31L	0.664	0.31	2.14
31R|31L to 31L,31R|31L	1.02	0.359	2.85
4L,4R|4L to 31L,31R|31L	1.46	0.509	2.86
4L,4R|4L,31L to 4L,4R|4L	1.68	0.388	4.33
4R|4L to 4L,4R|4L	1.62	0.45	3.61
4R|4L,31L to 4L,4R|4L	2.79	0.496	5.63

a larger preference from air traffic control to choose runway configurations that eliminate tailwinds than to chose runway configurations that have high headwinds. The noise parameter indicates that during evening and morning hours, departures are likely to occur using runway 22 and arrivals are likely to occur on runway 4. Additionally, the switch proximity variables only converged for 4 out of the 30 possible combined configuration switches.

Arrival demand for combined configurations 13 Runways and 31 Runways are statistically significant and departure demand for configuration 22 Runways High Arrival and 22 Runways Low Arrival are statistically significant. Additionally, utilities

Table 4.2.3: Estimated utility function weights for JFK - Part III.

Parameters	Value	Std. error	t-statistic
Inter-airport coordination parameters			
EWR 21L,11|22R VS. JFK 13L,22L|13R	0.882	0.242	3.64
EWR 21L,11|22R VS. JFK 13L|13R	-0.718	0.439	-1.64
EWR 21L,11|22R VS. JFK 22L|22R,31L	0.479	0.2	2.39
EWR 21L|22R VS. JFK 4L,4R|4L,31L	-0.639	0.257	-2.48
EWR 21L|22R VS. JFK 4R|4L	0.621	0.236	2.63
EWR 21L|22R VS. JFK 22L,22R|22R	0.287	0.214	1.34
EWR 21L|22R VS. JFK 22L,22R|22R,31L	-0.979	0.238	-4.11
EWR 21L|22R VS. JFK 22L|22R	0.434	0.177	2.45
EWR 4R,11|4L VS. JFK 13L,22L|13R	1.5	0.432	3.48
EWR 4R,11|4L VS. JFK 4L,4R|4L,31L	1.37	0.366	3.74
EWR 4R,11|4L VS. JFK 31L,31R|31L	1.21	0.303	4
EWR 4R|4L VS. JFK 13L,22L|13R	-0.91	0.369	-2.46
EWR 4R|4L VS. JFK 4R|4L	0.574	0.228	2.52
EWR 4R|4L VS. JFK 22L,22R|22R,31L	-0.818	0.317	-2.58
EWR 4R|4L VS. JFK 22L|22R,31L	-0.652	0.249	-2.62
EWR 4R|4L VS. JFK 31L,31R|31L	0.529	0.175	3.02
LGA 22|13 VS. JFK 13L,22L|13R	-0.9	0.24	-3.74
LGA 22|13 VS. JFK 4L,4R|4L,31L	-0.843	0.335	-2.51
LGA 22|13 VS. JFK 4R|4L,31L	-1.22	0.338	-3.61
LGA 22|13 VS. JFK 22L,22R|22R,31L	-1.76	0.372	-4.72
LGA 22|13 VS. JFK 22L|22R	0.466	0.178	2.61
LGA 22|31 VS. JFK 4L,4R|4L,31L	-1.56	0.651	-2.39
LGA 22|31 VS. JFK 4R|4L,31L	-1.35	0.653	-2.06
LGA 22|31 VS. JFK 31L,31R|31L	0.413	0.188	2.2
LGA 22,31|31 VS. JFK 31L,31R|31L	1.66	0.335	4.96
LGA 4|13 VS. JFK 4L,4R|4L,31L	-0.949	0.326	-2.91
LGA 4|13 VS. JFK 4R|4L	0.581	0.237	2.45
LGA 4|13 VS. JFK 13L|13R	-1.19	0.497	-2.4
LGA 4|13 VS. JFK 22L,22R|22R,31L	-2.41	0.675	-3.57
LGA 4|13 VS. JFK 4L,4R|4	1.01	0.27	3.74
LGA 4|4 VS. JFK 4R|4L,31L	-0.934	0.382	-2.44

for configurations 4 Runways High Arrival and 4 Runways Low Arrival receive a negative contribution if VMC are in effect. Prior reports from the FAA confirm that configuration 13 Runways is primarily used under arrival priority operations, configurations 22 Runways High Arrival and 22 Runways Low Arrival are used for operations under departure priority, and configurations 4 Runways High Arrival and 4 Runways Low Arrival are used for IMC operations [25]. The fact that the data-

Table 4.2.4: Combined JFK configurations.

ID	Configuration	Freq.	% Freq.
1	13L,22L|13R	2,174	6.4%
2	13L|13R	1,395	4.1%
13 Runways		**3,569**	**10.4%**
3	22L,22R|22R	2,579	7.5%
4	22L,22R|22R,31L	1,785	5.2%
22 Runways High Arrival		**4,364**	**12.8%**
5	22L|22R	3,528	10.3%
6	22L|22R,31L	4,085	12.0%
22 Runways Low Arrival		**7,613**	**22.3%**
7	31L,31R|31L	7,064	20.7%
8	31R|31L	3,233	9.5%
31 Runways		**10,297**	**30.1%**
9	4L,4R|4L	1,927	5.6%
10	4L,4R|4L,31L	1,987	5.8%
4 Runways High Arrival		**3,914**	**11.4%**
11	4R|4L	2,592	7.6%
12	4R|4L,31L	1,835	5.4%
4 Runways Low Arrival		**4,427**	**13.0%**

driven approach aligns with previous research on FAA ATC is promising, because it reflects the potential of the discrete-choice models to capture the importance of various attributes in a choice selection. Furthermore, it opens the door for new insights if attributes that may not be obvious are statistically significant.

Additionally, coordination with EWR and LGA seems to be important to the runway configuration selection at JFK, moreso than with the other airport models. JFK is the largest airport in the New York Metropolitan area, which may make have a larger influence on the operations in the Metroplex. This could effectively make it more important for ATC at JFK to coordinate their operations with the other nearby airports.

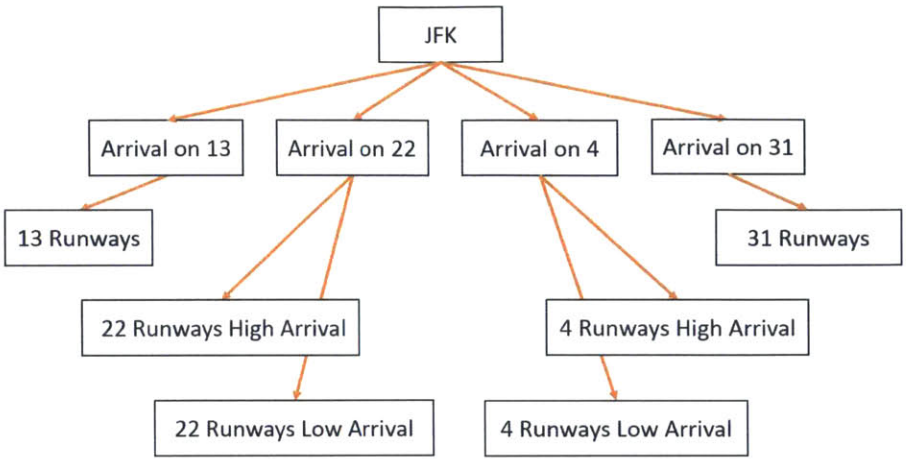

Figure 4.2.2: JFK Combined Configuration Model specification.

4.3 LGA

In the LGA model, configurations were removed if they were not seen in at least 2% of the decision selection periods throughout the year. Many model structures were tested, and the final model was chosen as a nested logit structure with a single nest containing all alternatives using runway 22 for arrivals with a scale parameter of $\mu_{Arr22} = 1.1$, shown in Fig. 4.3.1.

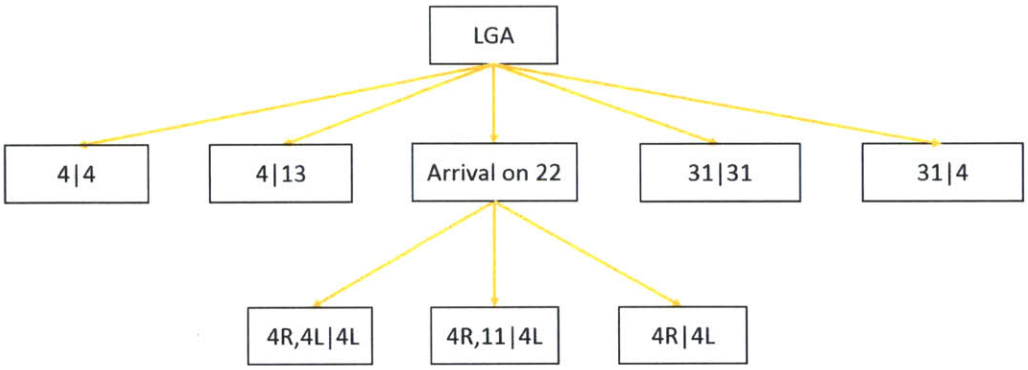

Figure 4.3.1: LGA model specification.

The estimated values of the of the weights on different attributes (i.e., $\vec{\beta}$), their standard errors and t-statistics are presented in Tables 4.3.1 - 4.3.2 . Again, the estimated weights indicate that inertia is the most important factor when selecting a runway configuration at LGA, particularly for configurations 22|31 and 22,31|31.

Table 4.2.5: Estimated utility function weights for JFK's Combined Configuration Model.

Parameters	Value	Std. error	t-statistic
Inertia parameters			
Config. 13 Runways	4.86	0.185	26.31
Config. 22 Runways High Arrival	4.31	0.0931	46.30
Config. 22 Runways Low Arrival	4.31	0.0931	46.30
Config. 31 Runways	3.82	0.224	17.08
Config. 4 Runways High Arrival	5.81	0.175	33.23
Config. 4 Runways Low Arrival	5.81	0.175	33.23
Wind parameters			
High headwind on arrival runway	0.0194	0.0148	1.32
Normal headwind on arrival runway	0.0356	0.0122	6.26
Tailwind on arrival runway	-0.159	0.0179	-8.91
Demand parameters			
Arrival demand on 13 Runways	0.151	0.0186	8.09
Arrival demand on 31 Runways	0.208	0.0175	11.86
Departure demand 22 Runways High Arrival	0.203	0.0176	11.55
Departure demand 4 Runways High Arrival	0.204	0.0222	9.17
Noise parameters			
Morning/evening arrivals on runway 22	-0.208	0.128	-1.63
Morning/evening departures on runway 4	-0.208	0.128	-1.63
VMC/IMC parameters			
VMC on 4 Runways High Arrival	-0.898	0.199	-4.52
VMC on 4 Runways Low Arrival	-1.09	0.198	-5.52
Switch proximity parameters			
31 Runways to 13 Runways	-0.885	0.335	-2.64
31 Runways to 22 Runways High Arrival	-1.36	0.327	-4.18
31 Runways to 22 Runways Low Arrival	-0.735	0.271	-2.71
31 Runways to 4 Runways Low Arrival	2.12	0.294	7.23
Inter-airport coordination parameters			
JFK config. 31 Runways VS. EWR config. 21L,11|22R	-0.666	0.218	-3.05
JFK config. 4 Runways High Arrival VS. EWR config. 21L|22R	-0.357	0.194	-1.84
JFK config. 13 Runways VS. EWR config. 4R,11|4L	0.942	0.391	2.41
JFK config. 4 Runways Low Arrival VS. EWR config. 4R,11|4L	-1.05	0.372	-2.82
JFK config. 22 Runways Low Arrival VS. EWR config. 4R|4L	-0.593	0.197	-3.01
JFK config. 22 Runways Low Arrival VS. LGA config. 22|13	0.651	0.133	4.88
JFK config. 4 Runways High Arrival VS. LGA config. 22|31	-1.77	0.544	-3.25
JFK config. 4 Runways Low Arrival VS. LGA config. 22|31	-1.19	0.513	-2.32
JFK config. 22 Runways Low Arrival VS. LGA config. 22,31|31	-1.13	0.371	-3.05
JFK config. 22 Runways High Arrival VS. LGA config. 31|4	0.324	0.246	1.32
JFK config. 4 Runways Low Arrival VS. LGA config. 4|4	-1.49	0.364	-4.10

Additionally, the results indicate that the headwind parameters are statistically significant for the primary arrival runway, but not for the primary departure runway or the extra arrival runway. This finding seems to suggest that the alignment of the primary arrival runway is more important than the alignment of the departure or extra arrival runways. This could be due to the fact that aircraft arrivals must be served, whereas departures could be held under extreme conditions. Furthermore, the negative influence of tailwinds was found to be statistically significant across the board. Note that the high headwind variable for the primary arrival runway had a slightly lower value than the normal headwind variable, which suggests that high headwinds are slightly less preferable due to compression.

Arrival demand effects were statistically significant for the low capacity runway configurations 31|31 and 4|4. During high demand scenarios, these configurations were less likely to be selected. VMC was seen to be important for configurations 31|31 and 31|4. This seems to suggest that VMC is an important consideration for arrivals on runway 31.

Switch proximity was only significant for 10 out of the 42 possible configuration switches. All had negative values, reflecting resistances to certain configuration switches. The relative weights suggest that, in general, air traffic controllers prefer not to reverse the direction of airport operations if at all possible.

4.4 SFO

4.4.1 SFO Model Using Raw ASPM Data

SFO Model Using Raw ASPM Data

In the SFO model, configurations that were removed if they were not seen in at least 1% of the decision selection periods throughout the year. Shown in Fig. 4.4.1, the chosen model had a nested logit structure with 6 possible runway configuration alternatives. The model structure grouped similar configurations 28R,28L|1R,1L , 28L|1R,1L , and 28R|1R,1L into a common nest with scale parameter $\mu_{Arr28,Dep1}$ =

Table 4.3.1: Estimated utility function weights for LGA - Part I.

Parameters	Value	Std. error	t-statistic
Inertia parameters			
Config. 22|13	4.58	0.187	24.5
Config. 22|31	7.41	0.36	20.57
Config. 22,31|31	7.41	0.36	20.57
Config. 31|31	4.91	0.401	12.24
Config. 31|4	3.16	0.25	12.6
Config. 4|13	3.99	0.196	20.34
Config. 4|4	5.44	0.416	13.1
Wind parameters			
High headwind on arrival runway	0.0952	0.0161	5.89
Normal headwind on arrival runway	0.123	0.0197	6.26
Tailwind on arrival runway	-0.0946	0.0199	-4.74
Tailwind on departure runway	-0.211	0.0173	-12.2
Tailwind on extra arrival runway	-0.348	0.07	-4.97
Demand parameters			
Arrival demand; 31|31	-0.101	0.0312	-3.24
Arrival demand; 4|4	-0.0807	0.0327	-2.47
VMC/IMC parameters			
VMC on 31|31	2.09	0.402	5.19
VMC on 31|4	1.36	0.231	5.9

1.16.

The estimated weighting parameters for the utility function (β values from (2.1.3)) for the SFO model are shown in Table 4.4.1. The table includes the estimated value, the standard error, and the t-statistics.

Similar to the EWR, JFK, and LGA models, the inertia variables were identified as the most important factors for the air traffic controller's decision selection at SFO. Note that all configurations within the common nest have the same inertia variable value because they were modeled with a common inertia variable. This was a necessary constraint because of the aforementioned VMC/IMC reporting challenges, which caused the runway configurations in this nest to be very sensitive to biases in the prediction model.

Wind was another significant factor for the SFO decision selection. During estimation, compression did not show a significant influence on the high headwind attributes. Additionally, the headwind and tailwind were linearly correlated. There-

Table 4.3.2: Estimated utility function weights for LGA - Part II.

Parameters	Value	Std. error	t-statistic
Switch proximity parameters			
31|4 to 31|31	-1.4	0.463	-3.03
4|13 to 31|31	-2.52	0.714	-3.53
4|4 to 31|31	-1.22	0.747	-1.77
22|13 to 31|31	-1.99	0.577	-3.45
4|13 to 31|4	-2.19	0.368	-5.94
4|4 to 31|4	-1.05	0.515	-2.04
22|13 to 31|4	-2.14	0.355	-6.04
4|13 to 4|4	-1.6	0.443	-3.61
22|13 to 4|4	-1.92	0.532	-3.6
31|31 to 22|13	-1.05	0.573	-1.84
Inter-airport coordination parameters			
LGA departures on runway 4 VS. JFK arrivals on runway 13	0.85	0.308	2.76
LGA departures on runway 13 VS. JFK arrivals on runway 13	1.27	0.464	2.75
LGA departures on runway 31 VS. JFK departures on runway 22	-1.99	0.224	-8.88
LGA departures on runway 22 VS. JFK arrivals on runway 22	-0.448	0.172	-2.6
LGA departures on runway 31 VS. JFK arrivals on runway 31	-1.61	0.222	-7.26
LGA departures on runway 13 VS. JFK arrivals on runway 31	0.796	0.25	3.19
LGA departures on runway 31 VS. JFK departures on runway 4	-2.5	0.341	-7.34
LGA departures on runway 4 VS. JFK arrivals on runway 4	-0.737	0.293	-2.51
LGA departures on runway 22 VS. JFK departures on runway 4	-1.15	0.312	-3.68
LGA arrivals on runway 22 VS. JFK arrivals on runway 13	0.85	0.308	2.76
LGA arrivals on runway 31 VS. JFK arrivals on runway 13	1.27	0.464	2.75
LGA arrivals on runway 13 VS. JFK departures on runway 22	-1.99	0.224	-8.88
LGA arrivals on runway 4 VS. JFK arrivals on runway 22	-0.448	0.172	-2.6
LGA arrivals on runway 13 VS. JFK arrivals on runway 31	-1.61	0.222	-7.26
LGA arrivals on runway 31 VS. JFK arrivals on runway 31	0.796	0.25	3.19
LGA arrivals on runway 13 VS. JFK departures on runway 4	-2.5	0.341	-7.34
LGA arrivals on runway 22 VS. JFK arrivals on runway 4	-0.737	0.293	-2.51
LGA arrivals on runway 4 VS. JFK departures on runway 4	-1.15	0.312	-3.68

fore, as shown in the table values, these variables were constrained linearly for the final model. Shown in Fig. 3.2.5, the wind at SFO is predominantly from the San Bruno Gap which corresponds to a headwind for arrivals on runway 28. Five of the six unique configurations in the SFO model serve arrivals using runways 28R and 28L due to this headwind advantage. During poor weather conditions, low pressure systems with high circulating winds disrupt the typical pattern, making runway configuration 19R,19L|10R,10L a more attractive runway configuration decision for air

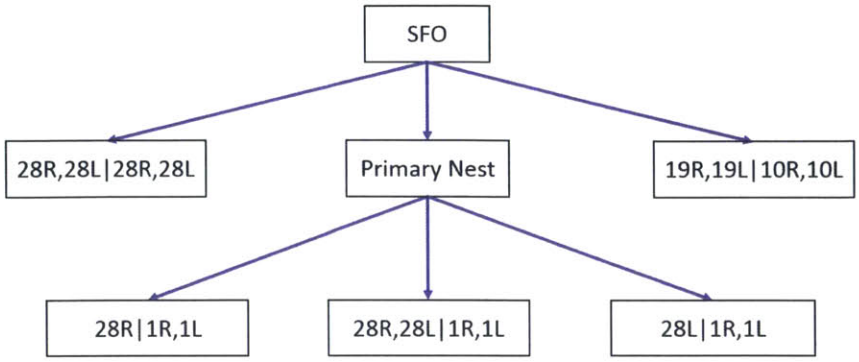

Figure 4.4.1: SFO model specification.

traffic controllers [30]. As a consequence, the headwind variables in the model are predominately used to distinguish between runway configuration options with arrivals on the 28's or arrivals on the 19's.

In the SFO model, attributes that grouped the effects of demand and visibility were estimated. As shown in Table 4.4.1, the utility functions for the 28R|1R,1L and 28L|1R,1L configurations received a negative demand bonus under VMC and a positive demand bonus under IMC. Additionally, the noise parameters indicate that flights over the water during the evening hours are preferred to flights departing over nearby communities. This seems to coincide with noise abatement procedures at SFO [31], and would likely be more prevalent if other configurations such as 28R,28L|10R,10L were modeled in the future.

The switch proximity variables also showed statistically significant effects for 10 of the 30 possible switches in the SFO model. In particular, the estimated coefficients reflected a preference to switch out of runway configuration 28R,28L|28R,28L which is likely because this configuration is used primarily for long haul arrivals and departures that fly over the Pacific Ocean. Also, configuration 19R,19L|10R,10L does not show a preference to switch back to the configurations with the arrivals on the 28's. This could be because the 19R,19L|10R,10L configuration would be used during significant shifts in wind patterns over the San Bruno Gap.

Table 4.4.1: Estimated utility function weights for SFO.

Parameters	Value	Std. error	t-statistic
Inertia parameters			
Config. 19R,19L|10R,10L	3.20	0.376	8.64
Config. 28R,28L|1R,1L	4.48	0.139	31.87
Config. 28R,28L|28R,28L	4.35	0.209	20.82
Config. 28R|1R,1L	4.48	0.139	31.87
Config. 28L|1R,1L	4.48	0.139	31.87
Wind parameters			
High headwind on arrival Runway	0.0415	0.0131	3.20
Normal headwind on arrival Runway	0.0415	0.0131	3.20
Tailwind on arrival Runway	-0.0415	0.0131	-3.20
High headwind on departure Runway	0.0608	0.0076	8.14
Normal headwind on departure Runway	0.0608	0.0076	8.14
Tailwind on departure Runway	-0.0608	0.0076	-8.14
Demand/visibility parameters			
VMC + high demand; 28R|1R,1L	-1.66	0.390	-3.44
VMC + high demand; 28L|1R,1L	-1.66	0.390	-3.44
IMC + low demand; 28R|1R,1L	-0.327	0.381	-0.93
IMC + low demand; 28L|1R,1L	-0.327	0.381	-0.93
Switch proximity parameters			
28R,28L|1R,1L to 28R,28L|28R,28L	-1.20	0.189	-5.98
28R|1R,1L to 28R,28L|28R,28L	-1.20	0.189	-5.98
28L|1R,1L to 28R,28L|28R,28L	-1.20	0.189	-5.98
19R,19L|10R,10L to 28R,28L|1R,1L	-0.775	0.649	-1.22
19R,19L|10R,10L to 28R|1R,1L	-0.775	0.649	-1.22
19R,19L|10R,10L to 28L|1R,1L	-0.775	0.649	-1.22
19R,19L|10R,10L to 28R,28L|28R,28L	-0.946	0.601	-1.57
28R,28L|28R,28L to 28R,28L|1R,1L	1.33	0.251	4.97
28R,28L|28R,28L to 28R|1R,1L	1.33	0.251	4.97
28R,28L|28R,28L to 28L|1R,1L	1.33	0.251	4.97
Noise abatement parameters			
Depart Runway 28 during evening	-0.356	0.176	-3.23
Arrive Runway 10 during evening	-0.356	0.176	-3.23

4.4.2 SFO Proof-of-Concept Study Using Airport AARs to Separate Sideby and Staggered Configurations

As stated previously, the nature of the airport layout, the ASPM data, and the weather at SFO made it a difficult airport to model accurately. The optimal capacity configuration at SFO is 28R,28L|1R,1L and the runways used in this operation are

closely-spaced at 750 feet apart. As per FAA regulations, simultaneous arrivals are not allowed under IMC conditions [29]. Therefore, one would expect that the 28R|1R,1L and 28L|1R,1L configurations would be favored under IMC, however the ASPM data suggests that configuration 28R,28L|1R,1L is used more times under IMC than configurations 28R|1R,1L and 28L|1R,1L (see table 4.4.2). The fact that SFO primarily reports configuration 28R,28L|1R,1L under IMC seems like a possible reporting error based on the FAA runway separation requirements, however it can be explained by the nature of operations at SFO and the reporting intervals in the ASPM dataset. Under IMC, simultaneous side-by landings are not possible, and the airport operates as if it would be in a single arrival runway configuration using a staggered arrival approach. Because the ASPM data is reported for an entire 15-minute interval, side-by and staggered approaches are both reported as configuration 28R,28L|1R,1L, when in reality they are 28R,28L|1R,1L-SIDEBY and 28R,28L|1R,1L-STAG respectively. Operationally, the 28R,28L|1R,1L-STAG configuration may have a small capacity benefit over the 28R|1R,1L and 28L|1R,1L configurations.

The ASPM data does not specify whether configuration 28R,28L|1R,1L is being operated in side-by or staggered. Because of this limitation on the data set, the prior discrete-choice model for SFO does not attempt to classify configurations as either side-by or staggered since the lack of any ground truth would make such a model very difficult to validate. Since future models may want to develop a side-by/staggered model in more rigor, a proof-of-concept analysis for SFO is evaluated by using the ASPM data to classify configuration 28R,28L|1R,1L as either side-by or staggered. Rather than specify a deterministic rule assigning all VMC flights to side-by and all IMC flights to staggered, the calculated Airport Arrival Rates (AARs) in the ASPM data were used as a tuning parameter. A histogram of the AARs under VMC and IMC at SFO for year 2011 is shown in Figure 4.4.2. As shown, all AARs have a higher count under VMC than under IMC. Note, however, that the number of counts at low AARs for both VMC and IMC are comparable. When viewing probabilities instead of counts, Figure 4.4.2 clearly shows that the low AAR ranges are much more likely under IMC than under VMC as would be expected. After further examination,

the runway configuration was classified as 28R,28L|1R,1L-SIDEBY if the AAR was greater than or equal to 18, and it was classified as 28R,28L|1R,1L-STAG if the AAR was less than 18.

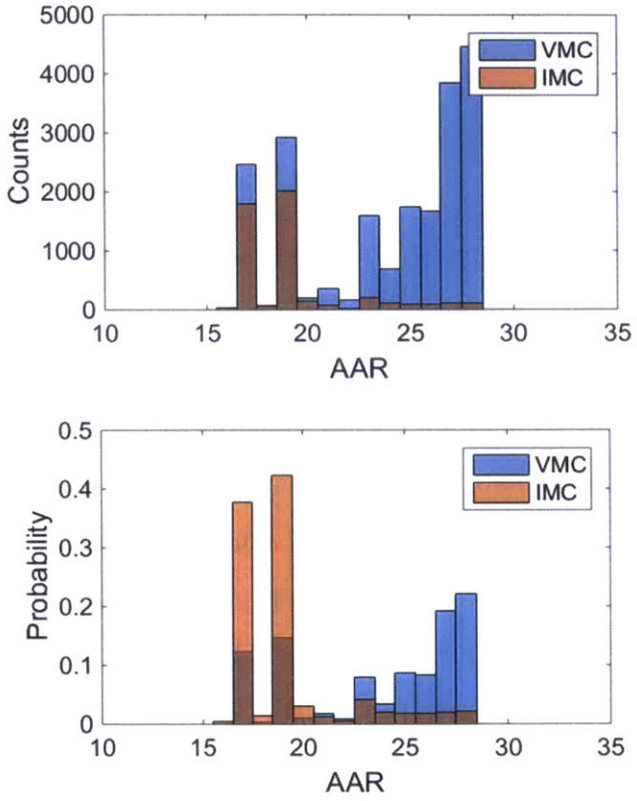

Figure 4.4.2: SFO AARs under VMC and IMC in 2011.

While the AAR rate reported in the ASPM data seems like the best way to classify side-by and staggered configurations for this proof-of-concept analysis, it should be noted that if this approach is taken to help develop future decision support tools or airport capacity models, side-by and staggered configurations should be classified more rigorously for several reasons. First, classifying side-by and staggered configurations using AARs, which could act as a substitute for the airport capacity at a given time, is a problem if the utilities are used to supplement airport capacity prediction models. In this case, the model would be useless because the airport capacity would essentially be included as an independent variable. Secondly, the AAR is determined using a proprietary model from the FAA, which may not be accurate under all situations. Third and most importantly, since the ASPM dataset does not include side-by

and staggered designations, there is no ground truth to be used for validating the accuracy of the prediction model. It would be very difficult to justify a model's accuracy without any solid data for validation. For these reasons, the utility functions under this AAR classification approach will be estimated, but they will not be shown within a prediction model.

After reclassifying the 28R,28L|1R,1L configuration as either 28R,28L|1R,1L- SIDEBY or 28R,28L|1R,1L-STAG according to the AAR cutoff determined above, the updated discrete-choice utility beta parameters were estimated and are shown in Table 4.4.2. Note that for the most part, the utility functions are consistent with the prior SFO model using the raw ASPM data. The utility function for 28R,28L|1R,1L-SIDEBY revived a positive demand bonus under VMC. By contrast, 28R,28L|1R,1L-STAG received a positive demand bonus under IMC. Interestingly, configurations 28R|1R,1L and 28L|1R,1L did not show any statistically significant demand bonuses or penalties under VMC or IMC. As with the previous SFO model, the switch proximity variables showed statistically significant effects. Here, they converged for configuration switches between 28R,28L|1R,1L-SIDEBY, 28R,28L|1R,1L-STAG, 28L|1R,1L, and 28R|1R,1L. This implies that switches between these configurations will occur with a higher probability than between the others. Intuitively, this makes sense because the predominant operational runway structure at SFO is handling arrivals on runway 28 and handling departures on runway 1. These configurations all hold this same structure, so switches between them would be operationally easy. Additionally, as before, the noise parameters indicate that flights over the water during the evening hours are preferred to flights departing over nearby communities.

Table 4.4.2: Estimated utility function weights for SFO utilities designating side-by and staggered.

Parameters	Value	Std. error	t-statistic
Inertia parameters			
Config. 19R,19L|10R,10L	6.94	0.481	14.42
Config. 28R,28L|1R,1L - SIDEBY	5.12	0.102	50.07
Config. 28R,28L|1R,1L - STAG	5.12	0.102	50.07
Config. 28R,28L|28R,28L	6.20	0.149	41.52
Config. 28R|1R,1L	5.12	0.102	50.07
Config. 28L|1R,1L	5.12	0.102	50.07
Wind parameters			
High headwind on arrival Runway	0.0944	0.0335	2.81
Normal headwind on arrival Runway	0.0944	0.0335	2.81
Tailwind on arrival Runway	-0.0944	0.0335	-2.81
Demand/visibility parameters			
Arrival demand for 28R,28L—1R,1L - SIDEBY under VMC	0.114	0.0371	3.07
Arrival demand for 28R,28L—1R,1L - STAG under IMC	0.105	0.0316	3.33
Switch proximity parameters			
28L|1R,1L to 28R,28L|1R,1L - SIDEBY	0.679	0.181	3.75
28R|1R,1L to 28R,28L|1R,1L - SIDEBY	0.679	0.181	3.75
28R,28L|1R,1L - SIDEBY to 28R,28L|1R,1L - STAG	1.36	0.133	10.20
28R,28L|1R,1L - SIDEBY to 28L|1R,1L	1.08	0.209	5.17
28R,28L|1R,1L - SIDEBY to 28R|1R,1L	1.08	0.209	5.17
28R,28L|1R,1L - STAG to 28L|1R,1L	1.70	0.167	10.13
28R,28L|1R,1L - STAG to 28R|1R,1L	1.70	0.167	10.13
28R,28L|1R,1L - STAG to 28R,28L|1R,1L - SIDEBY	1.65	0.138	11.93
Noise abatement parameters			
Arrivals on runways 28R and 28L between 23:00 and 7:00	0.531	0.308	1.73

Chapter 5

Discrete-Choice Prediction Models

5.1 3-Hour Forecast Using Actual Weather and Demand

The estimated utility functions from the discrete-choice model can be used within equations 2.1.6 - 2.1.9 to calculate the probability of choosing a runway configuration alternative during each 15-minute selection period. The utility functions require attributes such as the runway configuration in the previous time interval, as well as the wind conditions, weather conditions, and demand in the current time interval, to complete the calculations for every 15-minute selection period throughout the year. When predicting the runway configuration selection on a future time horizon, all possible evolutions of the runway configuration that could occur within that time horizon must be considered.

To consider all possible runway configuration evolutions, Bayes rule can be recursively applied using the runway configuration selection probabilities from each 15-minute interval in the specified time horizon. The model inputs include the actual arrival and departure demand, cloud ceiling, visibility conditions, wind speed, and wind direction for each 15-minute interval. The model outputs a probability of selection for each possible runway configuration alternative on the specified time horizon. The choice selection is taken as the runway alternative with the maximum

probability of being chosen. For example, consider predicting the runway configuration selection on a 3-hour time horizon at SFO. If the attributes of the current time at SFO, the runway configuration alternatives 28R,28L|1R,1L, 19R,19L|10R,10L, 28R,28L|28R,28L, 28R|1R,1L, and 28L|1R,1L may have selection probabilities of 5%, 75%, 5%, 7%, and 8% respectively, for the next 15-minute interval. The probabilities of being in each configuration 30 minutes (i.e., 2 time-periods) from now will have to be conditioned on the runway configuration in the next time period, and so on. In this manner, the probabilities of being in 28R,28L|1R,1L, 19R,19L|10R,10L, 28R,28L|28R,28L, 28R|1R,1L, or 28L|1R,1L 3 hours from now may have changed to 7%, 50%, 3%, 10%, and 30% respectively. Runway configuration 19R,19L|10R,10L is then taken as the 3-hour prediction because it has the highest probability.

The accuracy of the predictions are first evaluated assuming a perfect knowledge of the wind, visibility, and demand variables for the subsequent time horizon under consideration by using the meteorological and demand data from the 2012 ASPM data set. Note that the test sets from year 2012 are independent data sets from the training sets from year 2011. Accuracy is defined as the percentage of observed runway configurations modeled that were predicted correctly. In other words, if a runway configuration is observed 100 times throughout the year, and out of those times it was predicted correctly 70 times, the accuracy for that particular configuration would be reported as 70%. The overall accuracy of the entire model is the percentage of correct predictions throughout the year for all runway configurations modeled.

5.1.1 EWR

As shown in Table 5.1.1, the overall prediction accuracy for EWR in 2012 was 97.8% on a short-term 15-minute time horizon, 79.5% on a standard 3-hour time horizon, and 69.5% on a long-term 6-hour time horizon. The classification confusion matrix for EWR in 2012 for a standard 3-hour prediction horizon is shown in Figure 5.1.1. Note that the configurations which reported higher relative prediction accuracies had higher relative frequencies. Presumably, this phenomenon occurs because the utility parameters for configurations that are used more often than others have more accurate

estimations. Also, the low number of average switches per day at EWR (2.6 switches per day) increases the effects of inertia for more frequently used configurations.

Table 5.1.1: Prediction accuracy (using actual weather and demand) for EWR in 2012.

ID	Configuration	Frequency	Prediction Accuracy		
			15 min	3 hr	6 hr
1	22L,11|22R	3,705 (13.1%)	96.0%	61.3%	46.0%
2	22L|22R	13,245 (46.8%)	98.3%	83.5%	73.9%
3	22L|22R,29	147 (¡1%)	91.8%	36.7%	27.2%
4	4R,4L|4L	65 (¡1%)	90.8%	27.7%	4.6%
5	4R,11|4L	1,202 (4.2%)	93.7%	40.0%	13.1%
6	4R|4L	9,931 (35.1%)	98.3%	86.6%	68.6%
	Total	**28,295**	**97.8%**	**79.5%**	**69.5%**

Interestingly, on a 3 hour prediction horizon configuration 4R|4L had a higher accuracy than configuration 22L|22R (86.6% vs. 83.5%) even though configuration 4R|4L had a lower frequency than configuration 22L|22R (9,931 vs. 13,245). This effect does not last as the prediction horizon increases to 6 hours. As expected, the accuracies of all runway configuration predictions decreased as the prediction horizon increased. In the short term, the inertia parameter heavily influences the utility functions. Because configuration switches at EWR are infrequent, the short term accuracy is very high. As the time horizon increases, inertia becomes less important, and attributes such as wind, demand, and visibility have a larger influence on the predictions. Since the prediction model considers all possible evolutions of the possible runway configuration selection over time, the uncertainty present in the estimated utility functions gets magnified with time and lowers the prediction accuracy on longer time horizons.

Also, note that in 2011 configurations 4R,4L|4L and 22L|22R,29 were seen 2% and 1% of the year respectively. In 2012, they were both seen less than 1% of the selection periods throughout the year. Correspondingly, both had very low prediction accuracies at the 3 hour mark at 27.7% and 36.7% respectively. At the 6 hour mark 4R,4L|4L's prediction accuracy was only 4.6%. This demonstrates a major limitation when predicting the runway configuration selection using a discrete choice

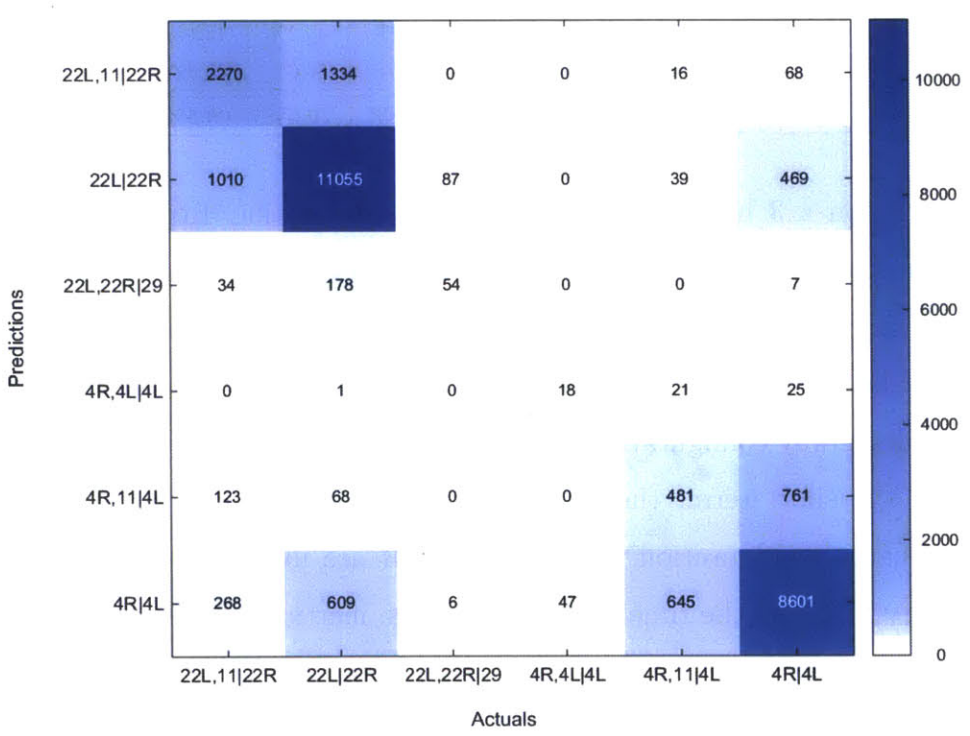

Figure 5.1.1: EWR classification confusion matrix for 3 hour time horizon (ASPM data).

methodology. Configurations that are seen very infrequently in the training data set will have poorly estimated utility functions which will result in a lower prediction accuracy in the test data set. When the number of occurrences are significantly lower in the test data set the effect is magnified because the estimated inertia and switch proximity parameters (which are already highly uncertain) overstate the preferences for the incumbent configuration staying or switching. New airport procedures, runway maintenance schedules, or capacity enhancements can cause significant shifts in the relative frequencies of the runway configurations from year to year. In these situations, the models would likely have to be re-estimated to stay accurate.

Figure 5.1.1 shows that configurations 4R,4L|4L and 4R,11|4L typically get confused with configuration 4R|4L. Similarly, configurations 22L,11|22R and 22L|22R,29 typically get confused for configuration 22L|22R. These two groups have very similar inertia, switch, and wind parameter effects. When the wind is favorable for either group, the choice of configuration is primarily driven by demand and VMC/IMC. The demand and visibility attributes predict well in extreme cases, however under normal conditions the EWR model still tends to over predict the high frequency configurations 4R|4L and 22L|22R.

To validate the benefit of the proposed models of the runway configuration selection, the overall accuracies of the discrete-choice utility prediction models was compared with a baseline heuristic. Operationally, since the runway configuration changes only occur a handful of times during a day, a baseline heuristic that assumes the airport remains in the current configuration for the allotted time-horizon is relevant. A comparison of the discrete-choice prediction model for EWR against the baseline heuristic prediction model for EWR has been shown for all 15-minute time intervals up to a 6-hour time horizon in Figure 5.1.2. As shown, in short term horizons when the airport usually stays in the same configuration, the accuracy of the two models are comparable. However, as the forecast horizon increases, the difference between the two models increases: At a 3-hour forecast horizon, there is a 3.3% improvement in the performance of the discrete-choice model, and at a 6-hour forecast horizon, the discrete-choice model outperforms the baseline heuristic by more than

7.3%.

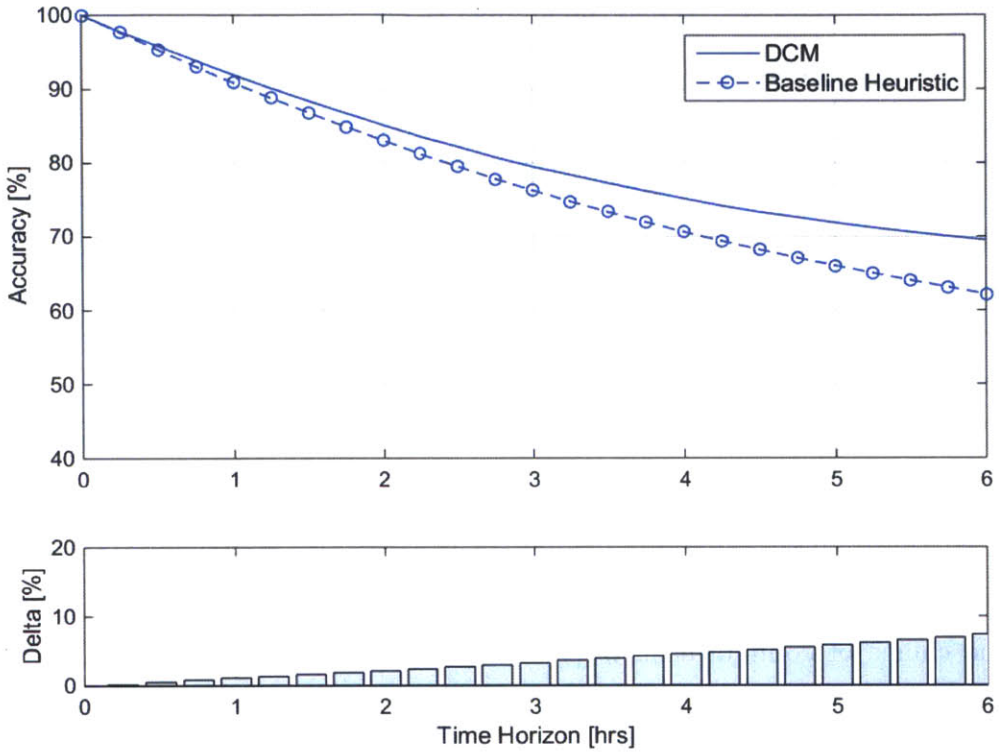

Figure 5.1.2: Comparison with baseline heuristic for EWR in 2012.

5.1.2 JFK

Table 5.1.2 shows the overall prediction accuracy for JFK in 2012 using the combined configuration model. The classification confusion matrix for JFK on a 3 hour horizon is shown in Figure 5.1.3. On a short term 15-minute prediction horizon, a standard 3-hour prediction horizon, and a long-term 6 hour prediction horizon, the model achieved an accuracies of 93.2%, 63.8%, and 48.2% respectively.

As with the EWR model, configurations that occurred more frequently throughout the year had higher individual prediction accuracies and longer time horizons had lower prediction accuracies. Configuration 31 Runways occurred the most during 2012, and as a result always had the highest individual configuration accuracy. Interestingly, configuration 22 Runways High Arrival had the lowest accuracy on a 6

hour time horizon even though it had one of the highest number of occurrences.

Overall, the accuracy of the JFK model is much lower than the accuracy of the EWR, LGA, and SFO models. On average, JFK has approximately 4.8 runway configuration switches per day - significantly more than EWR, LGA, or SFO. The high importance of the inertia variable in the short term, reduces the accuracy of the discrete choice models in situations when the meteorological or demand conditions do not highly favor a runway configuration switch. In many cases, JFK's switches are not under conditions that favor a certain runway configuration switch, reducing the models accuracy more quickly than the other models. While there seems to be a high amount of misclassification's, the confusion matrix in Figure 5.1.3 shows that typically the very similar configurations such as the 22's and the 4's are mistaken for one another which is a positive sign because either captures the overall runway being used at the airport which can still be useful within ATC decision support tools and maintenance support tools for ground crews.

Table 5.1.2: Prediction accuracy (using actual weather and demand) for JFK in 2012.

ID	Configuration	Frequency	Prediction Accuracy		
			15 min	3 hr	6 hr
1	13 Runways	3,109 (9.9%)	95.8%	54.8%	33.5%
2	22 Runways High Arrival	5,660 (18.1%)	95.8%	58.1%	28.0%
3	22 Runways	3,471 (11.1%)	94.3%	53.6%	53.2%
4	31 Runways	11,099 (35.5%)	97.3%	70.5%	57.9%
5	4 Runways High Arrival	5,983 (19.1%)	97.1%	68.9%	52.9%
6	4 Runways	1,955 (6.3%)	93.2%	58.5%	50.5%
	Total	**31,277**	**96.3%**	**63.8%**	**48.2%**

Fig 5.1.4 shows the benefit of the JFK against the baseline heuristic described above. The benefit between the short term and the long term models is practically the same, but as the prediction horizon increases the model starts to show a beneficial improvement: At a 3-hour forecast horizon, there is a 2.7% improvement in the performance of the discrete-choice model, and at a 6-hour forecast horizon, the discrete-choice model outperforms the baseline heuristic by more than 3.7%. Considering the high relative average number of switches per day at JFK, one would expect a larger benefit against the baseline heuristic (which is the case with the LGA

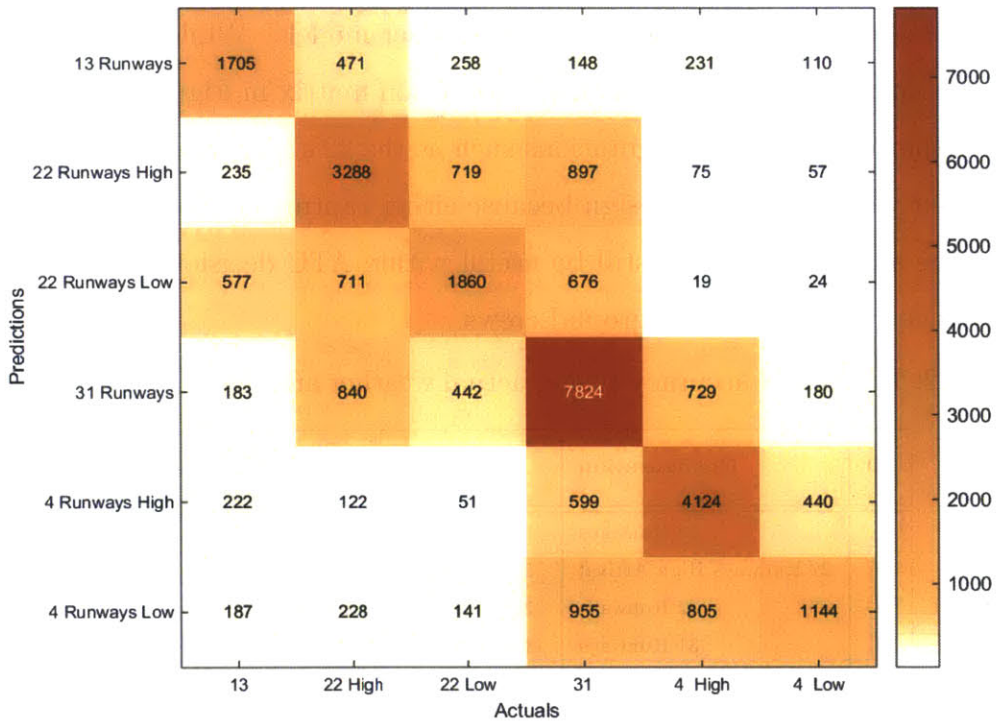

Figure 5.1.3: JFK classification confusion matrix for 3 hour time horizon (ASPM data).

model shown later). At JFK, this is not the case and is likely because these switches can appear somewhat unintuitive using only the ASPM data. Future models of JFK could benefit from adding additional sources of data that incorporate things such as wind gusts or aircraft size into the model. Additional data could give insight on what is driving some of the switches that appear to be "random" only given the ASPM data. It is also interesting that after the 5-hour prediction horizon, the accuracy of the discrete-choice model and the baseline accuracy begin to converge.

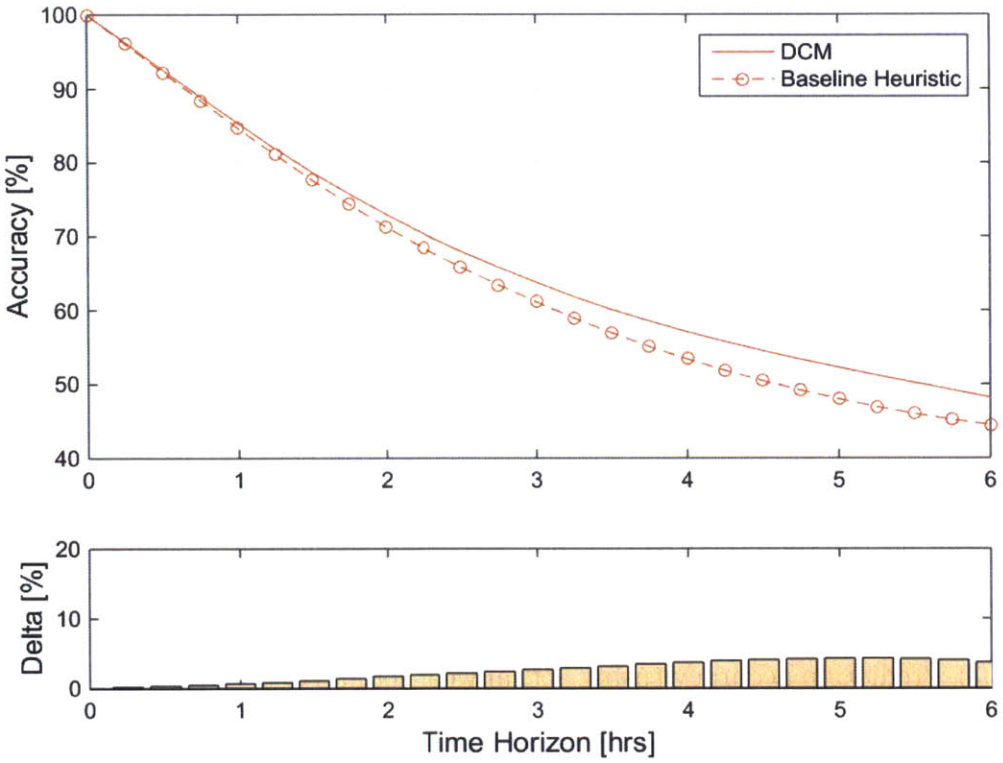

Figure 5.1.4: Comparison with baseline heursitic for JFK in 2012.

5.1.3 LGA

As shown in Table 5.1.3, the overall prediction accuracy for LGA in 2012 was 98.0% on a short-term 15-minute time horizon, 81.3% on a standard 3 hour time horizon, and 72.4% on a long-term 6 hour time horizon. The classification confusion matrix for LGA in 2012 for a standard 3 hour prediction horizon is shown in Figure 5.1.5.

These results are very promising considering prior research using logistic regression models for LGA without any look-ahead, achieved a prediction accuracy of 75% at LGA [16].

Similar to the EWR and JFK prediction models, configurations that were seen more often had higher relative prediction accuracies and the prediction accuracies typically lowered as the time horizon was increased. Configurations 22|13 and 31|4 were used the most frequently at LGA in 2012 having frequencies greater than 24% of the total selection periods and had accuracies approaching 90% on a 3 hour time horizon. Conversely, configurations 22|31, 22,31|31, 31|31, and 4|4 were each seen less than 10% of the decision selection periods throughout the year and had prediction accuracies from 65%-70% on a 3 hour time horizon. Interestingly, the prediction accuracy for configuration 31|4 increases from 89.5% on a 3 hour horizon to 90.9% on a 6 hour horizon. This may be because the model tends to bias the choice selection away from the much less frequent and lower capacity 31|31 and 4|4 configurations when arrival demand increases.

The confusion matrix in Figure 5.1.5 suggests that on a 3 hour prediction horizon, most runway configurations were correctly classified. Typically when runway configurations were not predicted correctly, they were misclassified as configurations with either a common arrival or common departure runway.

Table 5.1.3: Prediction accuracy (using actual weather and demand) for LGA in 2012.

ID	Configuration	Frequency	Prediction Accuracy		
			15 min	3 hr	6 hr
1	22|13	7,626 (29.1%)	98.3%	88.1%	85.7%
2	22|31	2,715 (10.3%)	97.1%	69.4%	47.6%
3	22,31|31	1,721 (6.6%)	96.8%	67.4%	46.0%
4	31|31	2,391 (9.1%)	97.3%	70.1%	37.1%
5	31|4	6,457 (24.6%)	98.6%	89.5%	90.9%
6	4|13	4,475 (17.1%)	98.1%	79.0%	73.1%
7	4|4	875 (3.3%)	96.9%	65.8%	42.0%
	Total	**26,260**	**98.0%**	**81.3%**	**72.4%**

Fig 5.1.6 shows the benefit of the LGA against the baseline heuristic described above. The same overall trend is seen at LGA as with the other prediction models.

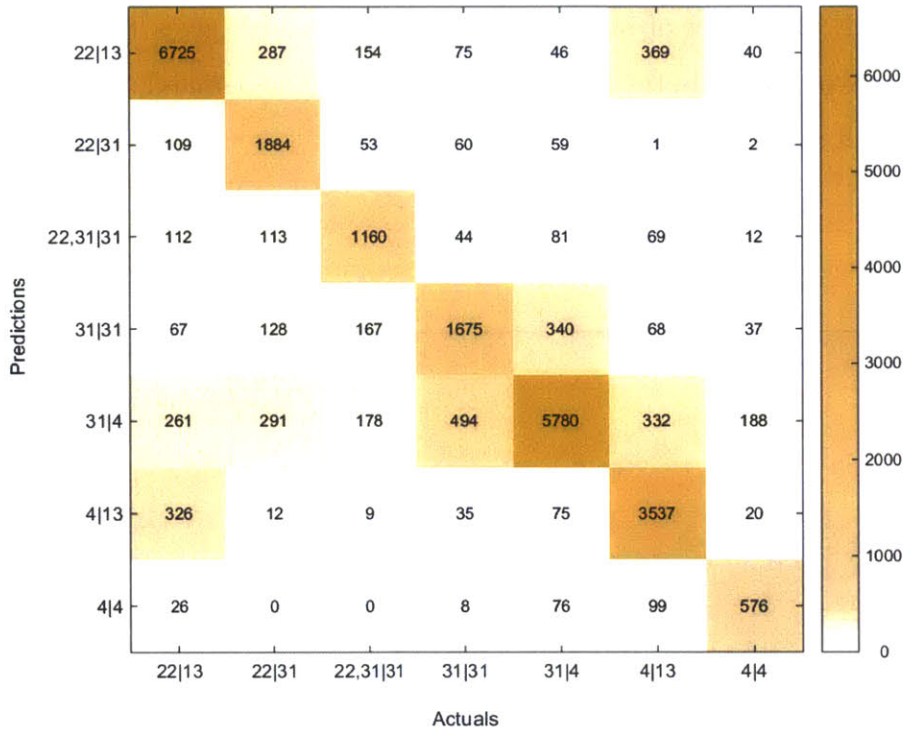

Figure 5.1.5: LGA classification confusion matrix for 3 hour time horizon (ASPM data).

In short term horizons when the airport usually stays in the same configuration, the accuracy of the two models are comparable. However, as the forecast horizon increases, the difference between the two models increases: At a 3-hour forecast horizon, there is a 7.5% improvement in the performance of the discrete-choice model, and at a 6-hour forecast horizon, the discrete-choice model outperforms the baseline heuristic by more than 17.4%. The LGA prediction model beats the baseline heuristic by a higher percentage than the other prediction models. This may be because LGA has more frequent switches (2.7 switches per day) than at SFO and EWR without reaching the high frequency of switches seen at JFK, which gives the model more opportunity to exploit attributes other than inertia such as demand, wind, visibility, coordination, and noise when predicting.

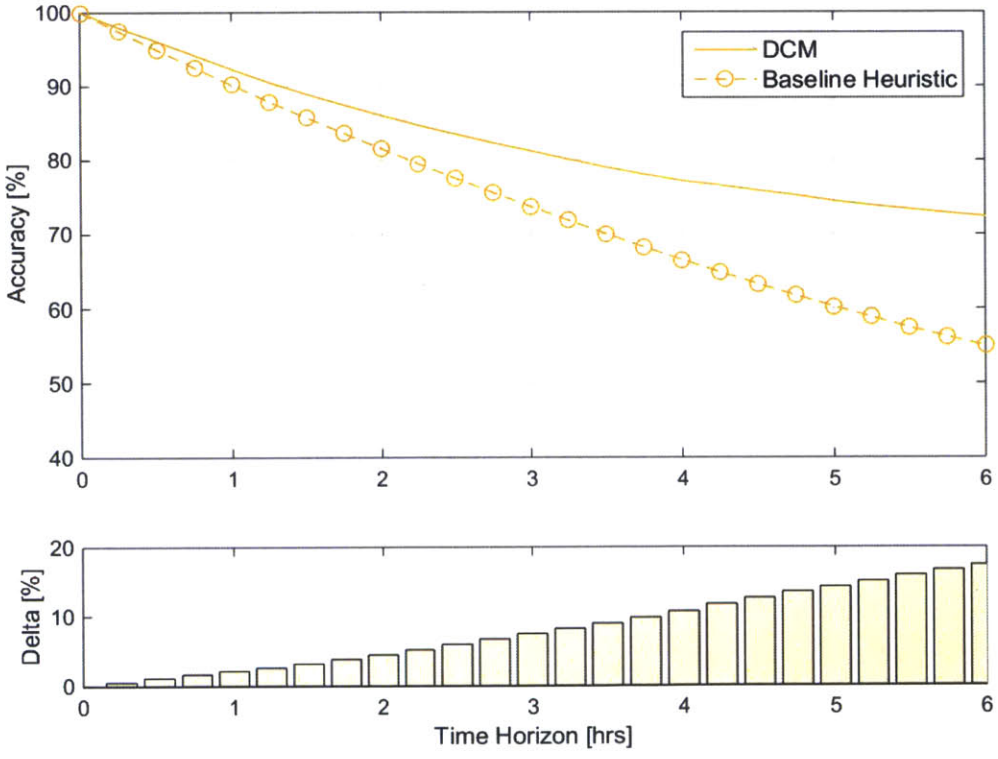

Figure 5.1.6: Comparison with baseline heuristic for LGA in 2012.

5.1.4 SFO

As shown in Table 5.1.4, the overall prediction accuracy for SF0 in 2012 was 98.1% on a short-term 15-minute time horizon, 82.8% on a standard 3-hour time horizon, and 72.8% on a long-term 6-hour time horizon. The classification confusion matrix for SFO in 2012 for a standard 3-hour prediction horizon is shown in Figure 5.1.7. Unlike the other models, relative frequency did not directly correlate with higher accuracy at SFO, which could be an effect of the relatively low average number of configuration switches per day at SFO (2.1 switches per day). The confusion matrix shows that the similar configurations 28R,28L|1R,1L, 28R|1R,1L, and 28L|1R,1L commonly got confused for one another. The bias is heavily weighted toward predicting configuration 28R,28L|1R,1L because it occurred much more frequently throughout 2012. This result is reflective of some of the modeling challenges present when predicting the runway configuration at SFO.

A major challenge when modeling SFO is predicting between 28R,28L|1R,1L with arrivals on the closely spaced parallel runways, and the single arrival runway configurations, 28R|1R,1L and 28L|1R,1L. As mentioned before, simultaneous (side-by) landings are not possible under IMC, and the airport operates almost as it would in a single arrival runway configuration, even in 28R,28L|1R,1L. The reported configurations in the ASPM data set do not differentiate between simultaneous and staggered parallel approaches, even though staggered approaches have a capacity that would be closer to 28R|1R,1L or 28L|1R,1L. This fact, along with the other similarities between these two runway configuration alternatives, makes it difficult to predict either of these alternatives accurately without introducing a selection bias. As stated above, predictions using the discrete-choice model which attempted to classify "side-by" and "staggered" configurations using the AARs given in the ASPM datasets is not shown due to lack of any ground truth in the data. Additionally, accurately predicting 28R,28L|28R,28L is also challenging due to the limitations from the ASPM dataset. This runway configuration is typically only used for long-haul departures over the Pacific Ocean and to Hawaii, and the aggregate flight counts in ASPM are not sufficient

to account for this factor. Despite these challenges, the overall prediction accuracy in 2012 for SFO is promising and opens the door for future models that may potentially include these types of data.

As with the other models, short prediction horizons were very accurate and the longer horizons had lower relative prediction accuracies. Over time, configurations 19R,19L|10R,10L, 28R,28L|1R,1L, and 28R,28L|28R,28L all held relatively high prediction accuracies. The two low capacity configurations 28R|1R,1L and 28L|1R,1L accuracies dropped significantly over time. This is partly due due to their relatively low number of occurrences and because they get confused for one another during prediction (see Figure 5.1.7). Interestingly, configuration 19R,19L|10R,10L which only occurred 3.6% of the selection periods in 2012 had the highest prediction accuracy over all time horizons tested. It is very likely that configuration 19R,19L|10R,10L is the preferred configuration under poor weather conditions because the other four configurations all align into the wind very similarly. This suggests that when unfavorable weather patterns approach SFO, the preferred configuration switch is to 19R,19L|10R,10L.

Table 5.1.4: Prediction accuracy (using actual weather and demand) for SFO in 2012.

ID	Configuration	Frequency	Prediction Accuracy		
			15 min	3 hr	6 hr
1	19R,19L|10R,10L	1,072 (3.6%)	99.3%	95.9%	95.5%
2	28R,28L|1R,1L	21,007 (71.6%)	98.7%	87.0%	78.6%
3	28R,28L|28R,28L	2,389 (8.1%)	98.2%	88.7%	84.7%
4	28R|1R,1L	921 (3.1%)	93.2%	50.5%	44.6%
5	28L|1R,1L	3,958 (13.5%)	96.0%	60.5%	35.0%
	Total	**29,348**	**98.1%**	**82.8%**	**72.8%**

Figure 5.1.8 shows the benefit of the SFO against the baseline heuristic. In short term horizons when the airport usually stays in the same configuration, the accuracy of the two models are comparable. However, as the forecast horizon increases, the difference between the two models increases: At a 3-hour forecast horizon, there is a 3.1% percentage point improvement in the performance of the discrete-choice model, and at a 6-hour forecast horizon, the discrete-choice model outperforms the base-

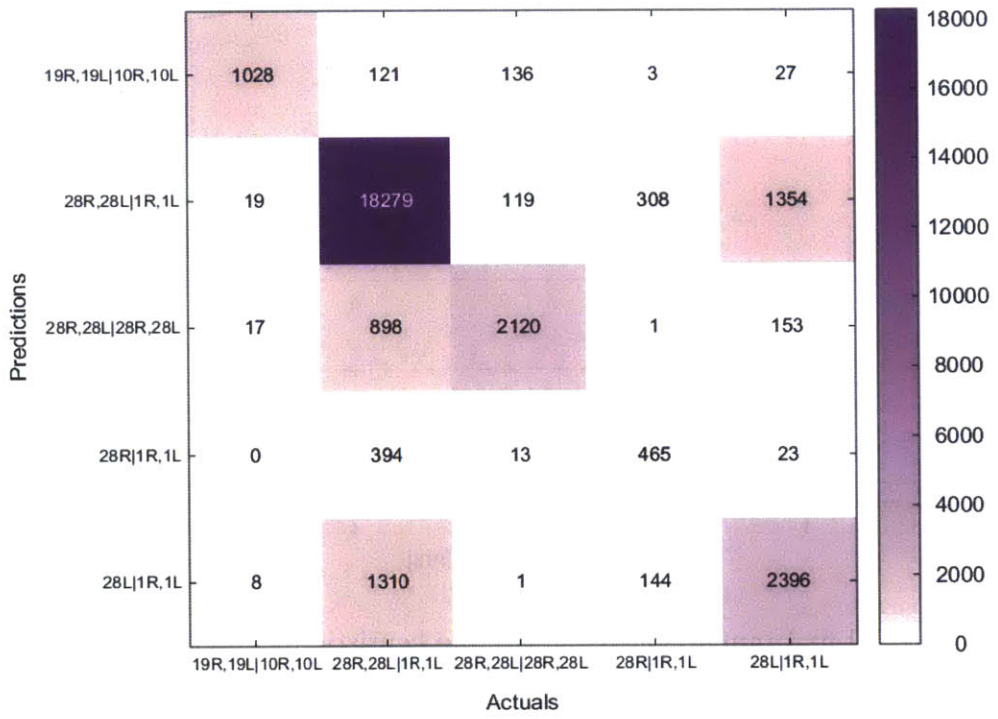

Figure 5.1.7: SFO classification confusion matrix for 3 hour time horizon (ASPM data).

line heuristic by more than 4.0% percentage points. The SFO model's improvement against the baseline heuristic is lower than with the other models because configuration 28R,28L|1R,1L was used 71.6% of the selection periods throughout 2012 which keeps the baseline heuristic fairly accurate in most cases.

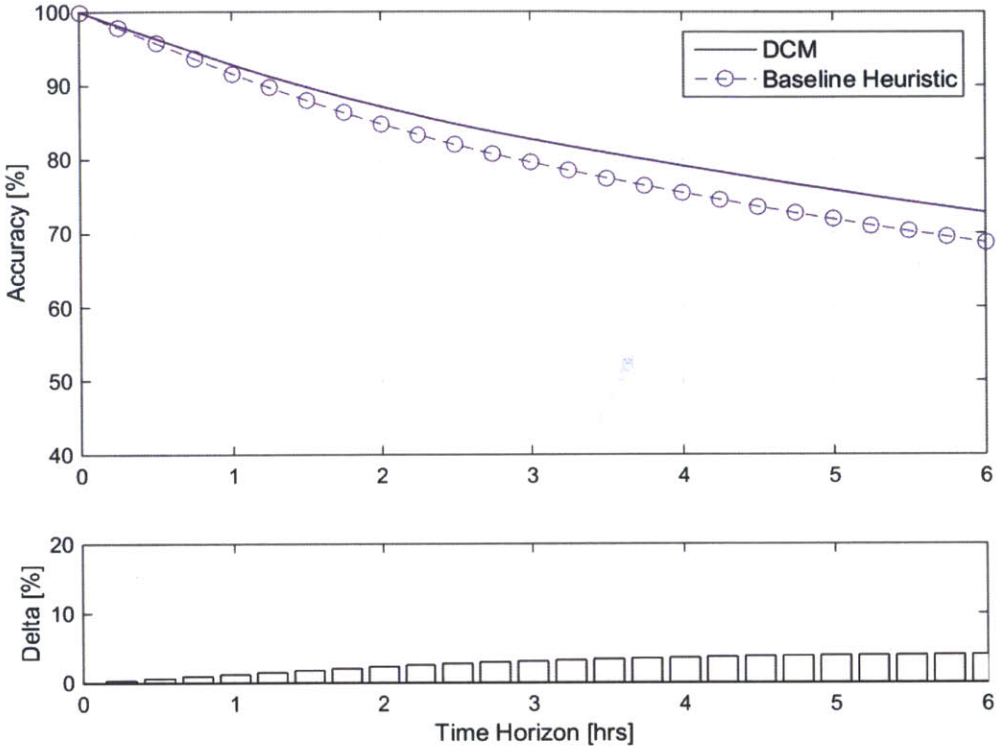

Figure 5.1.8: Comparison with baseline heuristic for SFO in 2012.

5.2 3-Hour Forecast Using Weather and Demand Forecast Data

5.2.1 Data Pre-Processing

The discrete choice models developed in this study are also tested on Terminal Aerodrome Forecast (TAF) data in order to test their predictive capability on data that is not perfectly precise. This corresponds to a practical implementation, where the actual values of the attributes for the utility functions used within the prediction

model are not available, and weather forecast data and schedule data must be used instead. Wind speed, wind direction, and visibility conditions was taken from the TAF data for the prediction models using forecasts. Unfortunately, the TAF data has many gaps, where the quality of the weather forecasts made available to air traffic control personnel may be several minutes or several hours old. When this was observed, the TAF data points were selected to replicate what forecasts would have actually been available to air traffic control in real time. Additionally, future aircraft arrival demand would not realistically be available to air traffic controllers for future predictions. Therefore, when predicting the runway configuration using forecast data, scheduled demand data is used in place of active arrival demand data. The scheduled demand data used is taken from the ASPM data set.

5.2.2 EWR TAF Results

The overall accuracy using the TAF data for EWR is shown in Table 5.2.1 for prediction horizons of 3 hours, 6 hours, 9 hours, 12 hours, and 15 hours. As shown, the overall accuracy using TAF forecast data on a standard 3 hour prediction horizon was 78.9% for EWR in 2012. The accuracy of the EWR model at the 3 hour prediction horizon using TAF forecast data is only 0.6 percentage points lower than the prediction model using actual weather and demand data. It is promising that the accuracy of the model is not substantially degraded by using forecast data, which is inherently prone to error. Note that configuration 22L|22R,29, which was seen much less than 1% of the year according to the ASPM data, was counted as a reporting error when using the TAF forecast data and the availability criterion. This means that the TAF data forecast wind conditions that were outside of range for any feasible use of configuration 22L|22R,29.

5.2.3 LGA TAF Results

The overall accuracy using the TAF data for LGA is shown in Table 5.2.2 for prediction horizons of 3 hours, 6 hours, 9 hours, 12 hours, and 15 hours. The overall

Table 5.2.1: Prediction accuracy (using forecast weather and scheduled demand data) for EWR in 2012.

ID	Configuration	Frequency	Prediction Accuracy				
			3 hr	6 hr	9hr	12hr	15hr
1	22L,11|22R	3,530 (12.6%)	60.7%	45.4%	22.8%	17.2%	52.5%
2	22L|22R	13,345 (47.6%)	83.2%	67.0%	63.1%	58.0%	37.2%
3	22L|22R,29	0 (0%)	N/a	N/a	N/a	N/a	N/a
4	4R,4L|4L	65 (¡1%)	27.7%	4.6%	0%	0%	0%
5	4R,11|4L	1,123 (4.0%)	41.2%	27.1%	2.6%	4.9%	76.7%
6	4R|4L	9,959 (35.5%)	84.1%	72.6%	61.0%	55.4%	46.0%
	Total	**28,022**	**78.9%**	**64.5%**	**55.7%**	**49.1%**	**43.8%**

accuracy using TAF forecast data on a standard 3 hour prediction horizon was 78.9% for LGA in 2012. The accuracy of the LGA model using TAF forecast data is only 2.4 percentage points lower at the 3 hour mark than the prediction model using actual weather and demand data. On a very long 15 hour prediction horizon, the accuracy of the model is primarily driven by the changes in forecast weather conditions and the accuracy of the model is degraded to 53.7%. The uncertainty in the forecast is magnified over longer time horizons, but it is very promising that the accuracy of the models is not significantly lower than the models using actual weather and demand data. For comparison, other research models for LGA achieved a prediction accuracy at 75% using logistic regression and no look ahead [16]. As with LGA model predicted on the ASPM data, the accuracy of specific configurations increased as the frequency of their occurrence increased with one exception. Runway configuration 31|4's prediction slightly increased as the time horizon increased. This is likely for the same reason described in the LGA model using actual weather and demand data - the model's estimated utilities are such that it will bias the choice selection away from the low capacity and less frequently observed configurations 31|31 and 4|4 when arrival demand increases. Over long time horizons, this effectively lowers the accuracy of 31|31 and 4|4 and does not degrade the accuracy of 31|4.

Table 5.2.2: Prediction accuracy (using forecast weather and scheduled demand data) for LGA in 2012.

ID	Configuration	Frequency	Prediction Accuracy				
			3 hr	6 hr	9 hr	12 hr	15 hr
1	22|13	7,360 (28.3%)	77.4%	77.9%	74.1%	72.8%	69.1%
2	22|31	2,645 (10.2%)	73.8%	47.8%	30.3%	22.5%	14.6%
3	22,31|31	1,679 (6.5%)	68.2%	47.7%	30.3%	19.4%	13.2%
4	31|31	2,471 (9.5%)	65.1%	33.1%	28.9%	26.7%	25.2%
5	31|4	6,401 (24.6%)	77.6%	87.0%	88.4%	88.5%	87.4%
6	4|13	4,533 (17.5%)	76.2%	61.7%	48.0%	42.4%	40.8%
7	4|4	885 (3.4%)	61.2%	35.8%	28.2%	23.2%	22.1%
	Total	**25974**	**78.9%**	**66.6%**	**59.9%**	**56.7%**	**53.7%**

5.2.4 SFO TAF Results

The overall accuracy using the TAF data for SFO is shown in Table 5.2.3 for prediction horizons of 3 hours, 6 hours, 9 hours, 12 hours, and 15 hours. As shown, the overall accuracy using TAF forecast data on a standard 3 hour prediction horizon was 80.8% for SFO in 2012 and is only 2.0 percentage points lower on a 3 hour time horizon than the prediction model using ASPM data. Again, the accuracy of specific configurations increased as the frequency of their occurrence increased.

The TAF model results for EWR, JFK, LGA, and SFO all suggest that the prediction models using ASPM data are not significantly degraded when using real-time forecasts. This is a very promising result because, realistically, future decision support tools which use the estimated utility functions from the discrete-choice models will not have perfectly precise data as an input. The TAF analysis shows that forecast data will still provide a reasonably high level of prediction accuracy and, therefore, can still be viable for decision support tools in the field.

Table 5.2.3: Prediction accuracy (using forecast weather and scheduled demand data) for SFO in 2012.

ID	Configuration	Frequency	Prediction Accuracy				
			3 hr	6 hr	9hr	12hr	15hr
1	19R,19L|10R,10L	1,063 (3.6%)	87.0%	77.2%	45.0%	73.4%	24.1%
2	28R,28L|1R,1L	21,362 (71.8%)	87.5%	80.2%	76.4%	73.9%	75.2%
3	28R,28L|28R,28L	2,840 (9.5%)	67.8%	49.7%	31.4%	43.2%	28.9%
4	28R|1R,1L	918 (3.1%)	50.1%	44.5%	40.8%	37.8%	36.0%
5	28L|1R,1L	3,566 (12.0%)	57.2%	33.6%	19.8%	15.6%	17.7%
	Total	**29,750**	**80.8%**	**70.3%**	**64.3%**	**62.7%**	**61.7%**

Chapter 6

Modeling the New York Metroplex

6.1 Introducing the New York Metroplex Models

Because the New York Metroplex is very congested, a high level of coordination is required between EWR, JFK, and LGA. It is relevant to discuss characterizing and predicting the state of the entire New York Metroplex as a whole, particularly because improving the capacity utilization of the New York Metroplex is one area of focus for NextGen research.

This paper describes two different approaches to predicting the New York Metroplex runway configurations (mentioned in Section 1). The first approach, denoted the Configuration Model, estimates a new discrete-choice model for the runway configurations of the entire New York Metroplex as a whole. It then predicts the New York Metroplex runway configuration using the same method used in the individual airport models. The second approach, denoted the Stackedd Model, combines the previously shown individual airport discrete-choice models for EWR, JFK, and LGA to predict the runway configuration for the entire New York Metroplex.

6.2 New York Metroplex Configuration Model

6.2.1 Utility Function Estimation

The discrete-choice model for the New York Metroplex configurations was taken as a nested logit structure with a nest containing the South Flow configurations, a nest containing the North Flow configurations, and a nest containing the Mixed Flow configurations. The South Flow and North Flow nests have scale parameters of $\mu_S = 1.03$ and $\mu_N = 1.13$, respectively.

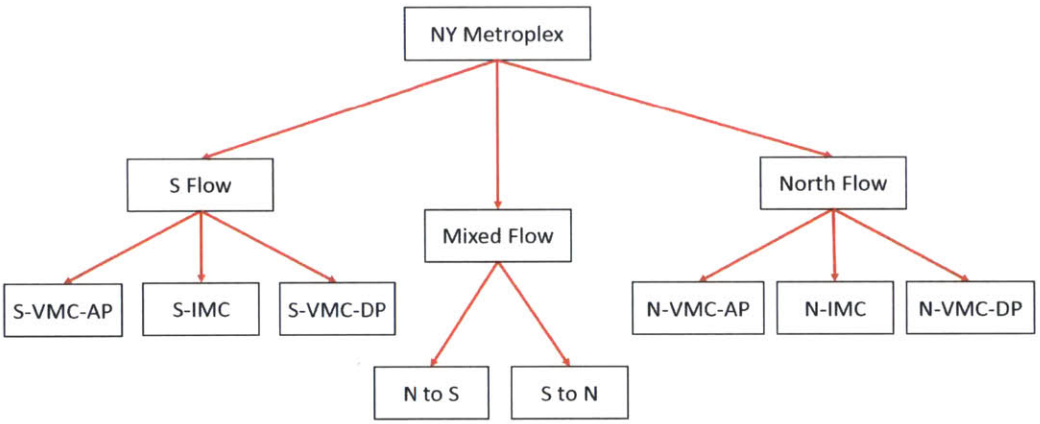

Figure 6.2.1: New York Metroplex Configuration Model specification.

The discrete-choice model is estimated for each 15-minute interval in the data set. The estimated values of the attribute weights, their standard errors, and their t-statistics are shown in Tables 6.2.1 - ??. Similar to the individual airport models, the inertia variables for the New York Metroplex have the largest statistical significance compared with the other attributes; however, the relative difference is not nearly as large as with the individual airport models. Other parameters, such as wind, demand, and VMC/IMC, are rather close to the level of significance of the inertia variables in the New York Metroplex models. This effect is likely a result of the frequent switching that occurs between New York Metroplex configurations (average 5.7 switches per day).

The wind variables are also very significant, particularly the negative penalty associated with tailwinds. The high headwind and normal headwind weights indicate

that there is a slightly lower preference to operate in high-headwinds rather than normal-headwinds. This is likely driven by the compression effect mentioned in the individual airport models. Note that the headwind variables in this model were chosen to capture the effects of the overall flow of the New York Metroplex instead of including wind parameters for each active runway at EWR, JFK, and LGA.

Both the total arrival demand and the total departure demand from EWR, JFK, and LGA have a very significant effect on the New York Metroplex configuration. The combined effect on the New York Metroplex appears to be more significant than individual demand effects at EWR, JFK, or LGA. The utility weights confirm the separation of New York Metroplex configurations into "Arrival Priority" and "Departure Priority" classes. Namely, arrival priority configurations received a positive utility benefit from high arrival demands and all departure priority configurations received a bonus for high departure demands. These demand variables also had high t-statistics. Interestingly, the Mixed Flow S to N configuration converged for total departure demand in 2011. It may be possible that departures are sometimes given priority when the overall airspace flow is changing, which is likely happening during a Mixed Flow configuration.

The Mixed Flow configurations both received a positive bonus under VMC, which may indicate that VMC is preferred if the arriving and departing aircrafts are flying in opposite directions. Also, as expected, both IMC configurations received a positive bonus under IMC.

The switch proximity variables show that the South Flow configurations are much more likely to switch with other South Flow configurations or the Mixed Flow S to N configuration. Similarly, the North Flow configurations are much more likely to switch with other North Flow configurations or the Mixed Flow N to S configuration. This makes sense because it likely requires a large amount of coordination between airports to switch from a North Flow configuration to a South Flow configuration and vice versa.

Table 6.2.1: Estimated utility function weights for New York Metroplex Configuration Model - Part I.

Parameters	Value	Std. error	t-statistic
Inertia parameters			
Config. S-VMC-AP	5.02	0.262	19.15
Config. S-IMC	6.77	0.376	17.98
Config. S-VMC-DP	6.82	0.588	11.60
Config. Mixed Flow S to N	6.74	0.450	14.97
Config. Mixed Flow N to S	5.90	0.473	12.47
Config. N-VMC-AP	6.41	0.338	18.96
Config. N-IMC	5.05	0.429	11.76
Config. N-VMC-DP	7.06	0.627	11.25
Wind parameters			
High headwind	0.0160	0.008	2.00
Normal headwind	0.0170	0.006	2.64
Tailwind	-0.125	0.011	-11.37
Demand parameters			
S-VMC-AP total arrival demand for EWR, JFK, and LGA	0.116	0.0107	10.87
N-VMC-AP total arrival demand for EWR, JFK, and LGA	0.0797	0.01	7.99
N-IMC total arrival demand for EWR, JFK, and LGA	0.0726	0.0124	5.87
S-VMC-DP total departure demand for EWR, JFK, and LGA	0.105	0.0105	9.95
S-IMC total departure demand for EWR, JFK, and LGA	0.0798	0.0107	7.47
Mixed Flow S to N total departure demand for EWR, JFK, and LGA	0.0320	0.0136	2.36
N-VMC-DP total arrival departure for EWR, JFK, and LGA	0.0368	0.01	3.81
VMC/IMC parameters			
VMC on Mixed Flow S to N	0.523	0.269	1.94
VMC on Mixed Flow N to S	1.88	0.210	8.94
IMC on S-IMC	2.29	0.403	5.67
IMC on N-IMC	1.87	0.411	4.54
Switch proximity parameters			
Mixed Flow S to N to S-VMC-DP	2.70	0.468	5.77
S-VMC-DP to Mixed Flow S to N	4.18	0.624	6.70
Mixed Flow N to S to N-VMC-AP	2.09	0.546	3.83
N-VMC-AP to N-VMC-DP	3.33	0.396	8.41
N-VMC-DP to N-VMC-AP	3.50	0.661	5.30

6.2.2 Prediction

As shown in Table 6.2.2, the overall prediction accuracy for the New York Metroplex Configuration Model in 2012 was 96.1% on a short-term 15-minute time horizon, 69.0% on a standard 3 hour time horizon, and 60.0% on a long-term 6 hour time horizon. The classification confusion matrix for the New York Metroplex Configuration

Model in 2012 for a standard 3 hour prediction horizon is shown in Figure 6.2.2.

As with the individual models, configurations that occurred more frequently had a higher accuracy during prediction. This effect is not as significant as with the other models because the frequencies of configurations in the New York Metroplex were all comparable. In the short term, the accuracy of the prediction model is still very high; however it falls quickly on longer time horizons. Uncertainties and errors present within the data at each of the three airports are magnified when modeling the combined New York Metroplex. The confusion matrix in Figure 5.1.5 suggests that on a 3 hour prediction horizon, the New York Metroplex configurations were typically confused among runway configurations in the same nest. The South Flow configurations were almost exclusively confused with other South Flow configurations and some were confused with the Mixed with South Priority configuration. Similarly, the North Flow configurations were typically confused with other North Flow configurations and some were confused with the Mixed with a North Flow Priority configuration. This is a promising result since many situations require a prediction of the overall type of airspace flow the system will be in on a future time horizon, rather than adding more specificity with VMC/IMC and arrival or departure priority.

Table 6.2.2: Prediction accuracy (using actual weather and demand) for NY Metro. Configuration Model in 2012.

ID	Configuration	Frequency	Prediction Accuracy		
			15 min	3 hr	6 hr
1	S-VMC-AP	1,199 (10.4%)	93.0%	60.6%	43.5%
2	S-IMC	2,367 (20.5%)	94.3%	79.3%	74.2%
3	S-VMC-DP	981 (8.5%)	92.4%	50.1%	39.7%
4	Mixed S to N	1,009 (8.8%)	92.2%	41.0%	29.2%
5	Mixed N to S	1,393 (12.1%)	93.0%	53.1%	31.7%
6	N-VMC-AP	2,380 (20.7%)	94.5%	90.4%	88.8%
7	N-IMC	1,733 (15.0%)	94.6%	86.0%	77.3%
8	N-VMC-DP	459 (4.0%)	90.6%	14.6%	12.4%
	Total	**11,521**	**96.1%**	**69.0%**	**60.0%**

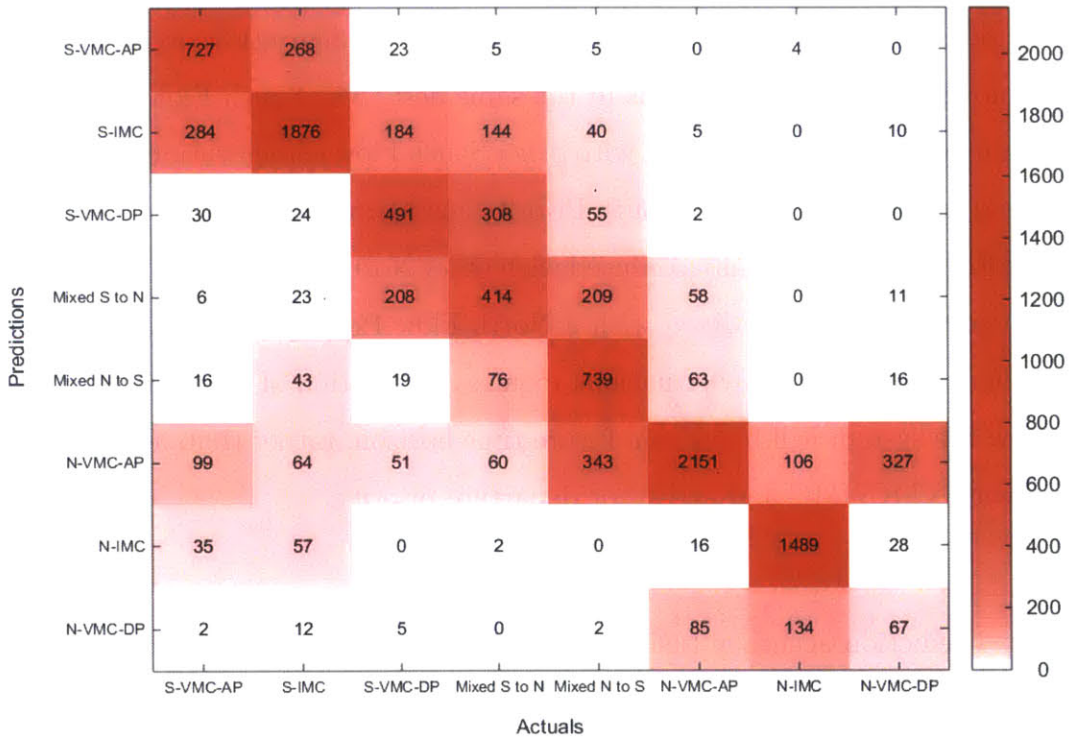

Figure 6.2.2: NY Metro. Configuration classification confusion matrix for 3 hour time horizon (ASPM data).

6.3 New York Metroplex Stacked Model

6.3.1 Prediction

The New York Metroplex Stacked Model predicts the runway configuration on a specified time horizon individually at EWR, JFK, and LGA using the individual airport discrete-choice models shown previously. It then combines each individual airport's runway configuration prediction into a prediction for the overall New York Metroplex configuration. To make for a relevant comparison, the New York Metroplex Stacked Model is run on the same filtered data set as the New York Metroplex Configuration Model. The New York Metroplex Stacked model also uses the same combined configurations used in the New York Metroplex Configuration Model (shown in Table 1.2.3).

Table 6.3.1 shows the prediction accuracies of the New York Metroplex Stacked Model. Figure 6.3.1 shows the corresponding classification confusion matrix for a 3 hour prediction horizon. Interestingly, the overall Stacked Model accuracy is much lower than the Configuration Model accuracy across the board. In the short term 15 minute prediction horizon, the Stacked Model only achieves an accuracy of 80.5% and in the long term 6 hour prediction horizon the accuracy plummets to 23.6%. The accuracies of each individual combined New York Metroplex configuration follow the same trends as with the other models; higher frequency configurations have higher accuracies and lower frequency configurations have lower accuracies. Since the New York Metroplex Stacked Model is limited by the configuration filtering cutoff of 1% at each individual airport rather than a cutoff of 1% for the entire New York Metroplex configurations, it has many more configurations that it can predict upon, and therefore a higher chance for error during each decision selection. Table 6.3.2 shows the number of combined New York Metroplex configuration predictions against the actual frequency for both the New York Metroplex Configuration Model and the New York Metroplex Stacked Model. Not surprisingly, the New York Metroplex Stacked Model under-predicts for all configurations (especially Mixed S to N) and has 49.2% (5,664 periods) of its data set predicted as "Other". The "Other" configuration refers

to times when the Stacked Model predicts a combined configuration that is not given in Table 1.2.3 and is therefore not used within the Configuration Model. For the same reasons as stated before, the major source of the New York Metroplex Stacked Model's inaccuracy stems from predicting these "Other" configurations since it has many more configuration possibilities. One could argue that since the data set is constrained to the filtered data from the New York Metroplex Configuration Model for comparison, the New York Metroplex Stacked model's overall accuracy is degraded. Initial tests using the entire data set and filtering by 1% at each individual airport in the New York Metroplex show that this is not the case, and resulted in the much lower overall accuracies of 27.1% using 26,500 decision periods. Threfore, for the purposes of comparison, the data set will remain consistent between both New York Metroplex Models.

Table 6.3.1: Prediction accuracy (using actual weather and demand) for NY Metro. Stacked Model in 2012.

ID	Configuration	Frequency	Prediction Accuracy		
			15 min	3 hr	6 hr
1	S-VMC-AP	1,199 (10.4%)	88.8%	35.0%	14.3%
2	S-IMC	2,367 (20.5%)	92.3%	45.4%	22.8%
3	S-VMC-DP	981 (8.5%)	72.3%	38.5%	19.7%
4	Mixed S to N	1,009 (8.8%)	74.5%	19.6%	5.7%
5	Mixed N to S	1,393 (12.1%)	78.5%	41.6%	28.1%
6	N-VMC-AP	2,380 (20.7%)	75.1%	36.0%	28.3%
7	N-IMC	1,733 (15.0%)	74.7%	39.8%	30.3%
8	N-VMC-DP	459 (4.0%)	85.6%	41.4%	35.5%
	Total	**11,521**	**80.5%**	**38.1%**	**23.6%**

Because the New York Metroplex Stacked model has more configurations it can predict upon, and therefore a lower accuracy than the New York Metroplex Configuration Model, it is relevant to examine the nature of the misclassified runway configuration predictions in the Stacked Model. Table 6.3.3 shows the number and percentage of decision selection periods in which the stacked model predicted 0 configurations correctly, 1 configuration correctly, 2 configurations correctly, and 3 configurations correctly between EWR, JFK, and LGA. Note that the 3 configurations predicted correctly percentage is not equal to the overall accuracy presented in Table

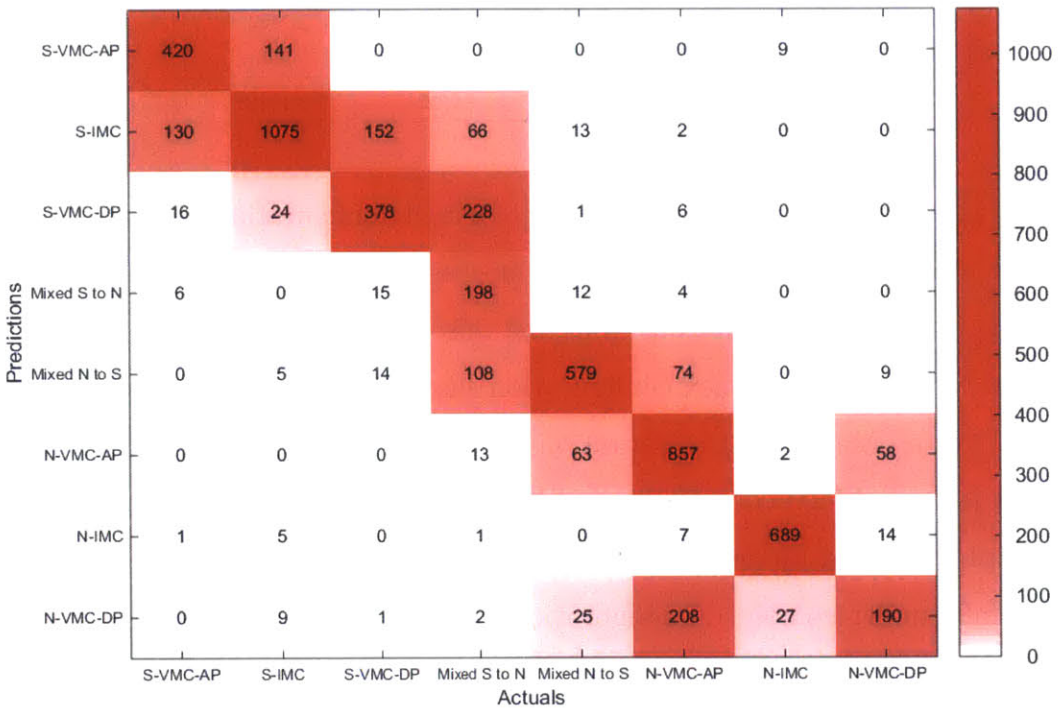

Figure 6.3.1: NY Metro. Stacked Model classification confusion matrix for 3 hour time horizon (ASPM data).

Table 6.3.2: Comparison between NY Metro. Configuration Model and NY Metro. Stacked Model predictions.

ID	Configuration	Frequency	Prediction Frequency	
			Config. Model	Stacked Model
1	S-VMC-AP	1,199 (10.4%)	1,032	570
2	S-IMC	2,367 (20.5%)	2,543	1,438
3	S-VMC-DP	981 (8.5%)	910	653
4	Mixed S to N	1,009 (8.8%)	929	235
5	Mixed N to S	1,393 (12.1%)	972	789
6	N-VMC-AP	2,380 (20.7%)	3,201	993
7	N-IMC	1,733 (15.0%)	1,627	717
8	N-VMC-DP	459 (4.0%)	307	462
0	N-VMC-DP	0 (0.0%)	0	5,664
	Total	**11,521**	**11,521**	**11,521**

6.3.1. This is because the final model uses combined configurations for prediction, and in many cases when the Stacked Model misclassified one airport's runway configuration, it does so in such as way that the misclassification still falls under the same overall New York Metroplex configuration shown in Table 1.2.3. For example, if the actual overall Metroplex Configuration is 22L|22R - 31L,31R|31L - 31|31 which corresponds to a combined configuration of Mixed N to S (Table 1.2.3) and the New York Metroplex Stacked Model predicts configuration 22L|22R - 31L,31R|31L - 31|4 which also falls under combined configuration Mixed N to S, it would be classified as a correct prediction in Table 6.3.1, but placed under scenario "2 correct" in Table 6.3.3. Additionally, Table 6.3.3 shows the percentage of times that each individual airport's prediction was wrong given that the Stacked Model predicted 0,1, or 2 of the individual airport configurations correctly.

Table 6.3.3 shows that the New York Metroplex Stacked Model very rarely predicted all three individual airports incorrectly during any given decision period. This fact is somewhat intuitive considering that the individual airport models all had a much higher accuracy than the Stacked Model. Additionally, it can be seen that the majority of the time (45.7%), the New York Metroplex Stacked Model predicts 2 of the 3 individual airport configurations correctly. This is promising because the overall dynamics of the New York Metroplex could potentially be characterized using

only 2 airports in future capacity models. In fact, the New York Metroplex Stacked Model predicts at least 2 individual airports correctly 73.7% of the time, making it competitive with the New York Metroplex Configuration Model in this sense. It should also be noted that Table 6.3.3 indicates that majority of incorrect predictions for all scenarios occur at JFK. This highlights some of the problems that were seen in the individual JFK model. If the accuracy of the JFK model is increased by adding more data periods or adding more attributes for prediction, it can be assumed that the accuracy of the New York Metroplex Stacked Model would also increase substantially.

Table 6.3.3: New York Metroplex Stacked Model statistics.

Scenario	Periods	% Periods	% Incorrect		
			EWR	JFK	LGA
0 Correct	368	3.2%	100.0%	100.0%	100.0%
1 Correct	2,662	23.1%	57.9%	81.9%	60.2%
2 Correct	5,261	45.7%	13.6%	59.7%	26.7%
3 Correct	3,230	28.0%	0.0%	0.0%	0.0%

Finally, it is relevant to examine if the New York Metroplex Stacked Model predicts correctly when the New York Metroplex Configuration model predicts incorrectly. It may be possible that each different model overcomes some of the shortcomings of the other model, and combining these two models in the future could be beneficial. Table 6.3.4 shows an overall confusion matrix for the New York Metroplex Configuration Model and Stacked Model. As shown, it appears that when the New York Metroplex Configuration Model is incorrect, the Stacked Model is also often incorrect. Furthermore, when the Stacked Model is incorrect, the Configuration Model is often correct. Going further, Figure 6.3.2 shows a confusion matrix between both the Stacked Model and Configuration Model predictions for each specific configuration on the data periods when the Configuration Model had correct predictions. Similarly, Figure 6.3.3 shows a confusion matrix between both the Stacked Model and Configuration Model predictions for each specific configuration on the data periods when the Stacked Model had correct predictions. These figures reinforce the statement above. When the Stacked Model is predicted correctly, the Configuration Model also predicts very accurately (except for the N-VMC-DP configuration). When the

Table 6.3.4: Confusion table comparing overall NY Metro. Configuration Model and NY Metro. Stacked Model.

		NY Metro. Config Model		
		Correct	*Incorrect*	***Total***
NY Metro. Stacked Model	*Correct*	3,818	568	4,386
	Incorrect	4,136	2,999	7,135
	Total	7,954	3,567	11,521

Configuration Model is predicted correctly, the Stacked Model is not nearly as accurate; usually predicting a configuration classified as "Other". This analysis seems to suggest the there would not be any major benefit from combining the Stacked Model and the Configuration Model when predicting the runway configuration of the New York Metroplex.

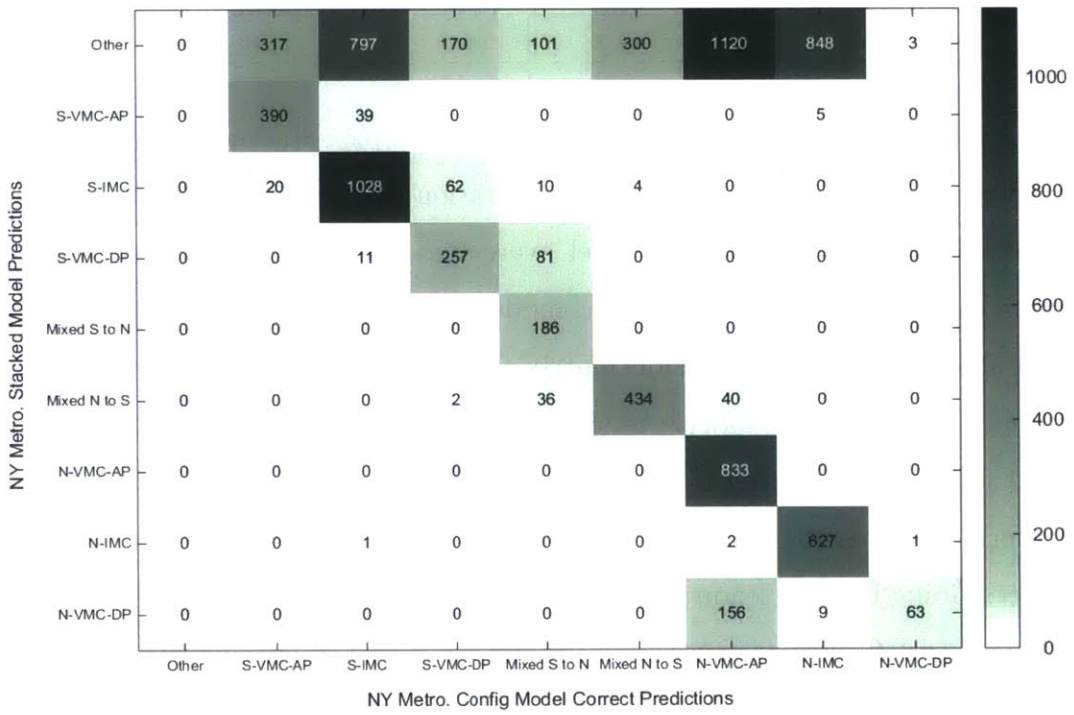

Figure 6.3.2: Comparison between NY Metro. Configuration Model and NY Metro. Stacked Model when Configuration Model model predictions are correct.

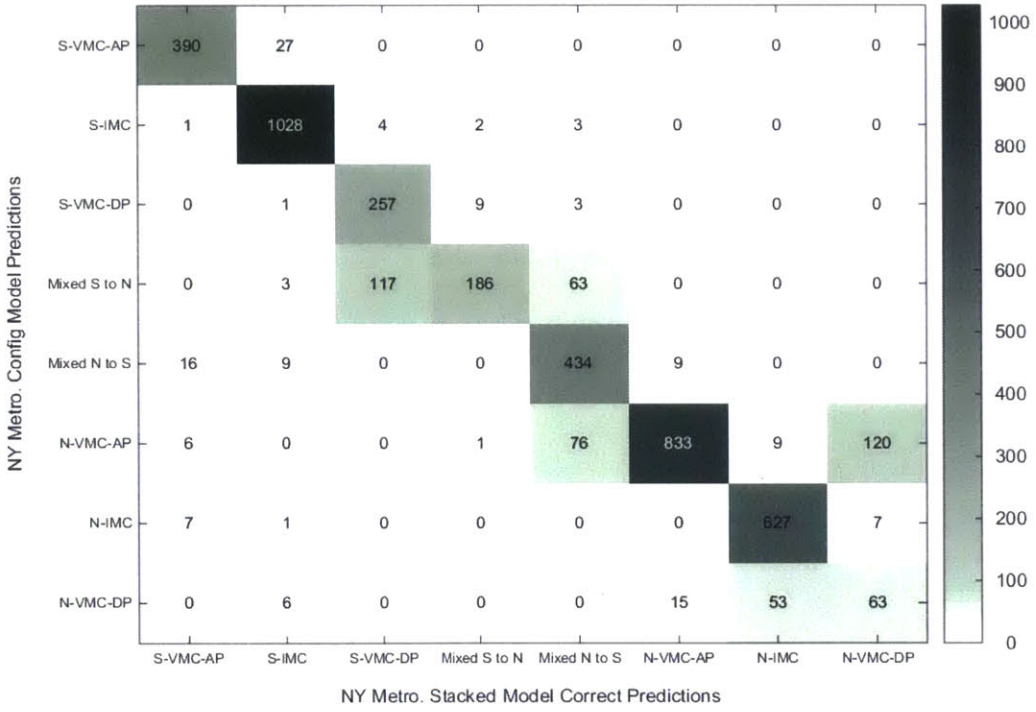

Figure 6.3.3: Comparison between NY Metro. Config Model and NY Metro. Stacked Model when Stacked Model predictions are correct.

Chapter 7

Conclusions

This paper develops models of the nominal runway configuration selection process of air traffic control personnel at Newark (EWR), John F. Kennedy (JFK), LaGuardia (LGA), and San-Francisco (SFO) airports using a discrete-choice framework with empirical observations and maximum likelihood estimation. The models for each airport are trained on year 2011 data from the ASPM database and are tested using year 2012 ASPM data. Utility functions for different runway configurations at each airport reflect the importance of various factors such as weather, wind speed, wind direction, airport demand, noise mitigation, inter-airport coordination, and the incumbent runway configuration. The weights assigned to the utilities are used to infer the relative importance of the different attributes. Across all models, the inertia variables were seen to have the highest importance when making the decision selection. Additionally, headwinds and tailwinds on the arrival runways were also found to be an important factor. For EWR, JFK, and LGA, high capacity configurations were favored under high arrival demand scenarios. At SFO, demand effects were coupled with visibility because of runway separation procedures. Switch proximity reflected a preference to reduce the operational effort required from a runway configuration switch.

The discrete-choice utility functions are then used within probabilistic models that predict the runway configuration selection on a specified forecast horizon. Predictions are first calculated assuming perfect knowledge of future weather conditions and airport demand. Assuming perfect knowledge, the EWR, JFK, LGA, and SFO models

achieve accuracies of 97.8%, 96.3%, 98.0%, and 98.1% respectively on a short-term 15 minute planning horizon. On a standard 3 hour planning horizon, the EWR, JFK, LGA, and SFO models achieve accuracies of 79.5%, 63.8%, 81.3%, and 82.8% respectively. Finally, on a longer-term 6 hour planning horizon, the EWR, JFK, LGA, and SFO models achieve accuracies of 69.5%, 48.2%, 72.4%, and 72.8% respectively.

To replicate the quality of predictions using information that is available to air traffic controllers in real-time practical applications, the prediction models are then calculated using forecast weather data and scheduled demand data. With forecast data, the EWR, LGA, and SFO models achieve accuracies of 78.9%, 78.9%, and 80.8% respectively on a 3 hour planning horizon. The results show that the accuracies of the models are not significantly degraded by a practical implementation. This is promising because future decision support tools would need to use forecast data in real-time.

After creating models for each airport in the New York Metroplex individually, two combined models of the overall New York Metroplex were created. Both models were predicted on overall combined metroplex configurations that grouped several specific configuration combinations together. The first model, denoted the New York Metroplex Configuration Model was developed by re-estimating a discrete choice model for the combined Metroplex configurations and achieved accuracies of 96.1%, 69.0%, 60.0% on a short-term 15 minute planning horizon, a 3 hour planning horizon, and a longer-term 6 hour planning horizon respectively. The second model, denoted the New York Metroplex Stacked Model was developed by combining the individual airport models and achieved much lower accuracies of 80.5%, 38.1%, 23.6% on a short-term 15 minute planning horizon, on a 3 hour planning horizon, and a longer-term 6 hour planning horizon respectively. A detailed analysis of the Stacked model suggested that it predicted at least 2 individual airports correctly 73.7% of the time and the JFK component of the model was the largest source of error. It also suggested that combining both the Configuration Model and the Stacked Model for prediction is likely not very beneficial.

7.1 Limitations of Approach

It is important to acknowledge the limitations that were seen in this study. Because the approach is inherently data-driven, the predictions are limited only the the runway configurations that have been observed in the past. The models should be re-estimated when there are major changes in the runway selection decision process, such as new procedures or capacity enhancements [37]. Similarly, runway configurations that are observed infrequently in the past will not have strong estimators and are therefore difficult to predict accurately in the future. For the same reasons, runway configurations that switch more frequently than others also have a high level of uncertainty and a lower prediction accuracy.

Additionally, the discrete-choice models do not account for variability among decision makers, who may have differing levels of experience, diverse concerns, and different rationales for selecting runway configurations. The proposed models capture nominal behavior which can cause instances of perceived reporting errors or prediction errors. In other words, the model assumes the presence of rational decision makers who share the same utility functions, even though this may not be the case.

Additionally, the recursive Bayesian approach to prediction inflates errors and biases present within the discrete-choice utility functions on longer time horizons. In many cases, the heavy influence of the inertia parameter biases the predictions to hold the incumbent configuration even when a switch is preferable. When the weather conditions do not heavily favor a configuration switch, these models tend to predict that the configuration will remain the same. As a result, the prediction accuracy is reduced in time periods close to a configuration switch [18].

7.2 Potential Extensions

Despite these limitations, the prediction performance of the proposed discrete choice models suggests that they are a promising approach to predict the runway configu-

ration a few hours ahead of time. Potential extensions of these models fall into two major categories: broad scope extensions and model specific extensions. The broad scope extension considers practically implementing these models for NAS-wide decision support. The model specific extensions focus on fine-tuning individual airport models for better estimators and more accurate predictions.

On a broad scope, estimated discrete-choice utility functions and their corresponding prediction models could be developed for all FAA core airports over multiple years of data. The beta weight parameters for each utility function can then be clustered for similarly defined variables such as inertia, wind, or demand and included in future NextGen decision support tools.

Model specific extensions could include improving the inertia term by limiting its effect as time progresses on long term prediction horizons. Biases within the estimated parameters could be reduced by estimating the utility parameters using a balanced training dataset. The effect of wind gust and aircraft size, which are currently ignored, can also be included in the utility models. Adding more sources of data could be particularly useful for the JFK model. The individual airport models could benefit from adding coordination variables between all other airports in their respective region. For example, the LGA model could include coordination variables between EWR and TEB in addition to the coordination parameters for JFK. This will likely increase the overall accuracy of the model, however, it should be noted that adding these variables could effectively overstate the accuracy of the model if they remove too much of the variability in the data set. Additionally, the New York Metroplex Models could be improved by including effects from TEB on the Metroplex. Furthermore, new models for SFO could be estimated by defining "side-by" and "staggered" classes with more rigor. Challenges will still be present with this type of modeling, because "side-by and "staggered operations are not reported in the ASPM data. The lack of any ground truth will make it difficult to evaluate the predictive performance of this type of model, however new insights could be gained from the estimated utility functions.

Bibliography

[1] E. Gilbo, "Airport capacity: representation, estimation, optimization," *IEEE Transactions on Control System Technology*, vol. 1, no. 3, pp. 144-153, 1993.

[2] D. Bertsimas and S. Stock Patterson, "The air traffic flow management problem with en-route capabilities," *Operations Research*, vol. 46, no. 3, pp.406-422, May-June 1998.

[3] G. Lulli and A. Odoni, "The European Air Traffic Flow Management Problem," *Transportation Science*, vol. 41, no.4, pp. 431-443, November 2007.

[4] Y. Jung, T. Hoang, J. Montoya, G. Gupta, W. Malik, L. Tobias, and H. Wang. "Performance evaluation of a surface traffic management tool for Dallas/Fort Worth International Airport," *9th USA/Europe Air Traffic Management RD Seminar*, 2011.

[5] J. Post, J. Gulding, K. Noonan, D. Murphy, J. Bonn, and M. Graham, "The Modernized National Airspace System Performance Analysis Capability (NASPAC)," in *26th International Congress of the Aeronautical Sciences (ICAS)*, 2008.

[6] Flight Transportation Associates, "Enhanced Preferential Runway Advisory System (ENPRAS)." [Online]. Available: http://www.ftausa.com/enpras.htm

[7] D. G. Southgate and S. J. Sedgwick, "Time stamped aircraft noise prediction replacing the 'average day' wit the 'composite year'," *INTER-NOISE*, December 2006.

[8] L. Li and J. P. Clarke, "A stochastic model of runway configuration planning," *AIAA Guidance, Navigation and Control Conference*, August 2010.

[9] C. A. Provan and S. C. Atkins, "Optimization models for strategic runway configuration management under weather uncertainty," in *AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, 2010.

[10] C. Weld, M. Duarte, and R. Kincaid, "A runway configuration management model with marginally decreasing transition capabilities," *Advances in Operations Research*, 2010.

[11] M. J. Frankovich, D. Bertsimas, and A. R. Odoni, "Optimal selection of airport runway configurations," *Operations Research*, vol. 59, no. 6, pp. 1407-1419, November-December 2011.

[12] R. Oseguera-Lohr, N. Phojanamongkolkij, G. Lohr, and J. W. Fenbert, "Benefits assessment of tactical runway configuration management tool," *AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, 2013.

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

[14] G. Buxi and M. Hansen, "Generating probabilistic capacity profiles from weather forecast: A design-of-experiment approach," *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, 2011.

[15] H. Hesselink and J. Nibourg, "Probabilistic 2-day forecast of runway use: Efficient and safe runway allocation based on weather forecast," *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM 2011)*, 2011.

[16] S. Houston and D. Murphy, "Predicting runway configurations at airports," *Transportation Research Board (TRB) Annual Meeting*, 2012, paper number 12-3682.

[17] V. Ramanujam and H. Balakrishnan, "Estimation of maximum-likelihood discrete-choice models of the runway configuration selection process," *American Control Conference*, 2011.

[18] V. Ramanujam and H. Balakrishnan, "Data-driven modeling of the airport configuration selection process," *IEEE Transactions on Human Machine Systems*, 2014.

[19] J. Quayson, D. Gafurson, J. Debrosse, and T. Japi, "The Port Authority of NY NJ, July 2015 Traffic Report", [Online]. Available: http://www.panynj.gov/airports/pdf-traffic/JUL2015EWR.pdf.

[20] "Top 25 U.S. Freight Gateways, Ranked by Value of Shipments: 2008", *Bureau of Transportation Statistics*, [Online]. Available: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications.

[21] "John F. Kennedy International Airport," *FAA Information Effective 20 August 2015*, [Online]. Available: http://www.airnav.com/airport/KJFK.

[22] "LaGuardia Airport: Facts and Information", *The Port Authority of New York & New Jersey*, [Online]. Available: http://www.panynj.gov/airports/lga-facts-info.html.

[23] "ACI Annual World Airport Traffic Report," *Airports Council International*, [Online]. Available: http://www.aci.aero/Data-Centre/Monthly-Traffic-Data/Passenger-Summary/Year-to-date.

[24] "NextGen: San Francisco International Airport", *Federal Aviation Administration*, [Online]. Available: https://www.faa.gov/nextgen/snapshots/airport/?locationId=46.

[25] "John F. Kennedy International (New York)", *Federal Aviation Administration*, [Online]. Available: https://www.faa.gov/airports/planningcapacity/profiles/media.

[26] M. Ben-Akiva and S. Lerman, "Discrete choice analysis: theory and application to travel demand". *MIT Press*, 1985.

[27] M. Bierlaire, "BIOGEME: a free package for the estimation of discrete choice models," *3rd Swiss Transport Research Conference*, 2003.

[28] "Aviation System Performance Metrics (ASPM) Database." *Federal Aviation Administration*, [Online]. Available: aspm.faa.gov.

[29] R. A. DeLaura, R. F. Ferris, F. M. Robasky, S. W. Troxel, and N. K. Underhill, "Initial assessment of wind forecasts for Airport Acceptance Rate (AAR) and Ground Delay Program (GDP) planning," Lincoln Laboratory, Tech. Rep. ATC-414, January 2014.

[30] "Airport capacity benchmark report." *Federal Aviation Administration*, 2004.

[31] "Airport noise and emissions regulations." *The Boeing Company*, [Online]. Available: www.boeing.com/commercial/noise/san-francisco.html

[32] P. Bonnefoy and R.J. Hansman. "Scalability of the air transportation system and development of multi-airport systems: A worldwide perspective." 2008.

[33] X. Ning, "Method for deriving multi-factor models for predicting airport delays," Ph.D. dissertation, George Mason University, December 2007.

[34] A. Donaldson and R.J. Hansman. "Improvement of Terminal Area Capacity In The New York Airspace." ICAT Report No. ICAT-2011-4. February 2011.

[35] J. Quayson, J Cuneo, and D. Mesa."The Port Authority of NY & NJ December 2012 Traffic Report." December 2012.

[36] "Port Authority Airports Hit Second-Highest Annual Passenger Totals In 2012, While Also Setting Records For travelers at JFK and International Fliers." *The Port Authority of New York & New Jersey.* Press Release Number: 24-2013. February 2013.

[37] "FAA JO 7110.308 CHG 3: 1.5 Nautical Mile dependent approaches to parallel runways spaced less than 2,500 feet apart." *Federal Aviation Administration*, 2004.